Corporation for
NATIONAL &
COMMUNITY
SERVICE ★★★

IMPACT EVALUATION

Minnesota Reading Corps

K-3 Program

MARCH 2014

Impact Evaluation of the Minnesota Reading Corps K-3 Program

March 2014

Authors

This study was conducted by researchers from NORC at the University of Chicago and TIES:

Carrie E. Markovitz, Ph.D., Principal Research Scientist, NORC at the University of Chicago
Marc W. Hernandez, Ph.D., Senior Research Scientist, NORC at the University of Chicago
Eric C. Hedberg, Ph.D., Senior Research Scientist, NORC at the University of Chicago
Benjamin Silberglitt, Ph.D., Director of Software Applications, TIES

This report represents the work and perspectives of the authors and is the product of professional research. It does not represent the position or opinions of CNCS, the federal government, or the reviewers.

Acknowledgements

Citation

Markovitz, C.; Hernandez, M.; Hedberg, E.; Silberglitt, B. (2014). *Impact Evaluation of the Minnesota Reading Corps K-3 Program*. NORC at the University of Chicago: Chicago, IL.

Table of Contents

List of Figures

List of Tables

Executive Summary

Minnesota Reading Corps (MRC) is the largest AmeriCorps State program in the country. The goal of MRC is to ensure that students become successful readers and meet reading proficiency targets by the end of the third grade. To meet this goal, the MRC program, and its host organization, ServeMinnesota Action Network, recruit, train, place and monitor AmeriCorps members to implement research-based literacy enrichment activities and interventions for at-risk Kindergarten through third grade (K-3) students and preschool children.

Starting in 2011, the Corporation for National and Community Service (CNCS) sponsored a randomized controlled trial (RCT) impact evaluation of over 1,300 K-3 students at 23 participating schools who were determined to be eligible for the MRC program during the 2012-2013 school year. The goal of the impact evaluation was to determine both the short- and long-term impacts of the MRC program on elementary students' literacy outcomes.

Key Study Findings

Kindergarten, first, and third grade students who received MRC tutoring achieved significantly higher literacy assessment scores than students who did not.

The magnitude of MRC tutoring effects differed by grade, with the largest effects found among the youngest students (i.e., Kindergarten and first grade students), and the smallest effects among the oldest students (i.e., third grade students). Significant effects were not found for second grade students. In later grades (second and third), when students begin the more complex task of reading connected text, the MRC program appears to take longer than a single semester to produce significant improvements in student literacy. However, additional non-experimental analyses suggest that, over a longer period of time, the MRC program changes second and third grade students' growth trajectories towards increasing their reading proficiency.

MRC tutoring resulted in statistically significant impacts across multiple racial groups. In Kindergarten and first grade tutoring was effective despite important risk factors, including Dual Language Learner status and Free and Reduced Price Lunch eligibility.

A statistically significant impact of MRC tutoring was detected among Kindergarten and first grade students despite gender, minority group status, Dual Language Learner (DLL) status, and Free or Reduced Price Lunch (FRPL) eligibility. For each of these characteristics, students who received MRC tutoring significantly outperformed control students who did not receive tutoring on grade-specific literacy assessments. Third grade White, native English speaking (i.e., non-DLL), and eligible for FRPL students on average produced positive significant differences between program and control group students, while a statistically significant finding was not found for third grade Black and Asian students and third grade DLL students.

The MRC program is replicable in multiple school settings using AmeriCorps members with varied backgrounds.

Student assessment scores did not vary by AmeriCorps member characteristics (i.e., gender, race, age, years of education, full/part time status, or prior education experience) nor by the specific school at which the tutoring occurred. These results support the conclusion that the MRC program is replicable in multiple school settings using AmeriCorps members with diverse backgrounds. Many MRC AmeriCorps members have no previous experience working in schools, with students, or in the domain of literacy. The lack of member-level and school-level effects on student outcomes validates MRC's approach to training, coaching and supervision, as well as their intentional recruitment of members with diverse backgrounds who do not necessarily have formal training or experience in education or literacy instruction.

About the Minnesota Reading Corps

The MRC program was started in 2003 to provide emergent literacy enrichment and tutoring to children in four preschool (PreK) Head Start programs. In 2005, MRC expanded its program to serve students in Kindergarten through third grade (K-3). The core activities of MRC, and its host organization, ServeMinnesota Action Network, are to recruit, train, place and monitor AmeriCorps members to implement research-based literacy interventions for at-risk K-3 students and preschool children.

AmeriCorps members in the MRC program serve in school-based settings to implement MRC literacy strategies and conduct interventions with students using a Response to Intervention (RtI) framework. The key aspects of the MRC RtI framework are:

- Clear literacy targets at each age level from PreK through grade 3
- Benchmark assessment three times a year to identify students eligible for one-on-one interventions
- Scientifically based interventions
- Frequent progress monitoring during intervention delivery
- High-quality training and coaching in program components, and literacy assessment and instruction

In the RtI framework, data play the key roles of screening students' eligibility for services and then monitoring students' progress towards achieving academic goals (i.e., benchmarks). The Minnesota Reading Corps screens students for program eligibility three times a year (i.e., Fall, Winter, Spring) with two sets of grade-specific, literacy-focused general outcome measures (i.e., IGDI for PreK and AIMSweb for K-3) that possess criterion-referenced grade- and content-specific performance benchmarks. Program staff use scores from these general outcome measures to categorize students into one of three possible tiers (i.e., proficiency levels): 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.   Minnesota Reading Corps Response to Intervention Tiers
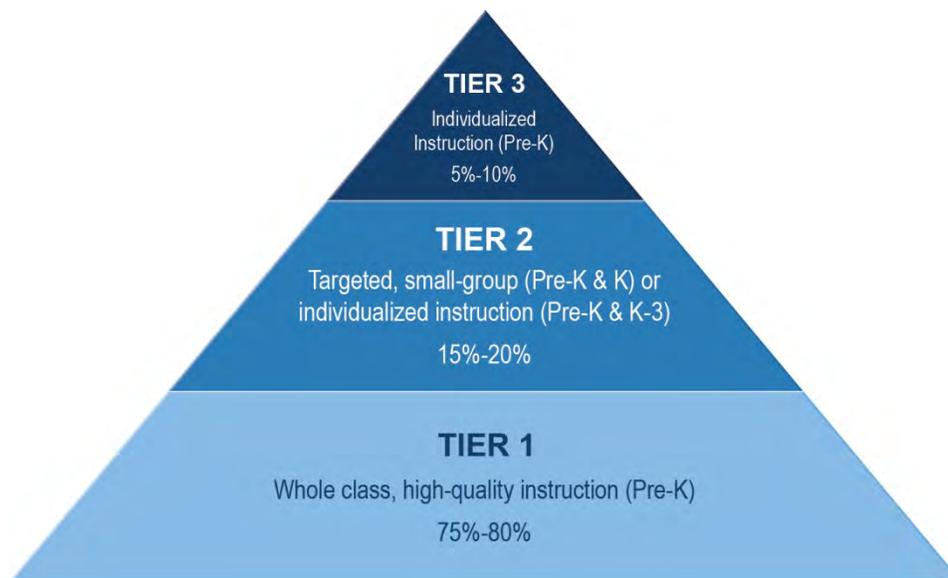
The MRC K-3 program provides one-on-one tutoring where members provide supplemental individualized literacy interventions to primarily Tier 2 students in Kindergarten through third grade. Members in the MRC PreK program provide whole-class literacy enrichment for all students (i.e., Tier 1) and a targeted one-on-one component, where members provide individualized interventions to students struggling with emergent literacy skills (i.e., Tiers 2 and 3).

At the K-3 level, which is the focus of this study, the program is focused on the "Big Five Ideas in Literacy" as identified by the National Reading Panel, including phonological awareness, phonics, fluency, vocabulary, and comprehension. AmeriCorps members serve as one-on-one tutors for Tier 2 students. Full-time members individually tutor approximately 15-18 K-3 students daily for 20 minutes each. The MRC tutoring interventions supplement the core reading instruction provided at each school. The goal of the tutoring is to raise individual students' literacy levels so that they are on track to meet or exceed the next program-specified literacy benchmark. One variation among K-3 members is the Kindergarten-Focus (K-Focus) position. K-Focus members continue to tutor first to third grade students, though they tend to spend a majority of their time providing Kindergarteners with two small-group (20-minute) sessions daily, for a total of 40 minutes of literacy-focused intervention.

Impact Evaluation Methodology

The K-3 impact evaluation is one of several complementary studies being completed on the MRC program: a process assessment of the MRC program in 20 PreK and K-3 sites (completed in Spring 2013);[1] a quasi-experimental impact evaluation of the MRC PreK program on preschool students' emergent literacy outcomes (forthcoming in Fall 2014); and a survey of AmeriCorps members (completed in Fall 2013). The impact evaluation focused on the following research questions:

1. What is the impact of the MRC program on student literacy outcomes?

    a. Does the impact vary by student characteristics/demographics?

    b. Do assessment scores vary by AmeriCorps member characteristics/demographics?

2. Does the impact of the program vary week to week? Does the number of weeks of intervention (i.e., dosage) impact student literacy outcomes?

3. Does participation in MRC have a longer-term impact on student literacy outcomes as measured at the end of the school year?

The methodology used was informed by the 2008-2009 and 2010-2011 annual Minnesota Reading Corps evaluations.[2]

School and Student Selection

A diverse and representative sample of 25 schools that had fully implemented the MRC K-3 program for at least two consecutive years were selected for the study. Due to the voluntary nature of the study, it was not possible to employ simple random sampling for the selection of schools; however, stratified random sampling was employed to select schools and school-level weights were used in our statistical models. Approximately 200 eligible schools were stratified by urbanicity (i.e., urban, suburban, and rural) using the MRC program regions and then selected using Probability Proportional to Size (PPS), whereby larger schools with a more pronounced need (defined as the number of students previously served by MRC) had a higher probability of selection. Using PPS ensured a statistically adequate sample size to conduct the K-3 impact evaluation. Participation in the evaluation was voluntary;

---

[1] Hafford, C., Markovitz, C., Hernandez, M, et al. (February 2013). *Process Assessment of the Minnesota Reading Corps Program.* (Prepared under contract to the Corporation for National and Community Service). Chicago, IL: NORC at the University of Chicago.

[2] Bollman, K. & Silberglitt, B. (2009). *Minnesota Reading Corps Final Evaluation 2008-2009.* Minneapolis, MN: MRC.; Bollman, K. & Silberglitt, B. (2011). *Minnesota Reading Corps Final Evaluation 2010-2011.* Minneapolis, MN: MRC.

consequently, 23 elementary schools agreed to participate in the K-3 impact evaluation during the 2012-2013 school year.[3]

All students at the 23 sampled schools identified by Fall benchmark scores as eligible for MRC services (i.e., Tier 2) were randomly assigned to either the MRC program (i.e., treatment) or control group at the beginning of the first semester prior to the start of tutoring. Each eligible student in each grade within a school was matched with another eligible student based upon their Fall benchmark score. Students within pairs were then randomly assigned to either the program or control condition. This matched pair design ensured that students in the program and control groups had similar Fall benchmark scores at the start of the school year. In the end, a total of 1,530 eligible students were selected to participate in the evaluation. During the school year, some students left the school area (i.e., moved) or were chronically absent and did not receive regular MRC tutoring or assessments. These students and their matched pair were removed from the analytic sample (i.e., pairwise deletion). Thus, the final sample of students included in the evaluation totaled 1,341 students.[4]

Table 1.   Demographic characteristics of students in the MRC K-3 Impact Evaluation (Fall 2012)

	Kindergarten (N=359)	1st Grade (N=409)	2nd Grade (N=265)	3rd Grade (N=308)
	Mean	Mean	Mean	Mean
Female	56%	52%	46%	46%
Race/Ethnicity				
White	30%	39%	36%	40%
Black	33	22	29	23
Asian	27	26	25	27
Hispanic	7	12	9	8
Other	3	1	1	2
Dual Language Learner (DLL)	26%	37%	37%	31%
Free and Reduced Price Lunch (FRPL)	76	72	75	71

Data Collection

In the Fall, Winter, and Spring of each school year, AmeriCorps members collect general outcome measure data using the AIMSweb literacy assessments. The AIMSweb assessments evaluate three critical literacy skills that research on literacy development has confirmed are appropriate for specific grade levels and seasons: 1) *letter sound fluency* (Kindergarten), 2) *nonsense word fluency* (first grade –Fall/Winter), and 3) *oral reading fluency* (first grade –Winter/Spring, second and third grades). These assessments are collectively called curriculum-based

---

[3] Schools who did not participate referenced scheduling conflicts and staffing shortages as possible reasons.

[4] For each grade, we demonstrate low levels of attrition assuming the "liberal" standard outlined in the *WWC Evidence Review Protocol for Early Childhood interventions (Version 2):* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/ece_protocol_v2.0.pdf

measures (CBM), because they correspond closely with curricular expectations for literacy skills at each developmental level. These data are used in the Fall benchmark data collection period to identify students who are in need of MRC support. Tier 2 students receive MRC intervention services until their progress monitoring data shows that they have achieved 3 to 5 consecutive data points above projected growth trajectory (i.e., the aimline) and two scores at or above the upcoming season benchmark target (Winter or Spring).

AmeriCorps members were asked to collect both benchmark and weekly progress monitoring data from students in both the program and control groups comprising the primary data for the evaluation. Because the evaluation was designed to measure the impact of MRC program participation relative to nonparticipation, students in the control group were embargoed from receiving tutoring services during the first semester of the school year. It was not possible to continue the experimental RCT throughout the entire school year due to school apprehension about withholding MRC services. As such, all program and control students who were eligible at the Winter benchmark to participate in the MRC program were allowed to receive services during the second semester of the 2012-2013 school year (Winter 2013 – Spring 2013). However, benchmark and weekly progress monitoring data continued to be collected by AmeriCorps members on all students in the study throughout the entire school year.

Analysis

Three specific and separate analysis approaches were used to address the three major research questions of the study:

1. To address the impact of the MRC program on student literacy outcomes (RQ1), a Fall-Winter Experimental Study analyzed 16 weeks of assessment data collected on the program and control groups during which the control group was embargoed from participation in the MRC program (i.e., first semester of the 2012-2013 school year from the September 2012 Fall benchmark through the January 2013 Winter benchmark). We also conducted analyses to examine whether differential effects of the program existed for specific subgroups of students based on the following student characteristics: gender (male/female), race (White, Black, Asian; White/non-White), Dual Language Learner (DLL) status (yes/no), and Free or Reduced Price Lunch eligibility (yes/no).

2. To understand how the pattern of program impacts vary week to week (RQ2), a Full Year Non-Experimental Study analyzed the full year of assessment data collected on the program and control groups during both semesters of the 2012-2013 school year. This non-experimental analysis included weekly assessment scores from the second semester, during which all students in the control group became eligible for MRC tutoring services.

3. To estimate whether participation in MRC has a longer-term impact on student literacy outcomes (RQ3), an Exploratory Analysis of the full year of assessment data focused on the longer-term effects of the program

on students' proficiency levels by examining all students who received MRC services at any point throughout the school year, despite initial assignment to program or control groups.

Findings and Conclusions

Below, the evaluation team offers our conclusions based on the study findings and organizes them by the three major research questions, followed by final thoughts on the implications of these findings for the future of the MRC program.

Research Question #1: What is the impact of the MRC program on student literacy outcomes?

The results of the Fall-Winter Experimental Study showed that Kindergarten, first and third grade students who received MRC tutoring achieved significantly higher literacy assessment scores by the end of the first semester than did control students who did not participate in MRC tutoring. The magnitude of MRC tutoring effects differed by grade, with the largest effects found among the youngest students (i.e., Kindergarten and first grade students), and the smallest effects among the oldest students (i.e., third grade students). Significant effects were not found for second grade students.

Figure 2.   Mean scores for Kindergarten program and control students
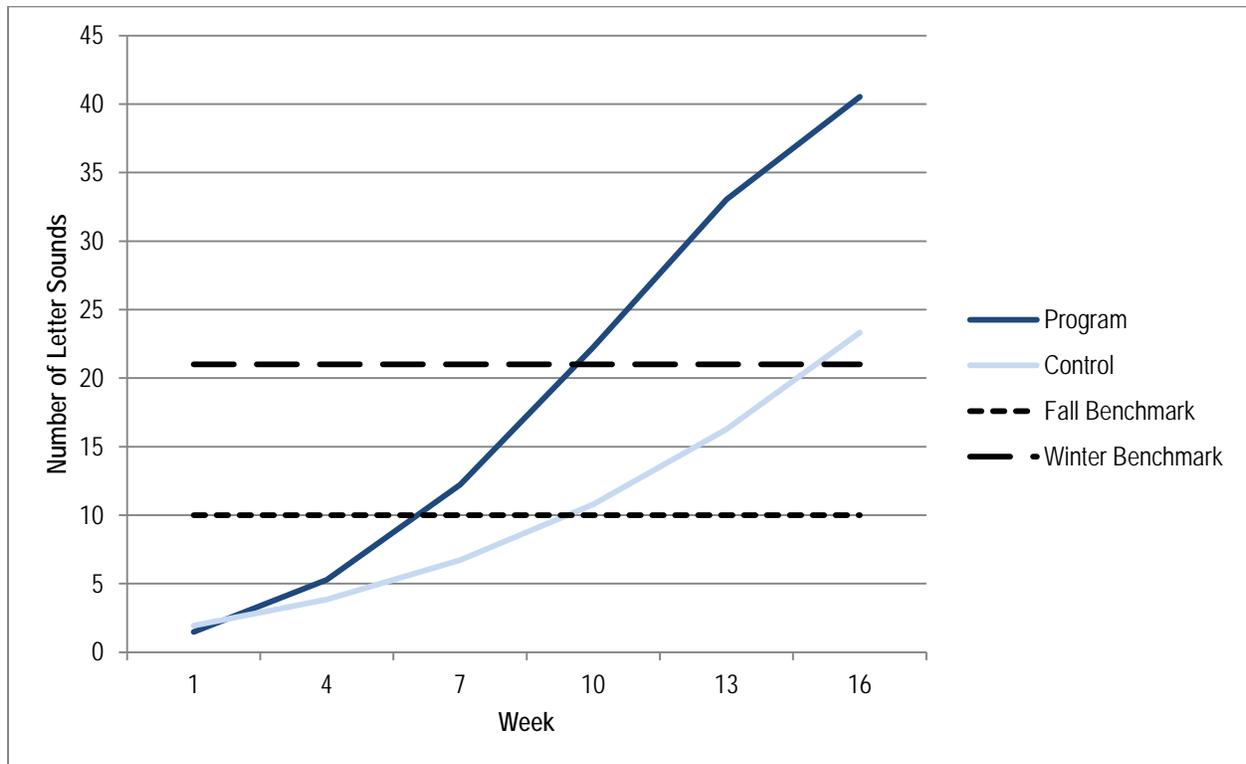
Kindergarten students who participated in the MRC program produced more than twice as many correct letter sounds by the end of the first semester than did students in the control condition. Similarly, first grade students participating in MRC tutoring demonstrated significantly higher letter sounds embedded within nonsense words than students in the control group. In contrast to the findings for Kindergarten and first grade students, the effect of the MRC program on *oral reading fluency* was significant but small for third grade students and not statistically significant for second grade students.

There are several possible explanations for the difference in program effects found in younger and older students. Younger students are more likely to qualify to receive MRC interventions due to a general lack of school readiness and insufficient exposure to academic language, books, and print. As students progress from Kindergarten into later grades, students are more likely to be eligible for MRC services because they are struggling to acquire or integrate needed skills, requiring significantly more in-depth intervention and time to remedy. When we consider that oral reading is a more challenging skill to acquire and that older children who are eligible for MRC services also are more likely to have experienced challenges mastering prerequisite skills, it is not surprising that it may take longer than a single semester for second and third grade students to accumulate substantial effects of the MRC program. Whereas lack of exposure in young children can be relatively quickly remedied by intensive and explicit instruction, learning challenges in older children can take longer to overcome.

1a. Does the impact vary by student characteristics/demographics?

A statistically significant impact of MRC tutoring was detected among Kindergarten and first grade students despite gender, minority status, DLL status, or FRPL eligibility. For each of these characteristics, students who received MRC tutoring significantly outperformed control students who did not receive tutoring on grade-specific literacy assessments. Among third grade students, an impact was not detected within all subgroups. Third grade White, native English speaking (i.e., non-DLL), and FRPL eligible students all produced significant differences between program and control group students. In contrast, among third grade Black and Asian students and third grade DLL students a statistically significant finding was not found between students who received MRC tutoring and control group students who did not participate in the program.

1b. Do assessment scores vary by AmeriCorps member characteristics/demographics?

Assessment scores did not vary by AmeriCorps member characteristics (i.e., gender, race, age, education, full/part time status, or years of education) nor by the specific school at which the tutoring occurred. These results support the conclusion that the MRC program is replicable in a variety of school settings using AmeriCorps members with diverse backgrounds. Many MRC members have no previous experience working in schools, with students, or in the domain

of literacy. In the *Process Assessment of the Minnesota Reading Corps,*[5] we concluded that the MRC program's high-quality training regime, research-based scripted interventions, regular objective assessment, ongoing on-site coaching, and multi-layered supervisory structure resulted in high levels of fidelity of program implementation and positive impacts on student literacy outcomes. The results of the Fall-Winter Experimental Study showed quantitatively that these critical program supports indeed reduced variability in the interventions delivered by AmeriCorps members within diverse school settings, such that the impact of member characteristics and individual school effects on K-3 students was minimized.

Research Question #2: Does the impact of the program vary week to week?

While the Fall-Winter Experimental Study examined the overall impact of the program during the first semester, the Full Year Non-Experimental Study examined estimates of week over week impacts to identify patterns in student growth by grade. After the Fall semester, all students in the control group became eligible for MRC tutoring services, were reassessed, and, if found eligible, could begin receiving MRC services in the second semester. Although the experimental portion of the evaluation ended after the Winter benchmark (i.e., first semester), weekly assessment data continued to be collected for the remainder of the school year from all students initially assigned in the Fall to the program and control groups. The full-year data was then used to estimate week over week growth in literacy outcomes for the average student who received tutoring during the 2012-2013 school year.

The analysis demonstrates that patterns in week over week gains among students receiving MRC tutoring vary by grade. Kindergarten students showed immediate and large gains, the largest of which occurred in the first few weeks of tutoring. In contrast, first, second and third grade students showed small, but steady week over week gains throughout the entire period of analysis. An important consideration when interpreting these findings is that roughly half of the schools in our study sample had K-Focus AmeriCorps members. While the typical MRC K-3 program provides students with one 20-minute session per day, in the K-Focus program, each Kindergarten student participates in two (20-minute) sessions daily, for a total of 40 minutes of literacy-focused instruction. Therefore, it is reasonable to consider that the more intensive, higher dosage intervention for Kindergarten students in some schools may have contributed to producing the large and early effects we observed in the week over week findings.

In contrast to the findings for Kindergarten students, the pattern of gains among first, second and third grade students continued to build throughout the study period. These findings are not unexpected, given older program eligible students are likely to need more intensive literacy intervention and practice, while younger students can benefit from increased exposure to literacy activities and time on task.

---

[5] Hafford, C., Markovitz, C., Hernandez, M, et al. (February 2013). *Process Assessment of the Minnesota Reading Corps Program.* (Prepared under contract to the Corporation for National and Community Service). Chicago, IL: NORC at the University of Chicago.

The findings for both second and third grade students indicate that the effects found in the Fall-Winter Experimental Study may have been more substantial if it had been possible to follow students for a longer period of time beyond 16 weeks. Thus, if the timeline for the experimental study could have been lengthened to allow observation of differences in scores between the program and control group of students over an entire school year, our finding for second and third grade students may have been more substantial.

Figure 3. Cumulative week over week growth in third grade words read aloud for students receiving MRC tutoring
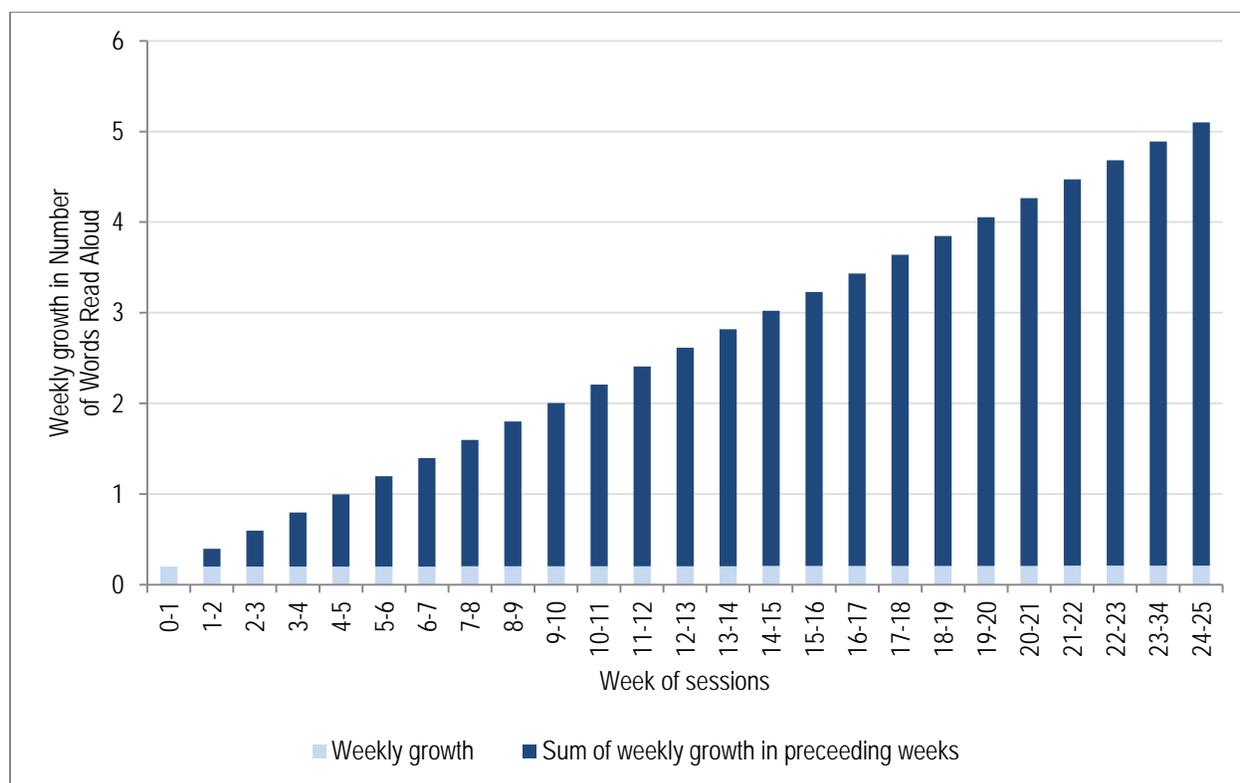
Research Question #3: Does participation in MRC have a longer-term impact on student literacy outcomes as measured at the end of the school year?

In the Exploratory Analysis, the evaluation team found evidence that participation in the MRC program results in longer-term effects on literacy outcomes when interventions begin earlier in the school year. When MRC interventions are implemented later in the school year (i.e., second semester), the probability of progressing and staying above benchmark decreases, while the likelihood of remaining chronically behind increases substantially. The findings showed that program group students who received tutoring assistance early in the school year have more than twice the likelihood of remaining above benchmark for the remainder of the school year compared to students assigned to the control group who received equal amounts of tutoring assistance, but later in the school year. The higher likelihood for program group students indicates that early intervention by the MRC program,

controlling for dosage, has a greater impact on students' longer-term literacy proficiency outcomes. Thus, a key conclusion from our analysis is that early intervention from the MRC program (i.e., in the first semester) for struggling students results in a higher likelihood of positive longer-term outcomes.

Final Thoughts

The Minnesota Reading Corps program is contributing to our nationwide goal of improving 3rd grade reading proficiency.

In sum, the results of the Fall-Winter Experimental Study suggest that the MRC program produces the largest effects most quickly with the youngest students, particularly Kindergarten students. The intensive one-on-one exposure to MRC tutoring produces large increases in young students' *letter sound fluency.* In later grades (i.e., second and third), when students begin the more complex task of reading connected text, the MRC program may take longer to produce larger effects in *oral reading fluency.* While it was not possible to *experimentally* examine the full-year impact of the program on student outcomes in this study, the non-experimental analyses suggest that over the course of a longer period of time, the MRC program could produce larger improvements in second and third grade students' *oral reading fluency.*

One of the most critical findings for program replication is MRC's successful deployment of AmeriCorps members lacking any specialized background in education or literacy. The results of the member analysis revealed no significant differences in student impacts due to the characteristics of the members providing the tutoring. The lack of member effects suggests that if similar program-based infrastructure and resources are provided and specialized interventions are accurately implemented and closely monitored, members with diverse backgrounds can serve without possessing any specialized prerequisite technical skills. The combination of MRC program elements that resulted in positive impacts on student literacy outcomes can be considered an effective model for the development of other successful reading intervention programs for K-3 students.

Given the smaller impacts found among older students, future research may wish to examine the impact of specific MRC interventions on *oral reading fluency,* as well as the number and timing of changes in the use of these interventions, to more fully explore which interventions may be more effective with older students. Additionally, it may be of interest to follow these same randomized students through later school years to assess the potential long-term effects of the MRC program on students' performance on future benchmark assessments, meeting or exceeding grade level proficiency on state literacy assessments, graduation rates, and other more distal educational and economic outcomes.

I. Introduction

Minnesota Reading Corps (MRC) is a statewide initiative with a mission to help every Minnesota child become a proficient reader by the end of third grade. MRC engages a diverse group of AmeriCorps members to provide literacy enrichment and tutoring services to at-risk Kindergarten through third grade (K-3) elementary school students and preschool children (PreK). As of the 2012-2013 school year, more than 1,100 AmeriCorps members implemented the MRC program in 652 schools or sites.[6] and 184 school districts across the state of Minnesota..[7]

This report, funded by the Corporation for National and Community Service (CNCS), describes the findings from a randomized controlled trial (RCT) impact evaluation of over 1,300 K-3 students who were determined to be eligible for the MRC program during the 2012-2013 school year. The students were enrolled at a representative sample of 23 schools that were experienced implementers of MRC programs. The goal of the impact evaluation was to determine both the short- and long-term impacts of the MRC program on elementary students' literacy outcomes. The K-3 impact evaluation is one of several complementary studies being completed on the MRC program: a process assessment of the MRC program in 20 PreK and K-3 sites (completed in Spring 2013);[8] a quasi-experimental impact evaluation of the MRC PreK program on preschool students' emergent literacy outcomes (forthcoming in Fall 2014); and a survey of AmeriCorps members (Fall 2013). The impact evaluation focused on the following research questions:

1. What is the impact of the MRC program on student literacy outcomes?
   a. Does the impact vary by student characteristics/demographics?
   b. Do assessment scores vary by AmeriCorps member characteristics/demographics?

2. Does the impact of the program vary week to week? Does the number of weeks of intervention (i.e., dosage) impact student literacy outcomes?

3. Does participation in MRC have a longer-term impact on student literacy outcomes as measured at the end of the school year?

To address these questions, we begin in Chapter II by presenting a brief overview of the MRC program and its role in the recruitment, training, placement and monitoring of AmeriCorps members as they implement the program in

---

[6] According to the Minnesota Department of Education (MDE), during the 2011-2012 school year, 942 public schools served grades K-12. Of those schools, 912 offered PreK services. The total number of preschools in the state of Minnesota (i.e., public schools and non-public schools) was not available. http://w20.education.state.mn.us/MDEAnalytics/Summary.jsp

[7] According to MDE, during the 2011-2012 school year, there were 333 public operating elementary & secondary independent school districts, 3 intermediate school districts, and 148 charter schools (which are considered public school districts in Minnesota).

[8] Hafford, C., Markovitz, C., Hernandez, M, et al. (February 2013). *Process Assessment of the Minnesota Reading Corps Program.* (Prepared under contract to the Corporation for National and Community Service). Chicago, IL: NORC at the University of Chicago.

preschool and elementary school settings. We then describe the K-3 component of the MRC program, which is the focus of this evaluation, MRC's multi-layered supervisory structure, and their Summer Training Institute. Chapter III then provides information on the impact evaluation's methodology for selecting sites and students, randomization procedures for forming program and control groups for comparisons, collection and use of program data, and analysis of findings.

This background information sets the context for the presentation of findings from the three analyses of assessment data from the K-3 program in Chapters IV, V, and VI. Chapter IV provides the findings from the examination of the shorter-term impacts of the program in our Fall-Winter Experimental Study. These findings are based on a comparison of 16 weeks of data from the RCT program and control groups, during which control group students did not receive any MRC services. Our examination of the program impacts also includes analysis of key subgroups, including gender, race, Dual Language Learner (DLL) status, and Free and Reduced Price Lunch (FRPL) status. In this chapter, we also examined whether impacts vary due to the characteristics of the AmeriCorps members who conducted the tutoring or the schools where the tutoring took place.

In addition to the presentation of findings on the RCT results in Chapter IV, we provide in **Appendix C** a separate analysis of the data tailored to the requirements of the U.S. Department of Education's Institute of Education Sciences' What Works Clearinghouse (WWC). We share the WWC's goal to provide educators with the information they need to make evidence-based decisions. Therefore, we have developed this appendix to specifically demonstrate that our study meets WWC's rigorous standards.

Chapter V provides the findings of results from the entire year of program data. This analysis includes data from all students, including the control group students who were eligible to receive MRC services in the second semester of the school year. For this Full Year Non-Experimental Study, we examined the effect of the program for each week of tutoring and the cumulative effect over the school year both for the program group, which began receiving tutoring earlier in the school year, and the control group, which included many students who received tutoring in the second half of the school year.

The results presented in Chapter VI focus on the longer-term effects of the MRC program on students' literacy outcomes. Using two different exploratory analysis approaches, the evaluation team attempted to estimate the longer-term probability of successfully maintaining proficiency after exiting the MRC program. For these analyses, we used the entire year of weekly assessment data collected from both the program group, all of whom received tutoring, and those control group students who received tutoring later in the school year.

We conclude our report in Chapter VII by returning to the research questions. The evaluation team addresses whether the MRC K-3 program appears to have an impact on students' literacy proficiency and whether there are

differential effects by grade, gender, race, DLL status, and/or FRPL status based on the findings from the Fall-Winter Experimental Study. We also draw on the findings from the Full Year Non-Experimental Study to answer the evaluation's other key research questions on the week over week effects of the program and its longer-term impact on students' literacy outcomes. Finally, we discuss the implications of the findings for the MRC program. A glossary of terms to assist the reader is provided in Appendix E.

II. About Minnesota Reading Corps

A. Statewide Implementation of MRC: 2003-2013

Minnesota Reading Corps (MRC) is the largest AmeriCorps State program in the country. The goal of MRC is to ensure that students become successful readers and meet reading proficiency targets by the end of the third grade. The MRC program was started in 2003 to provide reading and literacy tutoring to children in four preschool (PreK) Head Start programs. In 2005, MRC expanded its program to serve students in Kindergarten through third grade (K-3). The core activities of MRC, and its host organization, ServeMinnesota Action Network, are to recruit, train, place and monitor AmeriCorps members to implement research-based literacy interventions for at-risk K-3 students and preschool children.

Minnesota Reading Corps is a strategic initiative of ServeMinnesota. ServeMinnesota is the state commission for all AmeriCorps State programs in Minnesota, including the Minnesota Reading Corps, and helps leverage the federal, state and private dollars to operate MRC. As a catalyst for positive social change and community service, ServeMinnesota works with AmeriCorps members and community partners to meet critical needs in Minnesota. As a nonprofit organization, it supports thousands of individuals to improve the lives of Minnesotans by offering life-changing service opportunities that focus on education, affordable housing, employment, and the environment. The ServeMinnesota Action Network serves as fiscal host to provide statewide management and oversight for the MRC program. The Action Network is a nonprofit organization and serves as a home to incubate, replicate and scale evidence-based AmeriCorps programs that address critical state priorities. In addition, the Saint Croix River Education District (SCRED) and TIES have been funded by ServeMinnesota to conduct an annual evaluation of the MRC program..[9]

AmeriCorps members in the MRC program serve in school-based settings to implement MRC literacy strategies and conduct interventions with students. MRC members serve as AmeriCorps members, bound to the program's call to service. As a direct service program, MRC engages its members in service to work towards the solution of a social issue. In exchange for their service of 1700 hours a year (full-time) or 900 hours a year (part-time), members receive benefits that include a bi-weekly stipend, student loan forbearance, and an education stipend for the first two years of service.

In addition to AmeriCorps members serving in the schools, the MRC model provides supports for maintaining the fidelity of the intervention through the assignment of one or more Internal Coaches at each site or school to mentor

---

[9] ServeMinnesota 2011, Background document.

and guide members. Internal Coaches are typically specialists, teachers, or curriculum directors employed by the site or school. Expert-level Master Coaches are also assigned to each Internal Coach to provide consultation on literacy interventions and assessment, as well as ensure fidelity to the MRC model. The MRC Program Coordinators provide administrative support to individual sites (Principals, Internal Coaches, and Master Coaches) and assist members with their AmeriCorps responsibilities.

In the 2012-13 school year, the MRC program's more than 1,100 AmeriCorps members served over 30,000 students in 652 elementary schools, Head Start centers, and preschools, making it the largest AmeriCorps programs in the country. Based on the early success of the MRC program, replication is underway in Colorado, Massachusetts, Michigan, Santa Cruz County, CA, Washington DC, Virginia, Iowa, and North Dakota.

B. Foundational Framework and Staffing Structure in MRC

The MRC program utilizes a Response to Intervention (RtI) framework. The RtI model is based on a problem solving approach which was incorporated into the 2004 Individuals with Disabilities Education Act (IDEA) and has been gaining popularity among educators, policymakers, administrators, teachers, and researchers. The key aspects of the MRC RtI framework are:

- Clear literacy targets at each age level from PreK through grade 3
- Benchmark assessment three times a year to identify students eligible for one-on-one interventions
- Scientifically based interventions
- Frequent progress monitoring (formative assessment) during intervention delivery
- High-quality training and coaching in program components, and literacy assessment and instruction

In the RtI framework, data play the key roles of screening students' eligibility for additional services and then monitoring students' progress towards achieving academic goals (i.e., benchmarks). The Minnesota Reading Corps screens students for program eligibility three times a year (i.e., Fall, Winter, Spring) with two sets of grade-specific, literacy-focused general outcome measures (i.e., IGDI for PreK and AIMSweb for K-3) that possess criterion-referenced grade- and content-specific performance benchmarks. Program staff use scores from these general outcome measures to categorize students into one of three possible tiers (i.e., proficiency levels; see **Figure II.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 II.1. Minnesota Reading Corps Response to Intervention Tiers
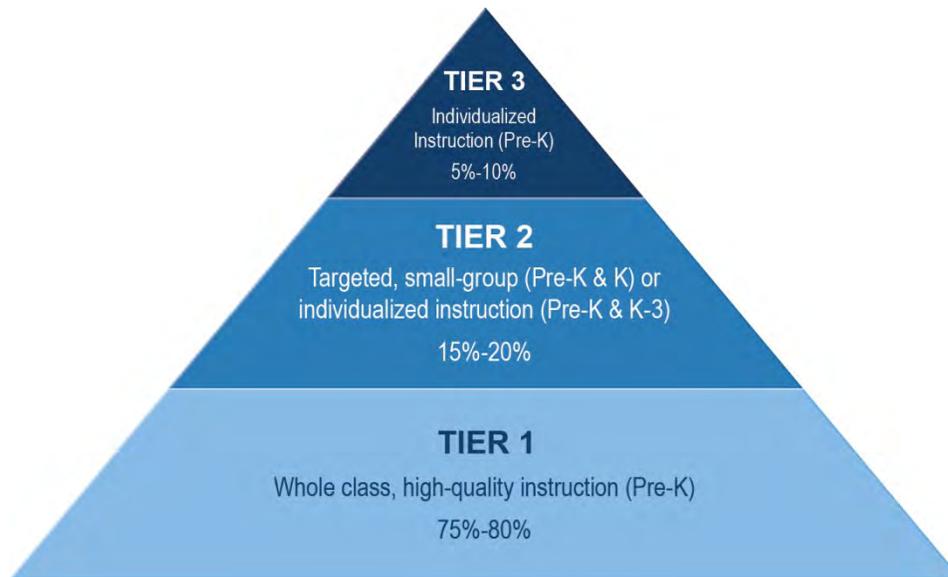
TIER 3
Individualized
Instruction (Pre-K)
5%-10%

TIER 2
Targeted, small-group (Pre-K & K) or
individualized instruction (Pre-K & K-3)
15%-20%

TIER 1
Whole class, high-quality instruction (Pre-K)
75%-80%

The MRC K-3 program provides one-on-one tutoring where members provide supplemental individualized literacy interventions to primarily Tier 2 students in Kindergarten through third grade. Generally, those Tier 2 students who score closest to the benchmark are offered MRC's intervention services first because they should require the least amount of intervention (i.e., time in program) to be set on a learning trajectory to achieve grade level proficiency. The students closest to the benchmark can be moved through the program more quickly than those students with greater need, allowing the schools to maximize support for students needing more intensive services. The MRC PreK program includes both an immersive "push-in" component, where members provide whole-class literacy enrichment for all students (i.e., Tier 1), and a targeted one-on-one component, where members provide individualized interventions to students struggling with emergent literacy skills (i.e., Tiers 2 and 3). Although the MRC program provides both PreK and K-3 interventions to students, the focus of this evaluation is on the MRC K-3 program. Therefore, the remainder of this report will focus on describing the K-3 program and evaluation. As previously mentioned, the findings from a quasi-experimental design (QED) evaluation of the PreK MRC program will be available in Fall 2014.

Overview of K-3 Program Literacy Focus and AmeriCorps Members' Role

At the K-3 level, the program is focused on the "Big Five Ideas in Literacy" as identified by the National Reading Panel, including phonological awareness, phonics, fluency, vocabulary, and comprehension. AmeriCorps members serve as one-on-one tutors and enact research-based interventions with students below grade-specific literacy benchmarks (i.e., Tier 2 students). Full-time members individually tutor approximately 15-18 K-3 students daily for 20 minutes each. The literacy interventions consist of a set of prescribed, research-validated activities such as

"Repeated Reading with Comprehension Strategy Practice" or "Duet Reading." The decision to change a student's interventions is based upon reviewing weekly progress monitoring data The tutoring interventions are supplemental to the core reading instruction provided at each school. The goal of the tutoring is to raise individual students' literacy levels so that they are on track to meet or exceed the next program-specified literacy benchmark. Meeting benchmark will allow the student to benefit fully from general (i.e., Tier 1) literacy instruction already provided in the classroom.

One variation among K-3 members is the Kindergarten-Focus (K-Focus) position. K-Focus members continue to tutor students in grades Kindergarten through third; however, they tend to spend a majority of their time providing Kindergarten students with a daily "double-dose" of MRC interventions. In the K-Focus program, each Kindergarten student participates in two (20-minute) sessions daily, for a total of 40 minutes. One session is a 5-day Repeated Read Aloud intervention that is conducted in a small group setting (typically four students) that includes dialogic reading to focus on phonemic awareness, phonics, and vocabulary instruction. The other session is a standard MRC early literacy intervention that is selected by the Internal Coach based on student needs (phoneme blending, phoneme segmenting, letter sounds or word blending) and is conducted in pairs of students.

Supervisory Staff

The Internal Coaches and Master Coaches play important roles in MRC program implementation (see **Figure II.2** for an illustration of the complete MRC supervisory structure). The Internal Coach is a school employee who is trained to provide on-site literacy support and oversight to AmeriCorps members serving as literacy tutors at the site. In order to ensure fidelity to the MRC model, the Internal Coach conducts monthly integrity checks for each intervention and scores the member using the Accuracy of Implementation Rating Scale (AIRS) before each benchmarking period. The Internal Coach provides the member with feedback based on these observations. The Internal Coach also ensures that the member is accurately reporting student data in AIMSweb and OnCorps. Throughout the school year, the Internal Coach works with assistance from the Master Coach to select appropriate interventions for each student and to determine if students are ready to exit the program. The Internal Coach also works closely with MRC program staff and school administration to address any concerns about member performance and to address disciplinary action if necessary. MRC estimates that the time commitment for Internal Coaches is 6-9 hours per member per month. The additional time commitment for required training is 32 hours for new K-3 Internal coaches and 16 hours for returning K-3 Internal coaches.

Figure II.2.   MRC Supervisory Structure
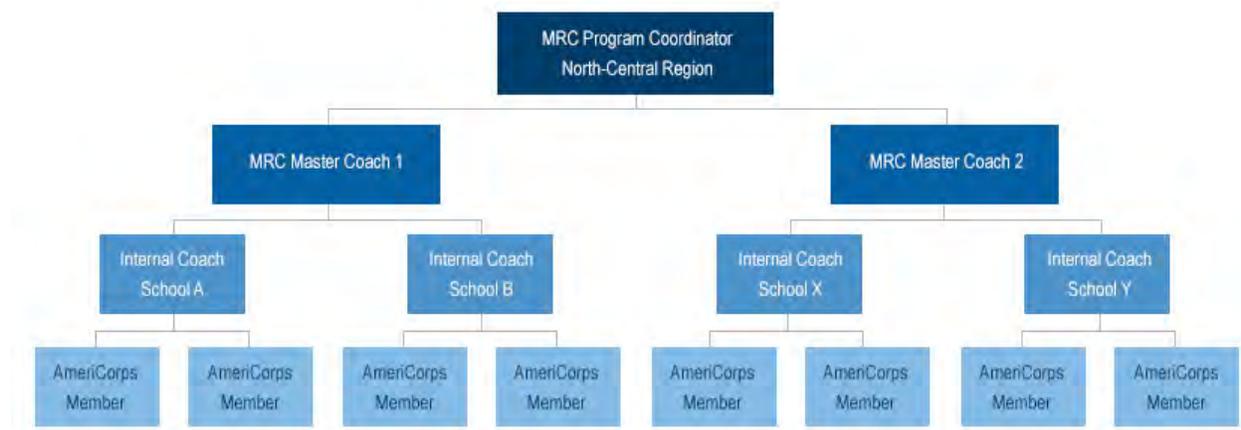
The Master Coach is a literacy expert employed by MRC who serves as a literacy consultant to the Internal Coach and member(s). The Master Coach supports the Internal Coach and the member in making decisions about student eligibility and instruction by reviewing benchmark data. The Master Coach also helps to ensure fidelity to the MRC model. The Master Coach visits schools at different frequencies throughout the year depending on the schools' degree of experience implementing MRC, ranging from once a month for schools that have recently implemented MRC to three times a year for schools where MRC is well-established. Visits last approximately one hour per member, during which the Master Coach, Internal Coach and member(s) discuss students' assessment data, progress towards achieving benchmark goals, and implementation challenges.

Other Master Coach responsibilities include communicating with the Internal Coach and member(s) about preparing for benchmarking; performing member fidelity checks along with the Internal Coach to ensure appropriate administration of benchmark assessments and interventions; providing consultation as needed regarding the identification and prioritization of students to receive MRC tutoring; reviewing student progress monitoring graphs; and providing program updates to the Internal Coach and member. If the Internal Coach cannot answer a member's question, the Master Coach can often provide advice. The Master Coach can also answer questions about topics such as AIMSweb or scheduling.

For administrative issues, such as questions about training schedules and timesheets, the Internal Coach or member can contact their MRC Program Coordinator. The Program Coordinator also helps members answer questions about their community service requirement and requested leaves of absence. Program Coordinators also are to be notified about all member disciplinary issues.

C. Summer Institute Training

Each summer, the Minnesota Reading Corps hosts a multi-day Summer Institute for training returning and new Master Coaches, Internal Coaches, and AmeriCorps members.[10] ServeMinnesota and MRC staff orchestrates the organizational and administrative aspects of the Summer Institute, while Minnesota literacy experts conduct training sessions. This intensive, information-filled conference provides expert training in the research-based literacy interventions employed by MRC. In its most basic form, the Summer Institute is a learning forum for literacy interventions and teaching techniques. However, the Summer Institute also serves an important role in developing member, coach, and eventually, school adherence to the MRC model. Speeches from former and current members, funders, parents, and officials from the Minnesota Department of Education and local school districts encourage this process and enhance the inspirational atmosphere of the training sessions. At the Summer Institute, the members also meet with their Internal Coach, and sometimes Master Coach, with whom they will be working throughout the upcoming school year.

During several intensive sessions at the Summer Institute, members learn the essential skills, knowledge, and tools needed to serve as effective literacy tutors. These sessions introduce members to the MRC program model, the interventions that constitute the instructional core of the program, as well as the underlying research and theories supporting the interventions and program model. Importantly, members are provided with detailed *Literacy Handbooks* to serve as a resource for supporting program implementation. The handbooks provide an introduction to the MRC program, information on policies and procedures and service requirements, procedures for the benchmarking and progress monitoring of students, and specific direction and materials for conducting MRC strategies and interventions. In addition, members are provided with online resources that mirror the contents of the *Literacy Handbook* and supplement it with other resources such as videos of model interventions and best practices. Both the *Handbook* and website are intended to provide members with just-in-time support, as well as opportunities for continued professional development and skill refinement.

At the Summer Institute, K-3 AmeriCorps members are trained to provide the MRC research-based, reading interventions that help K–3 students reach grade-level literacy benchmarks. K-3 members are trained how to implement the majority of instructional interventions during the Summer Institute. However, members also participate in two additional trainings early in the fall where they learn to use the assessment tool, AIMSweb, and Great Leaps, a comprehensive intervention for struggling readers that focuses on sound awareness (phonological/ phonemic awareness), letter recognition and phonics, high frequency sight words and phrases, and stories for oral reading.

---

[10] Members attend all four days of the Summer Institute (one day orientation and three days of training). New Coaches attend three days, and returning Coaches attend one day.

In addition to member training, at the Summer Institute each Internal Coach receives a comprehensive orientation to MRC, including program and early literacy background, intervention delivery, benchmarking and progress monitoring. At their training sessions, Internal Coaches also receive information about their roles, responsibilities and expectations while serving in the program. The Internal Coaches are instructed in their responsibilities, including ensuring fidelity to the MRC model, orienting the member to the school, introducing school staff to the member, setting the tutoring schedule and coordinating school-based professional development opportunities for their members. Internal Coaches also are oriented to the layers of support provided by MRC, including the Master Coach and Program Coordinator.

D. The Role of Data in MRC Program Implementation and Improvement

In the Fall, Winter, and Spring of each school year, AmeriCorps members collect general outcome measure data using the AIMSweb literacy assessments. The AIMSweb assessments evaluate three critical literacy skills that research of literacy development has confirmed are appropriate for specific grade levels and seasons: 1) letter sound fluency (Kindergarten), 2) nonsense word fluency (first grade –Fall/Winter), and 3) oral reading fluency (first grade – Winter/Spring, second and third grades). These assessments are collectively called curriculum-based measures (CBM), because they correspond closely with curricular expectations for literacy skills at each developmental level. For example, literacy expectations for Kindergarten students focus on learning the phonetic relationships that exist between letters and sounds. As students progress through elementary school, they learn more complex (i.e. word-level) relationships between letters and sounds, and over time they are expected to demonstrate fluent reading of connected text. These expectations are broadly accepted by literacy experts and are reflected in such documents as the Common Core State Standards. The CBM measures used by MRC were developed to assess each of these skills, and extensive research has shown them to be sufficiently reliable and valid for making decisions within an RTI framework (see **Appendix C** for a description of the psychometric properties of outcome measures) .

Table II.1 lists the specific AIMSweb CBM assessments and corresponding benchmark scores used to identify program eligible K-3 students by grade and season. These benchmark scores correspond to empirically-derived target scores that, through large-scale research studies, indicate the level of literacy skill that needs to be demonstrated in a certain grade at a certain time period to have a 90% chance of passing a high-stakes state reading assessment in third grade. For example, in order for a second grade student to have at least a 90% chance of demonstrating proficiency on their future third grade reading proficiency assessment, they need to have a score of at least 42 on the Fall benchmark and at least 73 on the Winter benchmark in the oral reading fluency CBM.

These data are used in the Fall benchmark data collection period to identify students who are in need of MRC support. Given the sometimes large student to AmeriCorps member ratio at participating schools, the Internal Coach

typically prioritizes which students the members will assess using existing school data. Generally, Internal Coaches prioritize students who previously received MRC services, and any student the Internal Coach believes may benefit from MRC services. Members assess these students, and Internal Coaches review this data to then objectively determine eligibility based upon their benchmark score. Once selected to receive services, members collect weekly progress monitoring data using CBM assessments that are appropriate for their grade level.

Table II.1.   MRC K-3 CBM assessments and benchmarks by grade and season

	Assessment	Fall Target	Winter Target	Spring Target
Kindergarten	Letter Sound Fluency	10	21	41
1st Grade	Nonsense Word Fluency	32	52	n/a
	Oral Reading Fluency	n/a	22	52
2nd Grade	Oral Reading Fluency	42	73	90
3rd Grade	Oral Reading Fluency	70	91	109

The MRC program uses the OnCorps and AIMSweb internet-based data entry systems to record and store general outcome measure and progress monitoring data on all students served by the program. Progress monitoring allows members to chart student progress, assess effectiveness of current interventions, gauge if students require a change in interventions, or determine if they are ready to exit the program. Every student's progress monitoring scores are graphed and then reviewed monthly by a collaborative team consisting of the members, Internal Coach and Master Coach. In the K-3 program, Tier 2 students receive intervention services until their progress monitoring data shows that they have achieved 3 to 5 consecutive data points above the aimline (i.e., projected growth trajectory) and two scores at or above the upcoming season benchmark target. Similar criteria are used for the discontinuation of services with Kindergarten students, although the Spring rather than Winter target is used to determine eligibility for all seasons. Once these criteria are met, a student is deemed "on-track" to achieve appropriate grade-level benchmark at the next assessment window, and is "exited" from the MRC program (i.e., the member no longer provides intervention services). The Master Coach, Internal Coach, and AmeriCorps member discuss each student's assessment results over time before deciding to exit the student from service.

The data intensive orientation of the MRC program provides members, coaches, teachers and principals/directors with a consistent, objective means of identifying students to receive program services, tracking their progress toward achieving academic goals related to critical literacy skills, and informing instruction. The assessment data play an important role in garnering site-wide support from non-MRC-affiliated site staff, particularly as they see quantitative improvement in student outcomes. The data also provide members 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.

In addition to using assessment data to identify individual students for services and to inform instruction, the MRC program also uses data to evaluate and improve the program itself. This continued investment in research and development has led to a number of examples of innovations and program improvements at the systems level. For example, in the formative years of the MRC program, little to no definitive research existed on the reliability of and linearity of measurements of students' slopes of growth using the program's general outcome measurements (i.e., AIMSweb). MRC initially used the research-supported recommendation of two consecutive data points above a student's projected growth trajectory, measured from the most recent benchmark period to the next (e.g., first semester Fall to Winter or second semester Winter to Spring), for their exit criteria. This projected growth trajectory is referred to as an aimline. Over the years, MRC gathered data on the progress monitoring and benchmark assessments, as well as statewide reading assessments in third grade to examine whether the exit criteria were appropriate. The resulting analysis showed that the slope of growth of progress monitoring scores over the course of a school year among students who successfully exited the program, yet later did not reach grade-level criteria on either benchmark assessments or statewide reading assessments, was non-linear; thus, overestimating the student's end of year performance. In response, MRC raised its exit criteria, requiring three consecutive data points above the aimline with at least two of those data points also being above the upcoming season's benchmark target score to ensure that students who exit the MRC program remain on-track to perform at grade-level reading targets.

A second example of MRC's continued program innovation and improvement through the use of data pertains to its K-Focus program. K-Focus is an attempt to serve more students, increase the amount of time spent in intervention, and broaden the scope of interventions used to serve students in Kindergarten. K-Focus achieved these goals by modifying MRC's standard early literacy interventions to be delivered to pairs rather than individual students and by adding a 20-minute shared book reading intervention that includes dialogic reading to focus on phonemic awareness, phonics, and vocabulary instruction. A pilot study in the 2010-11 school year showed that student performance in K-Focus was stronger than in traditional interventions and that 5 to 7 times as many Kindergarten students could be served by the program. In the following year, the program was expanded and additional data were collected, which confirmed the first-year findings. The K-Focus program is now likely to become a standard component of the MRC K-3 program.

III. Impact Evaluation of the Minnesota Reading Corps K-3 Program

Building on the MRC program background provided in Chapter II, this chapter provides a detailed description of the methodology used to implement the MRC K-3 impact evaluation. The methodology was informed by the 2008-2009 and 2010-2011 annual Minnesota Reading Corps evaluations.[11] The chapter begins with the presentation of a logic model for the MRC program outlining key program and school inputs and activities, as well as the program's desired short-term outcomes and long-term goals. After establishing the key components of the MRC K-3 program, the evaluation team presents the three primary research questions for assessing the MRC program's impact on K-3 students' literacy proficiency (i.e., the logic model's K-3 short-term outcome). These research questions guided all aspects of the evaluation design, study implementation, and data analysis.

Following our presentation of the research questions, the next section in this chapter describes the process for sampling schools for the K-3 evaluation and presents key characteristics/demographics of the sampled schools. Next, the experimental design and randomization procedures for the Fall-Winter Experimental Study are provided. The section also includes a description of the process for selecting students at the sampled schools to participate in the evaluation and presents their pre-intervention (i.e., Fall benchmark) characteristics/demographics. A baseline analysis, which confirms the integrity of the randomization procedures (i.e., establishes the equality of the program and control groups at the start of the first semester in Fall 2012), also is provided.

Finally, data collection methods and the use of administrative data are discussed, along with the analytic methods and statistical models for both the Fall-Winter Experimental Study and the Full-Year Non-Experimental Study. The chapter concludes with a discussion of the limitations of the MRC K-3 impact evaluation.

A. Evaluation Logic Model

A logic model for the MRC program illustrating key program and school inputs and activities, as well as the program's desired short-term outcomes and long-term goals is provided in **Appendix A.1.**[12] The MRC logic model was developed jointly by the evaluation team and MRC program staff and served as the conceptual framework for the design of the K-3 impact evaluation. The logic model presents a comprehensive illustration of the complete MRC

---

[11]Bollman, K. & Silberglitt, B. (2009). *Minnesota Reading Corps Final Evaluation 2008-2009.* Minneapolis, MN: MRC.; Bollman, K. & Silberglitt, B. (2011). *Minnesota Reading Corps Final Evaluation 2010-2011.* Minneapolis, MN: MRC.

[12] A comprehensive description of the MRC logic model is available in the following two reports: *Feasibility Study of the Minnesota Reading Corps* (2013) and *Process Assessment of the Minnesota Reading Corps* (2013)

program, and includes inputs, activities, short-term outcomes and long-term goals for four primary program constituencies: PreK students, K-3 students, AmeriCorps members, and schools.

The focus of the K-3 impact evaluation was to assess the impact of MRC program participation on Kindergarten through third grade students' literacy scores. As such, the evaluation focused on only those components of the logic model relevant to K-3 students. These components include MRC program and school-based inputs and resources, Internal Coach and AmeriCorps member coaching and supervision, AmeriCorps member tutoring of K-3 students, and student proficiency outcomes.

Four key MRC program and school based inputs and resources are essential to successful K-3 program implementation: 1) MRC program selection of schools based on degree of student need and school capacity to partner effectively with the program; 2) school identification of at-risk (Tier 2) K-3 students within the school based on benchmark assessment of students' literacy skills; 3) web-based data management systems to track and monitor student progress with literacy interventions (i.e., OnCorps and AIMSweb); and 4) school use of research-based core literacy curriculum.

In addition to the MRC program and school-based inputs, three important MRC program inputs related to AmeriCorps members included: 1) joint MRC and school recruitment, screening and placement of members in schools; 2) comprehensive MRC training of members and Internal Coaches in literacy interventions, assessment, data-driven decision-making and program rules; and 3) school identification and assignment of dedicated Internal Coaches to support and monitor the members. The logic model also illustrates the multiple layers of supervision and coaching the MRC program provides to its school-based Internal Coaches and AmeriCorps members.

As shown in the logic model, the MRC program's primary activities include: 1) conducting benchmark testing three times per year (Fall, Winter and Spring) to identify students in need of literacy tutoring (i.e., Tier 2 students); 2) delivering one-on-one tutoring to eligible students 20 minutes a day, 5 days a week (i.e., tutoring); 3) assessing and charting weekly student progress on grade-specific literacy skills using AIMSweb (i.e., weekly progress monitoring); 4) "exiting" students from the program once they achieve assessment scores putting them on track to meet or exceed the next benchmark; and 5) identifying and tutoring new students eligible for the program. The intended short-term outcomes of these activities are demonstrated improvement on AIMSweb measures at the subsequent benchmarking period (i.e., Winter, Spring) and a successful (permanent) exit from MRC tutoring services. The desired long-term outcomes of the MRC program is for all third grade students to meet or exceed grade-level proficiency on the state's third grade reading test (MCA-III).

Thus, when reviewing the logic model, note that the K-3 impact evaluation assessed the cumulative impact of three specific program elements: 1) the MRC central operations and school-based inputs and resources; 2) Internal Coach and AmeriCorps member coaching and supervision; and 3) AmeriCorps member tutoring of Tier 2 K-3 students.

B. K-3 Impact Evaluation Research Questions

As the logic model presented above illustrates, the MRC program's short-term objective is to improve Tier 2 students' literacy skills so they are on track to achieve grade-level proficiency. The primary goal of the MRC K-3 impact evaluation was to independently and experimentally assess the impact of the MRC program on K-3 students' literacy proficiency scores. To achieve this goal, the K-3 impact evaluation focused on the following three research questions:

1. RQ1: What is the impact of the MRC program on student literacy outcomes?
    a. Does the impact vary by student characteristics/demographics?
    b. Do assessment scores vary by AmeriCorps member characteristics/demographics?

2. RQ2: Does the impact of the program vary week to week? Does the number of weeks of intervention (i.e., dosage) impact student literacy outcomes?

3. RQ3: Does participation in MRC have a longer-term impact on student literacy outcomes as measured at the end of the school year?

To explore these research questions, the evaluation team analyzed grade appropriate and semester specific literacy assessment scores collected weekly from a sample of 1,148 K-3rd grade students enrolled at 23 diverse elementary schools participating in the Minnesota Reading Corps program during the 2012-2013 school year. At the beginning of the school year, program eligible students were randomly assigned to either receive MRC tutoring (i.e., program group) or not receive any MRC interventions (i.e., control group) during the first semester (Fall 2012 –Winter 2013). Data from the first semester were used to answer the evaluation's first research question (RQ 1), which assesses the impact of the MRC program on student literacy outcomes. The results of the analysis from the first semester are presented in Chapter IV, "Fall-Winter Experimental Study Findings."

It was not possible to continue the experimental RCT throughout the entire school year due to school apprehension about withholding MRC services, AmeriCorps contract service requirements for MRC members, and ethical concerns about withholding supplemental reading assistance from needy students for an entire school year. As such, all students who were eligible at the Winter benchmark to participate in the MRC program were allowed to receive services during the second semester of the 2012-2013 school year (January 2013 – May 2013). Data from the first and second semester were combined to answer the second and third research questions aimed at assessing the week to week impact of the MRC program (RQ2), as well as the longer-term (i.e., full year) impact of the program on

student literacy outcomes (RQ3). The results of these full-year non-experimental analyses are presented in Chapter V, "Full Year Non-Experimental Study Findings" and Chapter VI, "Exploratory Analysis."

C. School Selection

The school selection process was designed to ensure representation of the diverse schools participating in the MRC K-3 program. Therefore, school selection was both purposive and involved a multi-step stratified sampling process. Since the MRC program began 10 years ago in 2003 and expanded rapidly over the years, there was a wide range in the number of years schools had participated in the MRC program. As it was not our intention to assess how well schools adopted/executed the program during their *first* year of participation, the school sampling process began by limiting the sampling frame to only the 200 schools that had fully implemented the K-3 MRC program for at least two consecutive years.

Next, to ensure geographical diversity, schools were stratified by urbanicity (i.e., urban, suburban, and rural) using the MRC program regions. The various locations of MRC sites are geographically organized by the program into 8 regions of the state. The only areas in the state of Minnesota that can be categorized as urban and suburban are in and immediately surrounding the cities of Minneapolis and St. Paul. Therefore, it was necessary to include the MRC Metro region that encompasses these areas in the evaluation's sampling frame. The areas in the state of Minnesota, both immediately outside the Metro region and in the farther reaches of the state are geographically similar and rural in urbanicity. Thus, the sampling frame for rural schools was limited to three MRC regions within a three-hour-drive radius of the Metro area (i.e., the Central, Southwest, and Southeast regions). This allowed the evaluation team to more readily and economically provide on-site support during the random assignment process and conduct follow-up site visits during the year-long evaluation.

Once the sampling frame was defined, the evaluation team selected 25 schools using Probability Proportional to Size (PPS), whereby larger schools with a more pronounced need (defined as the number of students served by MRC during the 2010-2011 school year) had a higher probability of selection. Using PPS ensured a statistically adequate sample size to conduct the K-3 impact evaluation. A set of alternative sites also was selected in case a school was determined to be ineligible or declined participation in the evaluation. It is important to note that while the MRC program staff encouraged sites to participate, participation in the evaluation was voluntary. In the end, 23 elementary schools agreed to participate in the K-3 impact evaluation during the 2012-2013 school year.[13] The list of participating schools and their key characteristics are provided in **Table III.1** below.

---

[13] Schools who did not participate referenced scheduling conflicts and staff shortages as possible reasons.

Table III.1.   Characteristics of schools participating in the MRC K-3 Impact Evaluation (Fall 2012)

School	Location in Minnesota	MRC Region	Urbanicity[1]	Year MRC began	Number of AmeriCorps Members[2] Full Time	Number of AmeriCorps Members[2] Part Time	% FRPL	School Enroll-ment[3]	Study Parti-cipants (N)
Becker Intermediate	Becker	Central	Rural	2008	1	-	23.1%	659	19
Becker Primary	Becker	Central	Rural	2008	1	1	22.8	605	12
Bel Air ES[4]	New Brighton	Metro	Suburban	2009	1	-	39.9	794	28
Bryn Mawr ES	Minneapolis	Metro	Urban	2009	2	-	81.7	465	62
Forest Hills ES	Eden Prairie	Metro	Suburban	2010	1	-	33.7	650	24
Franklin ES	Rochester	Southeast	Rural	2009	1	-	56.3	625	32
Frost Lake Magnet	St. Paul	Metro	Urban	2009	4	-	87.8	550	120
Green Central Park ES	Minneapolis	Metro	Urban	2007	2	1	95.0	617	56
Jackson Magnet	St. Paul	Metro	Urban	2007	4	1	91.4	487	134
Jefferson ES	Rochester	Southeast	Rural	2007	1	1	44.2	500	46
Jenny Lind ES	Minneapolis	Metro	Urban	2007	4	-	93.7	539	109
Kaposia ES	S. St. Paul	Metro	Suburban	2009	2	1	57.3	858	78
Nellie Stone Johnson ES	Minneapolis	Metro	Urban	2007	4	-	95.7	741	111
North ES	Princeton	Central	Rural	2008	-	2	31.7	745	32
Northrup ES	Minneapolis	Metro	Urban	2009	3	-	49.0	431	56
Oakdale ES	Oakdale	Metro	Suburban	2009	1	1	59.1	514	43
Wellstone ES	St. Paul	Metro	Urban	2008	3	1	94.2	694	104
Phalen Lake Hmong Studies Magnet	St. Paul	Metro	Urban	2009	4	-	93.8	725	120
Pine City ES	Pine City	Central	Rural	2006	2	1	48.6	880	78
Riverside ES	Rochester	Southeast	Rural	2009	1	-	70.5	522	30
Sheridan Arts Magnet	Minneapolis	Metro	Urban	2009	4	-	94.0	470	136
South ES	Princeton	Central	Rural	2010	2	1	33.1	792	80
Sunset Terrace ES	Rochester	Southeast	Rural	2007	1	-	46.3	703	30

Key: [1]Urban, Suburban or Rural; [2]Serving at the start of the 2012 school year; [3]Schools may have varying grade levels. (e.g., K-2, K-5, etc.); [4]ES = Elementary School

Sources: MRC Program Administrative Data 2011, Minnesota Department of Education Data Center

D. Random Assignment of Students Within Schools

All students at the 23 sampled schools identified by Fall benchmark scores as eligible for MRC services (i.e., Tier 2) were randomly assigned to either the MRC program (i.e., treatment) or control group at the beginning of the first semester prior to the start of tutoring.[14] During the first few weeks of the first semester, AmeriCorps members conducted the Fall CBM assessments on students who the school's staff identified as potentially eligible for MRC services.[15] Scores on these assessments determined students' eligibility for MRC services. The total evaluation sample size was determined by the number of students that AmeriCorps members at a school were capable of assessing on a weekly basis.[16] Students were added to the sample by grade in pairs, whereby each student in each grade within a school was matched with another student based upon their Fall benchmark score. Students within pairs were then randomly assigned to either the program or control condition. This matched pair design ensured that students in the program and control groups had similar Fall benchmark scores at the start of the school year. A figure illustrating the randomization process is provided in **Appendix A.2**.

When a school had more eligible students than the AmeriCorps members' had capacity to assess, the evaluation team recommended that the school follow their usual procedures for selecting students for the program. Most schools rank order eligible students within grade by benchmark score and provide tutoring to students closest to the benchmark first. Therefore, in instances where the number of students eligible for MRC services was greater than the number of students the AmeriCorps member(s) could assess on a weekly basis, a subset of students closest to benchmark was selected, and the students beyond the members' capacity to assess were excluded from the evaluation sample. Furthermore, in an effort to obtain as equal numbers of students in each grade as possible, the evaluation team paired students by similar benchmark score and added the two students closest to benchmark to the study sample in each grade (e.g., 2 kindergartners, 2 first grade students, 2 second grade students, and 2 third grade students for a total N of 8), repeating the process iteratively until the maximum number of students that could be assessed at the school was reached.[17] In some schools, the number of eligible students differed by grade; therefore, the total number of students selected to participate in the evaluation also differed by grade.[18] After students were

---

[14] Students who participated in the MRC program in previous years were eligible for participation in the evaluation. The important eligibility criteria for students was not whether they had received MRC services in the past, but instead whether they were eligible to receive services at the beginning of the 2012-2013 school year. Furthermore, since students were randomly assigned to condition, it was equally likely that a student who previously received services would be assigned to the program and control groups. As such, each group should have a roughly equal number of students who had and had not participated previously in the MRC program.

[15] Processes for identifying students can somewhat vary across schools. Some schools assess all students in grades K-3, while others may use previous years' test scores or other more subjective means for identifying students to be assessed. Schools were asked to use the same procedures they typically employ to identify students for Fall benchmark testing.

[16] On average, one full-time member is able to assess 30 students each week.

[17] The matched-pair random assignment design is described in section III.D.

[18] Note that as reported in Table III.2, the number of second grade students eligible for MRC services was lower than any other grade.

paired within grade based upon their Fall benchmark scores, they were then randomized to program or control conditions.

The final column in **Table III.1** lists the number of students at each of the 23 schools whose Fall benchmark scores made them eligible to receive MRC services during the first semester. In total, 1,530 eligible students were selected to participate in the evaluation. As the table shows, the number of students eligible to participate in the MRC program varied between schools. This variability resulted from differences in school enrollments (i.e., the larger the enrollment, the more students eligible for services) and in numbers of students who posted eligible benchmark scores (i.e., higher performing schools had fewer eligible students).

Table III.2 presents descriptive statistics for the K-3 students included in the evaluation. Demographics include gender, race/ethnicity, Dual Language Learner (DLL) status, Free and Reduced Price Lunch (FRPL) status, age, and Fall benchmark scores. The sample size for the evaluation is smaller than the 1,530 students eligible for program services. During the school year, some students left the school area (i.e., moved) or were chronically absent and did not receive regular MRC tutoring or assessments. These students and their matched pair were removed from the analytic sample (i.e., pairwise deletion). Thus, the final sample of students included in the evaluation totaled 1,341 students.[19]

The evaluation team conducted a power analysis prior to developing the evaluation's sampling plan, in which the number of students required to detect a difference between the treatment and control groups was calculated. The power analysis relied on the student, member, and school population numbers reported by the MRC program from the 2009-2010 school year. Assuming a Minimal Detectable Effect (MDE) of 15% based on findings from the MRC program's annual evaluations, the study required 600 treatment and 600 control members within the MRC K-3 program. In addition, the evaluation team conducted sensitivity tests to determine if the loss of students from the sample was systematic and, therefore, could potentially introduce bias into the evaluation results. The findings from the sensitivity tests are available in **Appendix A.3**.

---

[19] For each grade, we demonstrate low levels of attrition assuming the "liberal" standard outlined in the *WWC Evidence Review Protocol for Early Childhood interventions (Version 2):* http://ies.ed.gov/ncee/wwc/pdf/reference_resources/ece_protocol_v2.0.pdf

Table III.2.   Student participants for the MRC K-3 Impact Evaluation (Fall 2012)

	Kindergarten (N=359)		1st Grade (N=409)		2nd Grade (N=265)		3rd Grade (N=308)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Female	56%	-	52%	-	46%	-	46%	-
Race/Ethnicity								
White	30%	-	39%	-	36%	-	40%	-
Black	33%	-	22%	-	29%	-	23%	-
Asian	27%	-	26%	-	25%	-	27%	-
Hispanic	7%	-	12%	-	9%	-	8%	-
Other	3%	-	1%	-	1%	-	2%	-
Dual Language Learner (DLL)	26%	-	37%	-	37%	-	31%	-
Free and Reduced Price Lunch (FRPL)	76%		72%		75%		71%	
Age	5.53	0.30	6.57	0.48	7.57	0.35	8.57	0.35
Fall Benchmark Score[1]	4.58	6.66	23.02	7.00	26.97	10.11	52.31	13.48

[1]Kindergarten = Letter sound fluency; 1st grade = Nonsense word fluency; 2nd & 3rd grade = Oral reading fluency

Table III.2 shows that approximately 30% of students were White, 30% Black, and 30% Asian. Only 10% of students were Hispanic. It is important to note that unlike the broader American population, the Minneapolis/St. Paul region has a large Asian population (specifically, Hmong). Also, approximately one-third of the analytic sample included students who were classified by their schools as Dual Language Learners (DLLs). **Table III.3 s**hows that the majority of DLL students were Asian, and, as such, the comparatively high percentage of Asian students, and Asian DLL students specifically, reflects Minnesota's unique demographics.

Table III.3.   Student participants' DLL Status by race/ethnicity for the MRC K-3 Impact Evaluation (Fall 2012)

Race/ Ethnicity	Kindergarten				1st Grade				2nd Grade				3rd Grade			
	DLL		Not-DLL		DLL		Not-DLL		DLL		Not-DLL		DLL		Not-DLL	
	%	N	%	N	%	N	%	N	%	N	%	N)	%	N	%	N
White	3.6%	3	96.4%	80	2.7%	4	97.3%	142	3.5%	3	96.5%	83	2.0%	2	98.0%	99
Black	3.9	3	96.1	75	20.0	15	80.0	60	15.8	9	84.2	48	6.8	4	93.2	55
Asian	67.8	61	32.2	29	82.7	80	17.3	17	91.9	57	8.1	5	84.8	67	15.2	12
Hispanic	58.3	14	41.7	10	75.0	33	25.0	11	76.0	19	24.0	6	61.9	13	38.1	8
Other	33.3	2	66.7	4	0.0	0	100	4	66.7	2	33.3	1	20.0	1	80.0	4

The evaluation team used a NORC proprietary, centralized password-protected website to conduct the within pair randomization of students to the program or control groups. MRC Master Coaches entered Fall benchmark and demographic data on all MRC eligible students into the website. The website included grade-specific Fall benchmark assessment score range checks to ensure that only those students eligible for MRC services were included in the sample. Once all eligible students' data was entered into the website and confirmed, the website then automatically randomized students within pairs to either the program or control group and displayed each student's assignment. The Master Coach then shared the list of students with the school's Internal Coach and AmeriCorps members.

Table III.4 compares key demographic variables between the program and control groups. The overall lack of differences between program and control group students on most variables confirms successful randomization. The within-pair difference between Kindergarten control and program students on Fall benchmark scores, while significant, is small and favors the control group students. That is, the control group students posted slightly higher Fall benchmark scores on average than students in the program group. In the third grade, females posted slightly higher Fall benchmark scores than males. Any baseline differences between the program and control groups were accounted for in the final analysis models.

Table III.4. Differences between control and program group students by grade (Fall 2012)

	Kindergarten			1st Grade			2nd Grade			3rd Grade		
	C-P	SE(C-P)	Sig.	C-P	SE(C-P)	Sig.	C-P	SE(C-P)	Sig.	C-P	SE(C-P)	Sig.
Fall Benchmark												
Independent	0.16	0.68		0.44	0.69		0.71	1.24		0.62	1.54	
Within-pair	-0.27	0.10	**	-0.30	0.20		-0.24	0.18		0.05	0.25	
Race/Ethnicity												
White	-0.07	0.05		-0.03	0.05		-0.01	0.06		-0.01	0.06	
Black	-0.03	0.05		0.03	0.04		-0.02	0.06		0.03	0.05	
Asian	0.06	0.05		-0.04	0.04		-0.01	0.05		-0.02	0.05	
Hispanic	0.03	0.03		0.04	0.03		0.02	0.04		-0.01	0.03	
Other	0.03	0.03		0.04	0.03		0.02	0.04		-0.01	0.03	
Female	0.05	0.05		0.06	0.05		0.04	0.06		0.15	0.06	**
DLL	0.05	0.05		-0.03	0.05		0.06	0.06		0.02	0.05	
FRPL	0.10	0.04	*	0.05	0.04		0.03	0.05		-0.04	0.05	

Key: C=Control, P=Program, SE=Standard Error
*p<.05; **p<.01

Because the evaluation was designed to measure the impact of MRC program participation relative to nonparticipation, students in the control group were embargoed from receiving tutoring services during the first semester of the school year. Thus, Internal Coaches and AmeriCorps members were asked to tutor the students assigned to the program group as usual and to not tutor the control group students until after the Winter benchmark.

The evaluation team's Site Liaisons remained in regular contact with the Internal Coaches and Master Coaches to verify that program students received MRC tutoring services and control students did not. The evaluation's Principal Investigators also made periodic in-person site visits to further ensure that Internal Coaches and AmeriCorps members maintained the program and control group distinction and continued to collect weekly progress monitoring data each week. Also, while control students could not receive tutoring during the first semester, students in both the control and program groups were allowed to participate in any other tutoring or reading programs offered at the school. The evaluation team collected from the schools information about all alternative reading tutoring programs offered to students at the school. Further, AmeriCorps members recorded student participation in alternative programs during the weekly progress monitoring assessments. The results of the Fall-Winter Experimental Study are presented in Chapter IV.

It was not possible to continue the experimental RCT throughout the entire school year due to school apprehension about withholding MRC services. As such, all program and control students who were eligible at the Winter benchmark to participate in the MRC program were allowed to receive services during the second semester of the 2012-2013 school year (Winter 2013 – Spring 2013). The results of the Full-Year Non-Experimental Study are presented in Chapter V.

E.   Use of Administrative Data

The evaluation team partnered with the Minnesota Reading Corps program to utilize AmeriCorps members to collect all the benchmark and weekly progress monitoring data comprising the primary data for the evaluation. As mentioned previously, AmeriCorps members collect benchmark data three times a year on all K-3 students who the school identifies as potentially eligible for MRC services. This procedure identifies objectively those students who meet MRC eligibility criteria (i.e., Tier 2). Given that Members already collect this data and so as not to duplicate assessment efforts, the evaluation team requested access to the 2012 Fall benchmark data in order to identify the pool of students eligible to participate in the evaluation.

AmeriCorps members were asked to collect weekly progress monitoring data from students in both the program and control groups for the entire school year and enter the assessment scores in the MRC AIMSweb website and/or OnCorps website. Typically, AmeriCorps members collect and enter weekly progress monitoring assessment scores only while students are actively receiving tutoring services and up to three weeks after exiting the program. While the evaluation caused the AmeriCorps members to collect significantly more data than usual, the assessment and data entry procedures did not differ from normal.

The evaluation team obtained benchmark and weekly progress monitoring data directly from the MRC program throughout the school year. The data were reviewed to ensure that schools were collecting weekly progress

monitoring data on all students participating in the evaluation. At the end of the 2012-2013 school year, the complete dataset was de-identified and exported to the evaluation team for analysis. MRC provided the evaluation team with other administrative data including AmeriCorps member demographics, tutoring dosage and alterative intervention information. All assessment, demographic and intervention data was combined into a single dataset that is the source of all analyses and results presented in this report.

Table III.5 below lists the alternative interventions received by control group students as reported by the schools and supported with administrative data from MRC. At the beginning of the evaluation, the Internal Coaches at participating schools were asked to report all alternative intervention services offered by their school. The evaluation team made a determination as to whether these alternative interventions offered services similar to MRC; that is, where they were intended to support and/or improve students' literacy skills. If the alternative intervention offered literacy services, then AmeriCorps members at participating schools were asked to record on a weekly basis whether and for how long each control group students received the intervention.

Table III.5.   Alternative interventions received by control group students by school (2012-13 school year)

Intervention Category	Intervention Name	School Code+																			
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
Small Group Intervention	Fountas and Pinnell Leveled Literacy Intervention	x	x					x	x			x								x	x	
	SRA Reading Mastery (and Decoding)	x				x																
	Early Success					x																
	Direct phonics instruction													x								
	Early Intervention in Reading (EIR)													x								
	Spire																	x				
	Reading A-Z																	x				
	Jolly Phonics																	x				
	Making Connections																				x	
	General "Guided Reading" / Intervention Services		x	x						x						x						
	Hmong Language Reading Intervention Services			x																		
Individual Intervention	MRC-adapted intervention services				x		x															
	LDA Minnesota									x												
	Great Leaps																	x				
Computer Based Intervention	Read Naturally					x															x	
	Earobics	x																				
	Lexia																				x	
	Reading Eggs																				x	
After School/ Miscellaneous Interventions	After School Academy									x		x		x		x						x
	Extended Day/After School Learning Program									x	x											
	Weekly Conferencing									x												
	Experience Corps														x							
	Foundations																				x	

Note: 3 schools did not offer literacy-focused, supplemental alternative interventions.

+ The names of schools that correspond to the school codes are available in Appendix A.4.

F. Analysis

In this section, we provide a summary of the methods and analysis approaches used to analyze the student assessment data from the Fall-Winter Experimental Study, the Full-Year Non-Experimental Study, and the Exploratory Analysis. A more detailed methodology section is provided in **Appendix A.5**.

As explained in our presentation of the research questions in section III.B above, three sets of analyses were conducted to answer the primary research questions. The first set of analyses focus on the Fall-Winter Experimental Study data collected during the first semester of the 2012-2013 school year and was used to answer the first research question (RQ1), which assesses the impact of the MRC program on student literacy outcomes. The Fall-Winter experimental analyses employed growth models utilizing weekly progress monitoring data as primary outcomes and the Fall benchmarks as covariates to assess the impact of the MRC program on student literacy proficiency at each grade level.

Because it was not possible to continue the experimental RCT throughout the entire school year, all students who were eligible at the Winter benchmark to participate in the MRC program were allowed to receive services during the second semester of the 2012-2013 school year. Data from the first and second semester were combined to answer the second and third research questions aimed at assessing the week to week impact of the program on student literacy outcomes (RQ2), as well as the longer-term impact of the program on student literacy outcomes (RQ3).

Other analyses included robustness checks on the variability of the treatment effects and intraclass correlation analysis of Fall to Winter gains. We finally explored other possible analysis approaches to measure program effects using spline methods.

Analysis 1: Fall-Winter Experimental Analyses

Impact Analysis

The overall goal of the Fall-Winter Experimental Study was to obtain a measure of an average treatment effect of the MRC program by grade. In a simple experiment, this measure of impact would be a straightforward comparison of program (treatment) and control group means at the end of the experimental period. However, because the evaluation team was able to obtain weekly assessment data on all students in the study, we were able to estimate the impact of the program on participants as a varying quantity over the weeks of the experimental period (Fall-Winter). This approach required that we estimate the effect of the MRC program on a weekly basis concurrent with typical growth trajectories. Therefore, our experimental analysis consisted of two phases for each grade. In the first phase, a statistical model was developed to produce a set of regression coefficients that summarized the observed data. In the second phase, we then used the estimated model to predict marginal (average) scores for program

participants and control group students at several weekly points. From these points, the evaluation team estimated the difference between program and control group students to measure program impacts. We then calculated the precision of our estimates for statistical tests.

Two types of subgroup analysis were employed by the evaluation team. First, we conducted Chi-Square tests to determine if there were differential impacts across major subgroups, implying that the MRC program may be more or less beneficial for selected subgroups of students. For the second subgroup analysis, we subdivided the data by grade and by major subgroup (e.g., females, males, etc.) and used a similar two-phase approach as described above for measuring the average treatment effect of the MRC program by grade. Subgroups were defined based on characteristics provided at baseline and each level was determined by individual characteristics, including gender, race, Dual Language Learner (DLL) status, and Free and Reduced Price Lunch (FRPL) eligibility.

Robustness Analysis

The evaluation team also tested the variability of the program effect at the assigned pair level. For this analysis, we estimated the difference in the trajectory between program and control students, then estimate whether, on average, this difference was positive. The findings from the robustness analysis are provided in **Appendix B.2**.

Member- and School-Level Random Effects on Impacts

In order to determine whether members and/or schools were influential on program impacts for students receiving MRC tutoring, an intraclass correlation (ICC) analysis was conducted. Intraclass correlations for members measure the degree to which different students assigned the same member correlate on changes or gains in their assessment scores within a school. The school-level ICC measures the degree to which students, who are located within the same school, but assigned to different members, correlate with one another.

AmeriCorps Member Fixed Effects on Assessment Scores

To assess the impact of member fixed characteristics (i.e. observed demographics) on treatment students' assessment scores, the evaluation team estimated another statistical model on a set of data where the assessments were standardized within grade and combined into a single model. In other words, we estimated standardized assessment scores within each grade, combined the data, and fit a model predicting scores with member characteristics (age, education years, specialized degree, full time status, tenure status, race/ethnicity, and gender), a simple time trend, student gender, and the interaction between member and student gender.

Analysis 2: Full Year Non-Experimental Analysis

For the second set of analyses, data from the first and second semester were combined to answer the second research question, which focused on better understanding patterns of week over week growth in reading proficiency

by grade. After the first semester, all students who were eligible at the Winter benchmark to participate in the MRC program were allowed to receive services (program and control group members). Therefore, while the evaluation team employed similar analytic coding strategies and estimation techniques as applied in the first analysis, the second analysis differed in that we acknowledged that variation in dosage existed over the course of the year. Therefore, our primary predictor was not program participation as a moderator of trajectory, but instead the number of cumulative weeks of receiving the program. To summarize the results of this model, we estimated the tangential change in the outcome when moving from 0 to 1 weeks of sessions, 1 to 2 weeks of sessions, and so on.

Analysis 3: Exploratory Analysis

The ultimate goal of the program is to matriculate students into an "above benchmark" status. Therefore, our final analysis examines this question by asking which trajectory pattern are program participants most likely to follow. We conceived of two plausible methods to measure this. First, we consider nominal categories of trajectory patterns and estimate the likelihood of a student falling into one of these nominal categories based on the experimental assignment. Our second approach views trajectory as a more linear concept that is altered during and after exposure to the program.

In the first approach we coded each student as belonging to the following categories:

- No program effect (always below benchmark)
- Temporary program effect (moving above and below benchmark and ending the school year below)
- Final program effect (starting below benchmark, then eventually progressing and remaining above benchmark for the remainder of the school year)
- Always above benchmark
- Negative effect (above, then ending below benchmark)

Of these five patterns, the latter two represented a very small portion of students (4) and were not included in the final analysis. We performed a multinomial logistic regression analysis predicting the likelihood of students falling into each of the three major classes or groups as a function of program assignment, cumulative sessions, and demographics. We also included an interaction term between program assignment and cumulative sessions, again because some students that where initially assigned to the control group received sessions in the second semester.

Once we estimated the model (provided in **Appendix C**), the evaluation team then predicted the likelihood of following each group pattern for two average students who received the same dosage of the program, specifically 10 tutoring sessions, with the only difference being that one student was originally assigned to the program group and the other to the control group. The key comparison is the likelihood of falling into the "final program effect" category

for those students assigned to the program group versus the control group. Because all other differences between the two theoretical students were held constant in the model, including program dosage (i.e., 10 sessions of tutoring), the only difference is the timing in which the students would have received MRC tutoring (i.e., first vs. second semester). Therefore, the findings from this analysis indicate whether early intervention of the program results in a higher or lower likelihood of falling into the "final program effect" group.

The second analysis employed spline models, which allowed for a linear trajectory across the entire school year, but also allowed for this trajectory to be adjusted based on two phases: during treatment and after treatment. Spline models are typical in analyses that seek to fit a model to data measured over time, where specific events are hypothesized to change the trajectory. A non-education example would be political poll data altering trajectory after a major public event. In our case, the spline models measure not only the growth of students over the course of the year, but specifically how those growth patterns are altered by participation in the program.

The purpose of the analysis was to examine changes to the trajectory of student achievement prior to, during, and after receiving MRC tutoring, regardless of initial assignment to conditions (program or control group). We fit the model on a subset of cases in each grade that received at least one weekly session of tutoring during the school year. We also limited the analysis to only 10 weekly tutoring sessions to remove the cross-semester (i.e., first vs. second semester) effects of low performing students who would have required a larger dosage of the MRC intervention.

G.   Limitations of the Study

The primary objective of the MRC K-3 evaluation was to assess the impact of the MRC program on Kindergarten through third grade students' literacy scores (i.e., Research Question 1). In order to achieve this objective, the evaluation was designed to measure MRC program impacts in an experimentally rigorous and efficient manner. However, as discussed previously, unavoidable field limitations inherent in working with elementary schools and students constrained some aspects of the evaluation's design, implementation, and analysis.

One major constraint was that school participation in the evaluation was voluntary. As such, the evaluation had to be designed to be reasonable and appealing to schools in order to recruit them for participation. This implies two related limitations. First, as this is not a simple random sample of schools, it is plausible that our estimates may not generalize to the population. We address this issue by using stratified random sampling to select schools and employing school-level weights in our statistical models. Second, while a randomized controlled trial (RCT) was immediately identified as the ideal method for assessing program impacts, the reality that some students would be

barred from receiving MRC services during the entire 2012-2013 school year was unacceptable to school administrators. All schools eligible[20] to participate in the K-3 evaluation had participated in the MRC program for several years and observed the results of program participation on their students' literacy outcomes. Therefore, virtually none of the schools with existing MRC programs were likely to be willing to bar their students from receiving what they perceived as effective program support for an entire school year. Therefore, in order to recruit schools to participate in the evaluation, it was necessary to create a rigorous study design that could answer RQ1 without barring control students from receiving MRC services during the entire school year.

To address this potentially fatal design issue, the evaluation team examined the MRC annual internal evaluations that strongly suggested a single semester was sufficient time to observe change in student literacy outcomes as a function of program participation. With this additional evidence supporting a shorter embargo period, the evaluation team compromised with schools by asking them to maintain the distinction between the program and control groups for just the first semester of the school year (i.e., Fall 2012 through Winter 2013). Although a single semester Fall-Winter experimental RCT design allowed for a rigorous assessment of program impacts on student literacy outcomes over the course of 16 weeks, it is theoretically possible that maintaining the distinction between the program and control groups throughout the entire school year could have resulted in a different finding in terms of program impacts, even though, pragmatically, a year-long study was not feasible.

Given that the school recruitment-imposed constraint limited the overall amount of time we could *experimentally* examine program impacts between the program and control groups, two separate analytic strategies were employed to address all three research questions (see section III. F for a detailed description of each). As just discussed, one analysis utilized only the first semester student outcome data during which time the RCT was in effect (i.e., Fall-Winter experimental analysis). The other analysis, which was designed to answer key questions on the cumulative and longer-term effects of the program (RQs 2 and 3), used student outcome data collected throughout the entire school year, during which time all children, despite assignment to program or control groups, were eligible to receive MRC tutoring services at some point (i.e., Full-Year non-experimental analysis).

Although the Full-Year Analysis provides a longer-term picture of program impacts, the use of assessment data from the entire school year comes with important limitations in interpreting the results. Recall that at the end of the first semester, all program and control students were reassessed (i.e., Winter benchmark). This Winter benchmark score was then used to determine student eligibility to receive services during the second semester. At this point, the first semester RCT, which embargoed control group students from receiving MRC service, ended. Therefore, all eligible students, many of which were initially assigned to the control group, could now receive MRC tutoring. AmeriCorps

---

[20] Eligible schools had participated in the MRC K-3 program for at least two consecutive years prior to the 2012-2013 school year.

members enrolled new students in the program when spots became available in their tutoring schedule.[21] As mentioned in Chapter II, members enrolled students based upon their Winter benchmark score. Those closest to the benchmark were given priority over those further from the benchmark. Therefore, students' participation in MRC tutoring in the second semester was highly correlated with their benchmark assessment scores. This creates a statistical challenge, where the residuals of assessments are correlated with the benchmark and treatment predictors. Given the bias possibly created by the nonrandom method of program enrollment during the second semester and the correlation between scores and receipt of services, the results of the Full-Year non-experimental analysis should be interpreted with caution.

Unequal sample sizes across the four grades (Kindergarten through third grade) were another program-induced constraint that resulted in differences in statistical power to detect program effects within grades. As discussed above, while some schools assess all students for eligibility, others only assess K-3 students who they believe may be eligible for MRC services. AmeriCorps members assess these students at the beginning of the school year (i.e., Fall benchmark) to determine objectively whether the students are eligible for services. Because eligibility was determined by benchmark score, it was virtually always the case that the number of eligible students varied by grade.

In order to create as equal sample sizes as possible across grades, the evaluation team enrolled students in the evaluation by accepting pairs of students within grade matched on Fall benchmark scores (see Section III. C for more detail). However, if an unequal number of students within a grade tested eligible for program services, the sample sizes among grades would vary. **Table III.2** (above) shows that the natural variability in program eligibility indeed resulted in somewhat unequal sample sizes across Kindergarten through third grades. Second grade had the smallest sample size (N=233), whereas first grade had the largest (N=366). The evaluation team did not anticipate such a wide range of eligible students by grade. The smaller sample size at second grade reduced the statistical power to detect differences between the program and control groups compared to the other grades. Although the evaluation team oversampled schools so as to ensure a large enough pool of eligible students across schools, it was not always possible to predict the number of eligible students within each grade.

Another limitation inherent in the conduct of the evaluation was the necessity to assess students weekly in both the program and control groups. It is unclear what effect if any the act of being assessed might have had on student outcome scores. It is possible that simply participating in weekly oral reading fluency assessments, for example, might positively impact student oral reading proficiency. If such a test effect exists, the effect would benefit the control group, increasing their scores above what one would typically expect had they not been assessed.

---

[21] After winter benchmark, members continued to tutor eligible students who they previously tutored during the first semester. As students exited (i.e., graduated from) the program, spots became available for members to tutor a new student.

More common limitations in education impact evaluations that the evaluation team anticipated might occur in this study included: student attrition, data integrity, and tracking alternate interventions. As mentioned above, the evaluation team implemented a matched-pair experimental design for the Fall-Winter Experimental Study. Students within a school within grade were matched on Fall benchmark assessment scores and then randomly assigned within pair to either the program or control group. Statistical tests showed that the program and control groups did not differ on important demographic variables and proficiency measures at the beginning of the school year. As is inevitable, students moved away or were chronically absent from school (i.e., attrition). When this occurred, the affected student and the matched student within the pair were both dropped from the study's analytic sample.[22] This pairwise deletion procedure reduced the overall sample size by two students, but ensured the integrity of the RCT design. In addition to attrition, other pairs were eliminated from the analytic sample because of implementation error (i.e., members accidently provided a control student with tutoring, or failed to provide a program student with tutoring) or data entry error (i.e., incorrect student IDs). Pairwise deletion for these reasons reduced the sample size by 200 pairs, but guaranteed a non-biased balance between the program and control groups.

Because the use of pairwise deletion resulted in additional students being removed from the original sample, the evaluation team performed a sensitivity analysis to predict whether the removed sample members continued to be represented in the analyses using a logistic regression that employed demographics and treatment assignment as predictors (see **Appendix A.3**). We found that female and minority students and students assigned to the program group were less likely to be included in the final analysis sample (i.e., more likely to be represented in the dropped cases), while DLL students were more likely to be represented in the final sample (i.e., less likely to be dropped). Thus, while the randomization preserves the estimate of the treatment effect on the treated, a higher percentage of the dropped students were female and members of a minority group.[23] However, in the final analysis sample, we had sufficient balance on these factors between treatment and control groups. Therefore, while demographics played a role in selection into our final analysis sample, the effects were evenly spread across treatment and control groups.

The evaluation team employed Site Liaisons to regularly communicate with schools to safeguard data and RCT integrity. AmeriCorps members at 23 different schools were responsible for both maintaining the integrity of the program and control groups and collecting data on over 1,500 students throughout the school year. The Site Liaison communicated at first weekly and then monthly with their assigned schools' Internal Coaches and AmeriCorps members to reinforce and confirm maintenance of the distinction between the program and control groups, as well as the collection of weekly progress monitoring data from students in both groups. AmeriCorps members entered into

---

[22] The evaluation team did not use pairwise deletion for the WWC analysis provided in Appendix C.

[23] Since this analysis only modeled whether or not a sample member was included in the analysis, the precise mechanism of exclusion was not explicitly modeled as a function of covariates.

AIMSweb.[24] student's weekly progress monitoring data and notes about the types of interventions they provided to students. To further monitor data and RCT integrity, the evaluation team also received monthly data downloads from AIMSweb. These data downloads included all students' weekly progress monitoring scores and notes about interventions. If assessment data was missing, the Site Liaison immediately contacted the school Internal Coach to inquire about the reasons and make plans to restart weekly progress monitoring.

Another common limitation related to interpretation of impact evaluation results is the inability to assess whether alternative interventions impacted the study findings. To overcome this limitation, the evaluation team trained all AmeriCorps members to record all alternate interventions a student received in the AIMSweb program (see **Table III.5** for a list of alternative interventions). This record included the name of the intervention, when it was received and for how long. The Site Liaisons asked Internal Coaches to provide descriptions of each of the alternate programs offered at the school in order to determine if they could impact literacy proficiency. To counteract the effects of these alternative interventions on student outcomes, the participant-level measures were included as covariates in the statistical models used to estimate program impacts in the Fall analysis.

---

[24] AIMSweb is the web-based data management program utilized by the MRC program.

IV. Fall-Winter Experimental Study Findings

This section of the report presents the findings from the analysis of 16 weeks of assessment data collected on the program and control groups during the first semester of the 2012-2013 school year from the September 2012 Fall benchmark through the January 2013 Winter benchmark. As discussed earlier, the program and control groups were formed using randomization methods, which control for extraneous factors and result in groups that are considered statistically equal. Also, during this 16 week period, the control group students were barred from receiving any MRC services until after the Winter benchmark, so the statistical analysis presented below is a comparison of students receiving MRC services to a control group of students matched on Fall benchmark scores who received no services.[25]

To address the first research question (RQ1) of whether the MRC program had an impact on students' literacy outcomes, the evaluation team compared the assessment scores of program students who received MRC tutoring to those of a control group that did not receive MRC services during the first semester of the 2012-2013 school year (i.e., between Fall benchmark in September and Winter benchmark in January). As discussed in Chapter III, the evaluation team estimated the average treatment effect by developing a single model for each grade (Kindergarten, first, second, and third). To establish the average treatment effect, we calculated predicted scores by grade every three weeks of the program period using the weekly assessment scores, and we also provide estimated differences between program (treatment) and control, along with standard errors, confidence intervals and statistical tests.

Below the evaluation team provides by grade the findings from the final week of first semester assessment data collected during week 16. In order to more succinctly present these results in the figures below, we provide predicted score values for every third week during the Fall semester (i.e., weeks, 1, 4, 7, 10, 13 and 16) for both the program and control groups. Following the presentation of findings for all students, we provide the results of our subgroup analysis for gender, race, DLL status, and FRPL status. We conclude this chapter with the presentation of results from our examination of student impacts by AmeriCorps member characteristics/demographics and school assignment. More detailed tables, including scores for all 16 weeks of assessment data, are provided in **Appendix B.1**.

---

[25] As described in Chapter III, some schools provided other supplemental reading tutoring services similar to MRC services to a subset of students in the program and/or control groups; however, the MRC program did not provide any services to the control group students until after Winter benchmark.

Overall Impact Findings

The MRC K-3 program had a statistically significant positive impact on K-3 students' literacy skills, with large effects among Kindergarten students, moderate effects among first grade students, and no or small effects among second and third grade students. Below we provide a discussion of the findings by grade, and **Table IV.1** presents a summary of the data.

Table IV.1.   Mean scores for all program and control students at week 16 by grade

	Assignment to Condition	Mean Score	SE	Effect Size	Sig.	Confidence Interval	
	Control	23.32	1.18			21.02	25.62
Kindergarten (N=359)	Program	40.54	1.98			36.66	44.42
	Difference	17.22	1.48	1.06	***	14.30	20.12
	Control	58.83	0.61			57.64	60.01
1st Grade (N=409)	Program	65.34	0.66			64.04	66.64
	Difference	6.51	0.77	0.37	***	5.00	8.02
	Control	49.76	0.72			48.36	51.16
2nd Grade (N=265)	Program	51.09	0.74			49.64	52.55
	Difference	1.33	0.77	0.08		-0.18	2.85
	Control	82.94	0.56			81.84	84.03
3rd Grade (N=308)	Program	85.02	0.54			83.96	86.08
	Difference	2.08	0.66	0.10	**	0.79	3.37

***<.001

**<.01

*<.05

What is the impact of the MRC program on Kindergarten students?

The MRC program had a significant and large impact on Kindergarten students' *letter sound fluency* scores between Fall benchmark (September 2012) and Winter benchmark (January 2013). By the end of the first semester, Kindergarten students receiving MRC tutoring produced an average of 17.2 more letter sounds correctly in a one minute period than Kindergarten students in the control group. This represents a large effect size of about 1.1 standard deviations. Furthermore, by the end of the first semester, the program Kindergarten students' average score of 40.5 letter sounds was well above the Winter benchmark of 21, whereas the control Kindergarten students' average score barely exceeded benchmark(23.3 letter sounds) by the end of the first semester.

Figure IV.1.   Mean scores on Kindergarten program and control students
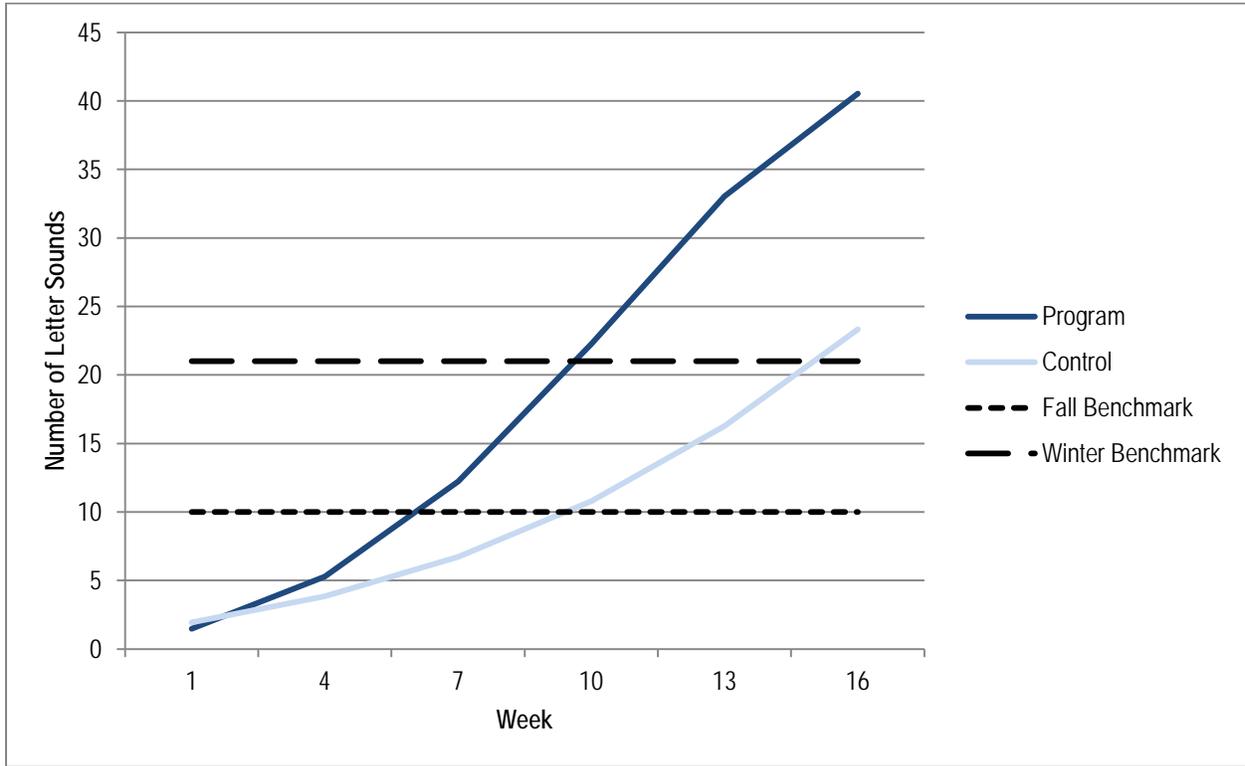
What is the impact of the MRC program on first grade students?

First grade students who received MRC tutoring attained significantly higher *nonsense word fluency* scores by Winter benchmark than did first grade students in the control group. By the end of the first semester, the MRC first grade students correctly produced in one minute an average of 6.5 more letter sounds embedded within non-real words than first grade control group students. This represents an effect size of about 0.37 standard deviations. By the end of the first semester, both the program and control first grade students' average scores (65.3 and 58.8, respectively) were above the Winter benchmark of 52 letter sounds within nonsense words.
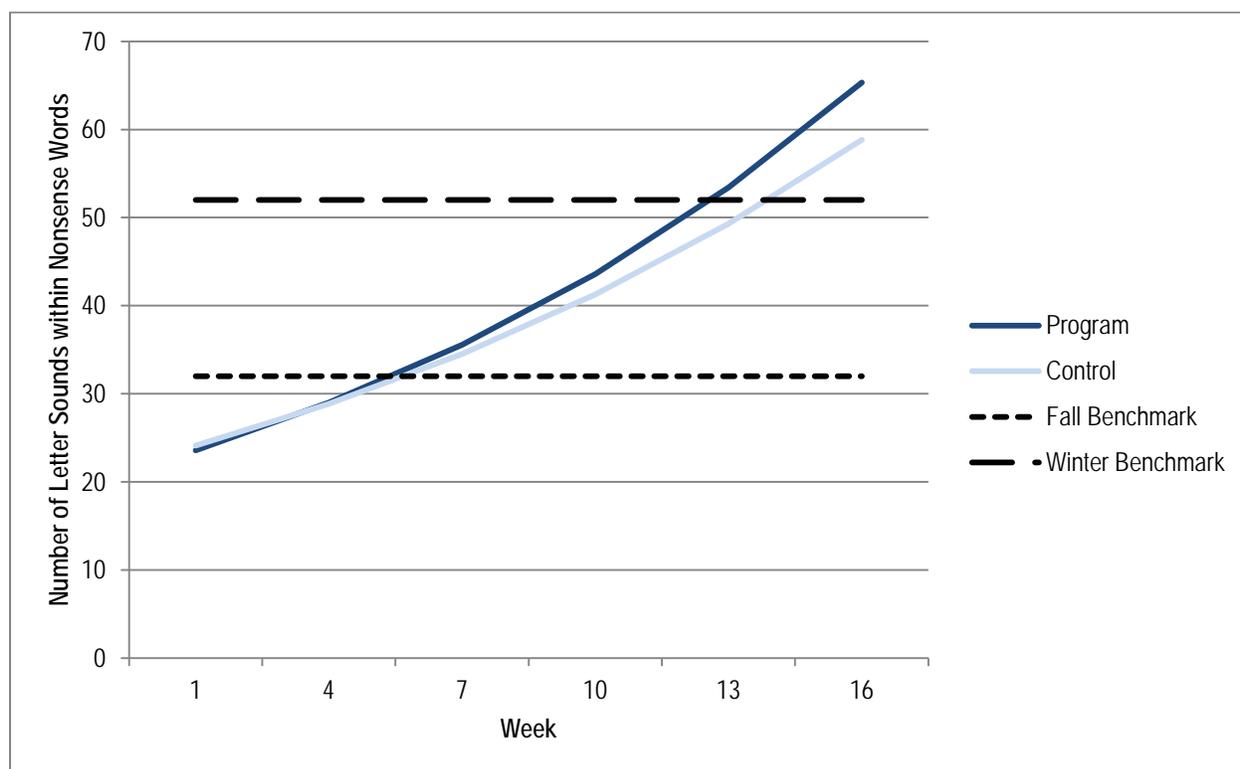
Figure IV.2. Mean scores on first grade program and control students

What is the impact of the MRC program on second grade students?

Second grade students who received MRC tutoring had significantly higher *oral reading fluency* scores for several, but not all, weeks of the first semester assessment data than did students in the control group. By the end of the first semester (week 16), second grade students who received MRC tutoring did not obtain high enough *oral reading fluency* scores compared to second grade students in the control group to maintain statistical significance. Despite the lack of significance, we found that second grade students in the MRC program group read aloud an average of 1.3 more words in one minute than their counterparts in the control group. The average scores of students in both the second grade program and control groups were well below the Winter benchmark of 73 word read aloud at the end of the first semester (51.1 and 49.8, respectively).

Figure IV.3.   Mean scores on second grade program and control students
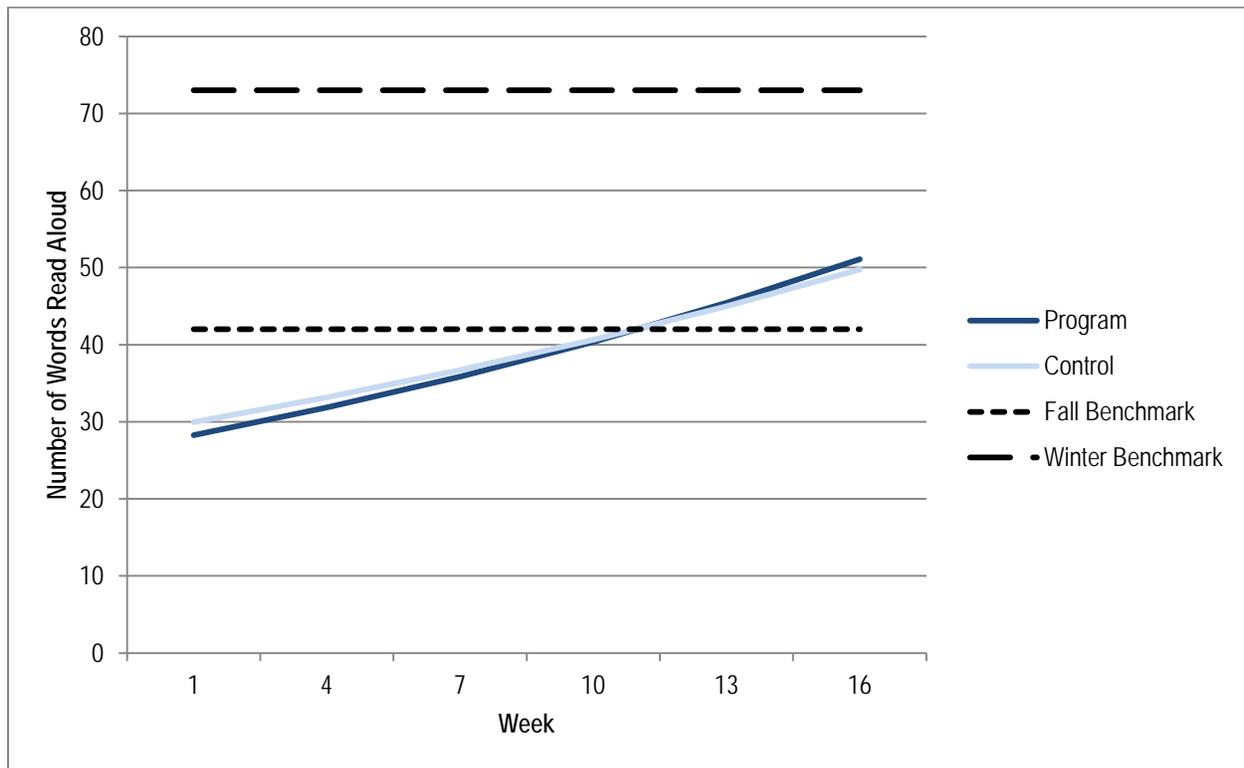
There are several possible explanations for the non-significant finding in second grade students at week 16. We explore these issues further in our findings from the Full Year Non-Experimental Study in Chapter V and address the lack of significance among second grade students in our conclusions section in Chapter VI.

What is the impact of the MRC program on third grade students?

Similar to the second grade students, third grade students who received MRC tutoring had significantly higher *oral reading fluency* scores for several, but not all, weeks of the first semester assessment data than did students in the control group. Among third grade students, however, students who received MRC tutoring obtained significantly higher *oral reading fluency* scores than students in the control group by the end of the first semester (week 16). Third grade students receiving MRC tutoring read aloud an average of 2.1 more words in a one minute period than third grade students in the control group. This represents a small effect size of about 0.10 standard deviations. Neither the program nor control group third grade students' average scores (85.0 and 82.9, respectively) reached the Winter benchmark of 91 words read aloud by the end of the first semester.

Figure IV.4.   Mean scores on third grade program and control students
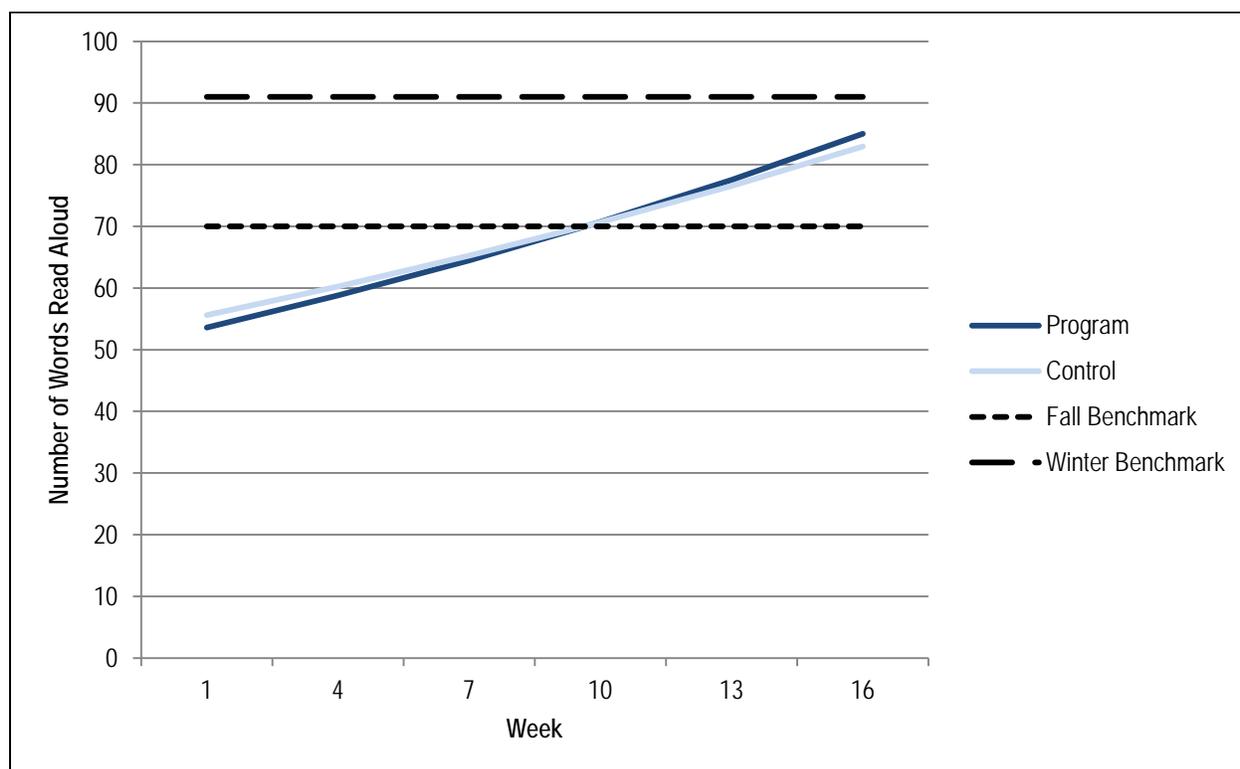
Findings by Major Demographic Groups

Our findings include a set of subgroup analyses of student study participants. The evaluation team identified subgroups for further examination based on the major demographic groups generally of interest in education research and also of interest to CNCS and the MRC program. We provide results by gender, race, Dual Language Learner (DLL) status, and Free and Reduced Price Lunch (FRPL) eligibility. Subgroups were defined based on characteristics provided at baseline and each level is determined by individual characteristics (e.g., male vs. female; race and ethnicity; DLL vs. non-DLL; FRPL vs. non-FRPL). In addition to examining subgroup differences by student demographic characteristics, we also present program effects for subgroups defined by a student's proximity to the Fall benchmark (baseline measure) prior to MRC tutoring. The purpose of this additional subgroup analysis is to examine how students closer and farther away from benchmark are impacted by the program.

Below the evaluation team provides two types of subgroup analysis. First, we conducted statistical tests (Chi-Square statistics to test the effect of moderated difference in difference coefficients) to determine if there are differential impacts across subgroups, implying that the MRC program may be more or less beneficial for selected subgroups of

students.[26] Results from the first set of subgroup statistical tests are provided in **Table IV.2** below. The results show statistically significant effects for most subgroup variables by grade.

Table IV.2. Chi-Square Test Results for Subgroup Variable Moderator Effects

Grade	Gender	Race	White/Non-White	DLL Status	FRPL Status
Kindergarten	15.49*** (N=359)	77.35*** (N=359)	0.90 (N=359)	16.24*** (N=359)	26.36*** (N=354)
Grade 1	1.71 (N=409)	10.03* (N=409)	6.26* (N=409)	1.47 (N=409)	1.20 (N=403)
Grade 3	26.84*** (N=308)	16.12** (N=305)	1.80 (N=305)	8.04** (N=308)	0.03 (N=301)

***<.001
**<.01
*<.05

For the second subgroup analysis, we subdivided the data by grade and by major subgroup (e.g., females, males, etc.), estimated the differences between program and control groups, and conducted significance testing to examine differences in patterns of findings by major subgroup.[27] The results from this second subgroup analysis should be interpreted with caution, in particular when the chi-square test is not significant, because it is possible that the differences found among certain subgroups are due to natural differences at baseline (prior to MRC tutoring) rather than due to actual program effects. When possible, we presented evidence of baseline differences to further explain potentially spurious program effects. The results of both subgroup analyses are presented below.

Gender

Both types of subgroup analyses found differences in the patterns of impacts by grade for females and males. The chi-square tests presented in **Table IV.2** above showed significant differences in effects by gender in both Kindergarten and third grade. When the sample was segmented by gender, significant findings persisted for Kindergarten, first, and third grade students for both males and females (p<.001). However, **Table IV.3** shows that the average difference between the treatment and control groups in week 16 were generally larger for males than females in Kindergarten and first grade and substantially larger in third grade. Most notably, by the end of the first semester, the average difference between program and control groups for third grade male students was substantially higher than the impact found in the general population. In sharp contrast, the findings showed significantly lower *oral reading fluency* scores for third grade female students who received MRC tutoring compared to female students in the third grade control group. While *all* third grade students receiving MRC tutoring read aloud

[26] Because significant findings were not found among second grade students in the main analysis of all students, subgroup analyses were not conducted for second grade students.

[27] It is important to note that sub-sectioning the data results in much smaller sample sizes; thus, significantly reducing power to detect statistically significant differences. Importantly, the study was not designed to conduct these analyses with sufficient power.

an average of 2.1 more words correctly in a one minute period (p<.01) by the end of the first semester, among third grade males the average difference between the program and control groups was 6.5 words (p<.001).

To further examine possible causes for these gender differences, we reviewed the average Fall benchmark scores (baseline measures) for third grade students by gender. We found that third grade males who received MRC tutoring had a much lower average Fall benchmark score than their female counterparts (46.5 for males and 51.4 for females), while the average score for third grade males and females in the control group was virtually the same (52.8 for males and 52.5 for females). These baseline differences suggest that the different average gains by gender may simply be due to the fact that male third grade students had further to grow than the female students. Thus, the evaluation team concluded that the larger average difference found among third grader males is likely due to this average difference prior to MRC tutoring.

Table IV.3.   Mean scores for program and control students at week 16 by grade and gender

Grade	Assignment to Condition	Male students			Female Students		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten (N=359)	Control	21.31	1.29		25.02	1.41	
	Program	41.38	2.34		39.75	2.22	
	Difference	20.07	2.06	***	14.73	1.92	***
1st Grade+ (N=409)	Control	59.74	0.86		58.10	0.76	
	Program	66.95	0.90		63.66	0.88	
	Difference	7.21	1.15	***	5.57	1.07	***
3rd Grade (N=308)	Control	80.55	0.72		85.31	0.76	
	Program	87.04	0.67		81.91	0.77	
	Difference	6.50	0.90	***	-3.41	1.00	***

+ The chi-square test for gender differences in impacts was not significant among first grade students.

Notes: Kindergarten sample includes 159 male and 200 female students, 1st grade sample includes 195 male and 214 female students, 3rd grade sample includes 165 male and 143 female students. ***<.001, **<.01, *<.05.

Race

For the analysis by race, we present two different subgroupings of the data: 1) by the three major racial groups, White, Black, and Asian;[28] and 2) a comparison between White and non-White groups, which allowed us to collapse all minority groups into one subgroup category. In our chi-square tests, where we examined whether race is a significant moderator of program impacts, we found race to be a significant factor in explaining variations in student outcomes (p<.001). Consistent with these findings, some differences were found in the patterns of impacts by grade for the three major racial groups when we subgrouped the sample into White, Black, and Asian and compared these findings to the analysis of all students. For example, **Table IV.4** shows that although large effects were found across

---

[28] Due to very small sample sizes, we were unable to conduct subgroup analysis for Hispanic students and for students in the "other" racial/ethnic category.

all Kindergarten students, the effects for Asian Kindergarten students were more than double the size of that found in the general population. Asian Kindergarten students in the program group produced 38.2 more letter sounds in a one minute testing period compared to the control group (<.001) by the end of the first semester. While still producing large gains, differences between the program and control groups for White and Black students were 11.1 letter sounds (<.001) and 13.6 letter sounds (<.001), respectively.

Among first grade students, Asian and Black students attained higher *nonsense word fluency* scores by the end of the first semester than those produced in the analysis of all students. While White students receiving MRC tutoring produced 1.9 more letter sounds embedded within non-real words in a one minute testing period than their first grade control group counterparts, the difference between Black program and control first grade students was 11.2 letter sounds (<.001) and between Asian program and control first grade students was 7.4 letter sounds (<.001). Thus, larger impacts of the MRC program were found for Black and Asian students in first grade compared to their White counterparts.

The analysis of third grade students' *oral reading fluency* showed that although significant differences were detected in the analysis of all students and among White students (p<.01), differences between the program and control groups in the number of words read aloud in a one minute testing period were not significant among Black and Asian students.

Table IV.4.   Mean scores for program and control students at week 16 by grade and select racial groups

Grade	Assignment to Condition	White students			Black students			Asian students		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten (N=359)	Control	24.00	1.65		18.94	1.36		26.14	1.77	
	Program	35.13	2.09		32.55	2.18		64.32	4.46	
	Difference	11.14	2.09	***	13.61	2.05	***	38.18	3.97	***
1st Grade (N=409)	Control	61.62	0.97		58.80	1.23		59.06	1.12	
	Program	63.55	0.94		69.98	1.54		66.50	1.16	
	Difference	1.93	1.20		11.18	1.87	***	7.44	1.45	***
3rd Grade (N=305)	Control	85.39	0.76		81.43	1.15		80.56	0.86	
	Program	88.16	0.71		83.85	1.08		81.01	0.84	
	Difference	2.77	1.02	**	2.42	1.56		0.46	1.17	

Notes: Kindergarten sample includes 107 White, 119 Black, and 97 Asian students; 1st grade sample includes 161 White, 88 Black, and 108 Asian students; 3rd grade sample includes 121 White, 71 Black, and 81 Asian students. ***<.001, **<.01, *<.05.

In addition to the analysis by racial category, the evaluation team also separated the sample into White and non-White.[29] subgroups and conducted similar analyses. This analysis allowed for the inclusion of racial subgroups whose

---

[29] Non-White students include Black, Asian, Hispanic, and "other" student racial categories.

sample sizes were too small (i.e., Hispanic and "Other") to include in the more detailed racial subgroup analysis presented above. The chi-square tests presented in **Table IV.2** above showed significant differences in effects by race only among the first grade students (p<.05). Consistent with this finding, **Table IV.5** shows that a significant effect persists for non-White first grade students, but not for White students. In fact the difference between program and control groups were much larger among non-White first grade students than those found among White first grade students. While not significant in the chi-square analysis, the subgroup analysis shows that the average impact on non-White Kindergarten students was almost double the size of that found in the White Kindergarten students. Specifically, an average of 20.5 more letter sounds were produced in a one minute testing period among non-White Kindergarten students in the program group (p<.001) compared to 11.0 more letter sounds among White Kindergarten students in the program group (p<.001). In first grade, the differences were even more pronounced with non-White first grade students in the program group producing on average 9.7 more letter sounds imbedded within nonsense words than the control group. However, we found that Kindergarten and first grade non-White students had much lower average Fall benchmark scores (approximately 2.7 for Kindergarten and 20 for first grade) than their White counterparts (approximately 9 for Kindergarten and 25 for first grade). Therefore, the differences found among Kindergarten and first grade students at 16 weeks are likely affected by differences at baseline.

Table IV.5.   Mean scores for program and control students at week 16 by grade and White and non-White racial group

Grade	Assignment to Condition	White			Non-White		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten+ (N=359)	Control	24.57	1.73		22.70	1.27	
	Program	35.59	2.18		43.24	2.42	
	Difference	11.02	2.18	***	20.54	1.92	***
1st Grade (N=409)	Control	61.79	0.97		56.92	0.73	
	Program	63.70	0.94		66.59	0.88	
	Difference	1.92	1.22		9.67	1.00	***
3rd Grade+ (N=305)	Control	85.34	0.76		80.58	0.64	
	Program	88.14	0.71		82.00	0.60	
	Difference	2.80	1.02	**	1.42	0.85	

+ The chi-square test for White/non-White differences in impacts was not significant among Kindergarten and third grade students.

Notes: Kindergarten sample includes 107 White and 252 non-White students; 1st grade sample includes 161 White and 248 non-White students; 3rd grade sample includes 121 White and 184 non-White students. ***<.001, **<.01, *<.05.

Dual Language Learners

Compared to the analysis of all students, some important differences also were detected in the patterns of findings by grade for DLL.[30] vs. non-DLL students. These differences in subgroup findings are supported by our chi-square

---

[30] It is important to note that Asian students and DLL status are highly correlated in this data set.

analysis showing DLL status to be a significant moderator of program impacts on Kindergarten and third grade student outcomes (p<.05 or lower depending on grade). While significant findings persisted for both DLL and non-DLL students in Kindergarten in week 16, the impact of the program on DLL students in Kindergarten was substantially larger than those found for non-DLL students (see **Table IV.6** below). By the end of the first semester, Kindergarten DLLs receiving MRC tutoring produced an average of 35.1 more letter sounds correctly in a one minute period than Kindergarten students in the control group (p<.001). In contrast, among non-DLL Kindergarten students, the average difference between the program and control groups was 13.2 letter sounds (p<.001). Furthermore, the impact of the MRC program for DLL students was far larger than found in the whole sample analysis, which showed a difference of 17.2 letter sounds between Kindergarten students in the program and control groups (p<.001).

In addition to Kindergarten students, substantial differences in impacts for DLL students were also found among students in first grade. While the chi-square test indicated that DLL status is not a significant moderator of program effects among first grade students, the individual subgroup analysis presented in **Table IV.6** suggests differences in program effects. Non-DLL first grade students demonstrated an average difference between the program and control groups of 4.5 more letter sounds in a one minute period (p<.001), while first grade students with DLL status who received MRC tutoring produced an average of 10.0 more letter sounds correctly than DLL first grade students in the control group (p<.001). However, these differences in the individual subgroup analysis may be due to baseline differences between DLL and non-DLL first grade students.

In contrast to the large gains found among DLL Kindergarten and first grade students, there were no significant findings for third grade DLL students.

Table IV.6.   Mean scores for program and control students at week 16 by grade and Dual Language Learner status

Grade	Assignment to Condition	Non-Dual Language Learner			Dual Language Learner		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten (N=359)	Control	22.23	1.20		26.66	1.79	
	Program	35.40	1.80		61.76	4.29	
	Difference	13.17	1.47	***	35.11	3.88	***
1st Grade+ (N=409)	Control	59.73	0.74		57.24	0.93	
	Program	64.23	0.80		67.18	1.06	
	Difference	4.50	0.98	***	9.95	1.28	***
3rd Grade (N=308)	Control	83.68	0.63		80.45	0.85	
	Program	86.70	0.58		80.37	0.83	
	Difference	3.02	0.81	***	-0.08	1.11	

+ The chi-square test for gender differences in impacts was not significant among first grade students.

Notes: Kindergarten sample includes 264 non-Dual Language Learners and 95 Dual Language Learners; 1st grade sample includes 263 non-Dual Language Learners and 146 Dual Language Learners; 3rd grade sample includes 213 non-Dual Language Learners and 95 Dual Language Learners. ***<.001, **<.01, *<.05.

Free and Reduced Price Lunch

As with the DLL status subgroup analysis, some important differences also were detected in the patterns of findings by FRPL eligibility. FRPL eligibility is a commonly used indicator of socio-economic status, as it is based on household income. The chi-square test results presented in **Table V.2** above showed significant differences in effects by FRPL eligibility only among the Kindergarten students (p<.001). However, **Table IV.7** shows that, when students were subgrouped by FRPL eligibility, differences in estimated impacts in both Kindergarten and first grade appear to be substantial. Among Kindergarten students, students eligible for FRPL had impacts almost four times those found for students who were not eligible for FRPL, and significant differences were found among first grade students eligible for FRPL, but not among ineligible students. Impacts for third grade students eligible for FRPL and ineligible for FRPL were both significant, but also similar in size.

Table IV.7   Mean scores for program and control students at week 16 by grade and Free and Reduced Price Lunch eligibility

Grade	Assignment to Condition	Non-FRPL			FRPL		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten (N=354)	Control	31.81	2.54		20.94	1.13	
	Program	37.33	2.34		41.54	2.28	
	Difference	5.52	2.82		20.60	1.81	***
1st Grade+ (N=403)	Control	61.94	1.10		57.52	0.69	
	Program	62.74	1.04		66.47	0.80	
	Difference	0.80	1.43		8.95	0.94	***
3rd Grade+ (N=301)	Control	84.04	0.91		82.11	0.64	
	Program	86.67	0.85		84.05	0.61	
	Difference	2.63	1.18	*	1.95	0.81	*

+ The chi-square test for gender differences in impacts was not significant among first and third grade students.

Notes: Kindergarten sample includes 85 non-FRPL and 269 FRPL students; 1st grade sample includes 112 non-FRPL and 291 FRPL students; 3rd grade sample includes 87 non-FRPL students and 217 FRPL students. ***<.001, **<.01, *<.05.

Proximity to Benchmark

In addition to examining subgroup differences by student demographic characteristics, the evaluation team also examined differences in program effects by student proximity to the Fall benchmark (baseline measure) prior to MRC tutoring. For this analysis, we divided students into three equal sized groups within each grade: low (farthest from) baseline, middle baseline, and high baseline. **Table IV.8** presents the findings from our baseline subgroup analysis, which showed differences in student impacts within grade by proximity to Fall benchmark.

Significant and large differences in program impact were observed on average among Kindergarten students who had low and middle baseline scores. In contrast, the program impacts for students in the high baseline group were small and not significant. However, due to their closer proximity to benchmark prior to program entry, these high baseline students were also more likely to reach the Winter benchmark (outcome) earlier and exit the MRC program

sooner. More importantly, the control group students on average in the middle and high baseline subgroups reached the Kindergarten Winter benchmark of 21 letter sounds without MRC tutoring. The largest impact of the program on benchmark achievement was observed in the low baseline group where the control group students' average score was far behind benchmark and the program students on average exceeded the 21 letter sound Winter benchmark after 16 weeks.

Similar to Kindergarten, significant differences were observed among first grade students who had low and middle baseline scores, but not high baseline scores. Again, the high baseline students were more likely to reach the Winter benchmark earlier and exit the program sooner due to their baseline proximity to Fall benchmark. The control group students on average in both the middle and high baseline subgroups exceeded the Winter benchmark of 52 nonsense words without MRC intervention. As with the Kindergarten students, the major difference in Winter benchmark achievement was observed in the low baseline group where the control group students were far behind Winter benchmark and the program group students on average reached the 52 word Winter benchmark after 16 weeks.

Significant differences were found among the low and high baseline subgroups for third grade students. However, only the program students in the high baseline subgroup on average achieved the third grade Winter benchmark (91 words read aloud).

Table IV.8   Mean scores for program and control students at week 16 by grade and proximity to Fall benchmark (baseline)

Grade	Assignment to Condition	Low Baseline			Middle Baseline			High Baseline		
		Mean Score	SE	Sig.	Mean Score	SE	Sig.	Mean Score	SE	Sig.
Kindergarten (N=359)	Control	8.00	0.56		24.01	1.46		37.08	2.40	
	Program	25.09	1.63		36.85	2.18		40.97	2.49	
	Difference	17.09	1.48	***	12.84	2.10	***	3.89	2.73	
1st Grade (N=409)	Control	48.70	0.94		61.39	1.06		64.51	1.19	
	Program	63.58	1.23		63.83	1.09		67.71	1.23	
	Difference	14.87	1.23	***	2.43	1.20	*	3.20	1.38	*
3rd Grade (N=308)	Control	65.35	1.10		89.00	1.18		90.67	1.36	
	Program	69.83	1.15		87.24	1.14		95.30	1.29	
	Difference	4.48	1.14	***	-1.76	1.16		4.64	1.36	***

Notes: Kindergarten sample includes 116 Low, 134 Middle, and 109 High Baseline students; 1st grade sample includes 129 Low, 154 Middle, and 126 High Baseline students; 3rd grade sample includes 89 Low, 113 Middle, and 106 High Baseline students. ***<.001, **<.01, *<.05.

AmeriCorps Member and School Level Effects

An additional analysis of interest to CNCS is the examination of whether student assessment scores varied as a function of the AmeriCorps member to whom they were assigned for tutoring or the school where they received tutoring as an indicator of program consistency. For this analysis, the evaluation team was only able to examine the assessment scores of students assigned to the program group because control group students did not receive tutoring. We examined this question using two different approaches: the calculation of member- and school-level intraclass correlations and an analysis of member characteristics as covariates within a regression model. First, while conducting our baseline analysis, we calculated the member- and school-level intraclass correlations (ICC) by grade using students' assessment scores at 16 weeks (i.e., at the end of the first semester) to determine whether the student outcome data indicated substantial member- or school-level effects, which if detected would need to be incorporated into our analysis models. Intraclass correlations measure the portion of the total variation in an outcome that is associated with various levels of a variable, such as member or school assignment. In this case, the member-level ICC measures the degree to which different students assigned the same member correlate on their program impacts (i.e., difference between Winter and Fall benchmark score). Next, the school-level ICC measures the degree to which students within the same school correlate on their impacts.

As shown in **Table IV.9**, both the member- and school-level ICCs are relatively low and present high standard errors of the estimates. The member-level ICCs for program impacts vary by grade, with the largest ICC among Kindergarten students (0.19).[31] and the smallest among first grade students (0.00). However, in each case, the standard error of the estimate approaches the same size as the estimate itself, effectively nullifying the ICC measures. We find a similar case with the school-level ICCs, again with the largest ICC among Kindergarten students (0.28) and the smallest among first grade students (0.05). However, the variance of these estimates is again so large that reliable inference is not possible.

Table IV.9.   AmeriCorps member- and school-level interclass correlations and standard errors based on student-level program effects (Winter-Fall benchmark) by grade

Grade	Member-level ICCs	SE for member-level ICCs	School-level ICCs	SE for school-level ICCs
Kindergarten (N=129)	0.19	0.18	0.28	0.21
1st Grade (N=174)	0.00	0.00	0.05	0.05
2nd Grade (N=108)	0.04	0.12	0.14	0.13
3rd Grade (N=128)	0.03	0.10	0.08	0.08

---

[31] High member- and school-level ICCs in Kindergarten may be due to differences in the presence of the K-focus intervention, which is generally implemented within an entire school.

Even though the within-grade member-level ICCs were presumed to be zero, the evaluation team conducted a second analysis to determine if any member characteristics (e.g., age, education, gender, race, prior education experience) were significant contributors to differences in student assessment scores. We standardized and pooled the data across grades and introduced member characteristics into the model as covariates. For this analysis, the evaluation team examined AmeriCorps members' years of education and whether they had a degree in an education-related field. In addition, we compared new (first year) vs. experienced members, part-time vs. full-time AmeriCorps service, and also examined members by race and gender characteristics. One additional variable of interest was whether matching members and students by similar gender (e. g., female member and female student) resulted in better student outcomes compared to mixing genders. **Table IV.10** provides our findings which show that none of the member characteristics resulted in significant differences in student impacts, implying that individuals with varying backgrounds can produce similar effects on students' literacy outcomes while serving in the MRC program.

Table IV.10.   Effects of AmeriCorps member characteristics on program students' weekly assessment scores

	Effect	SE(Effect)
Member age	0.00	0.00
Member has education degree	0.05	0.09
Member years of education	0.00	0.02
Member's first year	0.07	0.08
Member is part time	0.02	0.09
Member is Black	0.09	0.29
Member is Asian	0.04	0.21
Member is female	0.20	0.10
Student is female	0.13	0.05
Student is female × member is female	0.10	0.07
Observations		
Weeks	7,381	
Students	587	
Members	53	
Schools	22	

Notes: Other controls include student benchmark and week of assessment; outcome is standardized

***<.001

**<.01

*<.05

V. Full Year Non-Experimental Study Findings

This section of the report presents the findings from the analysis of the full year of assessment data collected on the program and control groups during both semesters of the 2012-2013 school year. As discussed above, after the Fall semester, all students in the control group became eligible for MRC tutoring services, were reassessed on the Winter benchmark in January, and then, if found eligible, could begin receiving MRC services in the second semester. At some schools, members were not able to serve all eligible students due to full tutoring case loads. Therefore, some control students never received MRC tutoring at any point during the 2012-2013 school year. Because the control group students no longer were embargoed from receiving MRC services after the first semester, the program and control groups no longer are considered statistically equal for the Full-Year non-experimental analysis presented in this chapter. For this reason, the results of the Full-Year non-experimental analysis should be interpreted with caution.

Also, in the second semester, students closest to the Winter benchmark were given priority entry into the program compared to those further from the benchmark. Therefore, students' participation in MRC tutoring in the second semester was highly correlated with their benchmark assessment scores. Given the bias possibly created by the nonrandom method of program enrollment during the second semester and the correlation between scores and receipt of services, the results presented in this chapter may be subject to differing interpretations.

While the first research question on the impact of the MRC program on student literacy outcomes was addressed in Chapter IV, the second research question, which address both week to week and cumulative impacts of the program, are addressed in this chapter using the full-year of assessment data. To address the second research question (RQ2) as to whether the MRC program impacts vary week to week and at what point the number of cumulative weeks of intervention (i.e., dosage) continue to show substantial impacts on student literacy outcomes, the evaluation team examined the impact of the program on student assessment scores for each week of tutoring received (irrespective of assignment to program or control group) over 16 weeks of tutoring sessions for Kindergarten and first grade students and 24 weeks of sessions for second and third grade students. We determined the number of weeks to include in the grade-specific analysis by examining both the average number of weeks of tutoring students received within each grade (provided in **Table V.1**) and for first grade, the amount of time during which assessment data was collected (i.e., among first grade students, *nonsense word fluency* and *oral reading fluency* are assessed for only one semester each).

Table V.1. Means and standard deviations (in parentheses) of number of weeks of tutoring by program assignment and semester

	First Semester[1]		Second Semester[2]	
	Control	Program	Control	Program
Grade (N, First Semester, Second Semester)				
Kindergarten (359. 356)	0 (0)	10.13 (2.91)	2.84 (4.66)	4.64 (6.28)
1st Grade (409, 406)	0 (0)	9.35 (2.71)	1.65 (4.44)	7.50 (8.47)
2nd Grade (265, 263)	0 (0)	11.02 (2.03)	3.00 (5.50)	13.45 (7.32)
3rd Grade (308, 301)	0 (0)	9.91 (2.89)	3.34 (6.41)	10.80 (8.59)

[1]Fall 2012 to Winter 2013; [2]Winter 2013 to Spring 2013

For this analysis, the average weekly gain in grade and semester specific literacy assessment scores was estimated by calculating the rate of change from week to week in predicted scores by grade using assessment data collected over the entire school year. These models controlled for semester effects, seasonal benchmark scores, and average growth as a function of time. Thus, the model controls for prior ability and typical growth, leaving the effect of tutoring sessions net of typical growth for the average student in each grade..[32] In the reported week to week differences, only the assessment scores of students who received at least some MRC tutoring, whether originally assigned to the program or control group, were included in the estimated gains analysis. Since the first grade literacy metric changed between the first and second semesters (from *nonsense word fluency* to *oral reading fluency)*, we report separate analyses for each semester of first grade instead of a full year analysis as for students in other grades. More detailed tables, including weekly gains for all weeks of assessment data, are provided in **Appendix B.1**.

Week Over Week and Cumulative Effects of MRC Program

Students participating in the MRC K-3 program showed differences by grade in patterns of week over week growth in reading proficiency. An important note is that these analyses examined the rate of change of the average growth rate itself. Throughout this section, and in our conclusions in Chapter VI, the rate of change of these growth rates are described in terms of whether they are increasing or decreasing. Because student scores tend to increase inherently over time, these findings do not indicate whether individual student scores themselves were increasing or decreasing. Rather, the findings tell us how the rate of growth, while controlling for average growth unrelated to the MRC program, increased or decreased week to week. For example, the average student could be growing each week, but could grow at a slower rate in week two than in week one. In this example, the average student's week over week difference in their rate of change would be negative, reflecting the deceleration in their growth rate.

---

[32] It is important to reiterate that this analysis models the effect of tutoring on an *average* student. Some students respond more rapidly to tutoring than others.

Kindergarten students showed large week over week gains in the average student growth rate, starting in the first week of tutoring, which then began to level off by week 8 and demonstrated a gradually decelerating pattern in later weeks. First, second, and third grade students showed slow, but steady week over week gains, which continued over the entire period of analysis (16 weeks for first grade and 24 weeks for second and third grade).

After calculating week over week gains in growth rates for students who received MRC tutoring, the evaluation team summed the week over week gains over time to demonstrate the cumulative effect of the program on the average student growth rate over time. Therefore, we also discuss our findings on the patterns of growth among students by grade over a 16 week (Kindergarten and first grade) or 24 week (second and third grade) period. For most students receiving MRC services, we observed consistent increases over time in weekly gains in growth rates.

What is the week over week and cumulative effect of the MRC program on Kindergarten students?

Figure V.1 below shows the average week over week changes in growth rates and cumulative gains (over all subsequent weeks) in *letter sound fluency* among Kindergarten students. As the figure demonstrates, Kindergarten students receiving MRC services produced an average of 2.7 more letter sounds correctly in a one minute period at the end of their first week of tutoring (i.e., growth between week 0 and week 1 of tutoring). Large week over week gains continue through week 4 with average gains of between 2.2 and 2.7 more letter sounds produced per week of tutoring. Students continued to show week over week gains in letter sounds up until week 12 when we begin to observe negative changes in week over week growth rate differences. Given the significant and large gains found in the Fall-Winter Experimental Study among Kindergarten students, these decreases are likely due to students not being able to maintain such high levels of week over week growth rate increases.

Figure V.1 demonstrates how the large week over week gains accumulate over the first three-quarters of the 16 week period and then begin to level off. The slight reduction in growth is likely due to students maximizing the *rate* at which they can continue to improve letter sound fluency over a minute period. It is important to note that even though the rate at which students improve their letter sound fluency levels off around week 12, students' overall letter sound fluency score continues to grow from one week to the next, culminating in observed cumulative gains (above expected average student growth) of over 16 letter sounds.

Figure V.1   Cumulative week over week growth in Kindergarten letter sounds for students receiving MRC tutoring
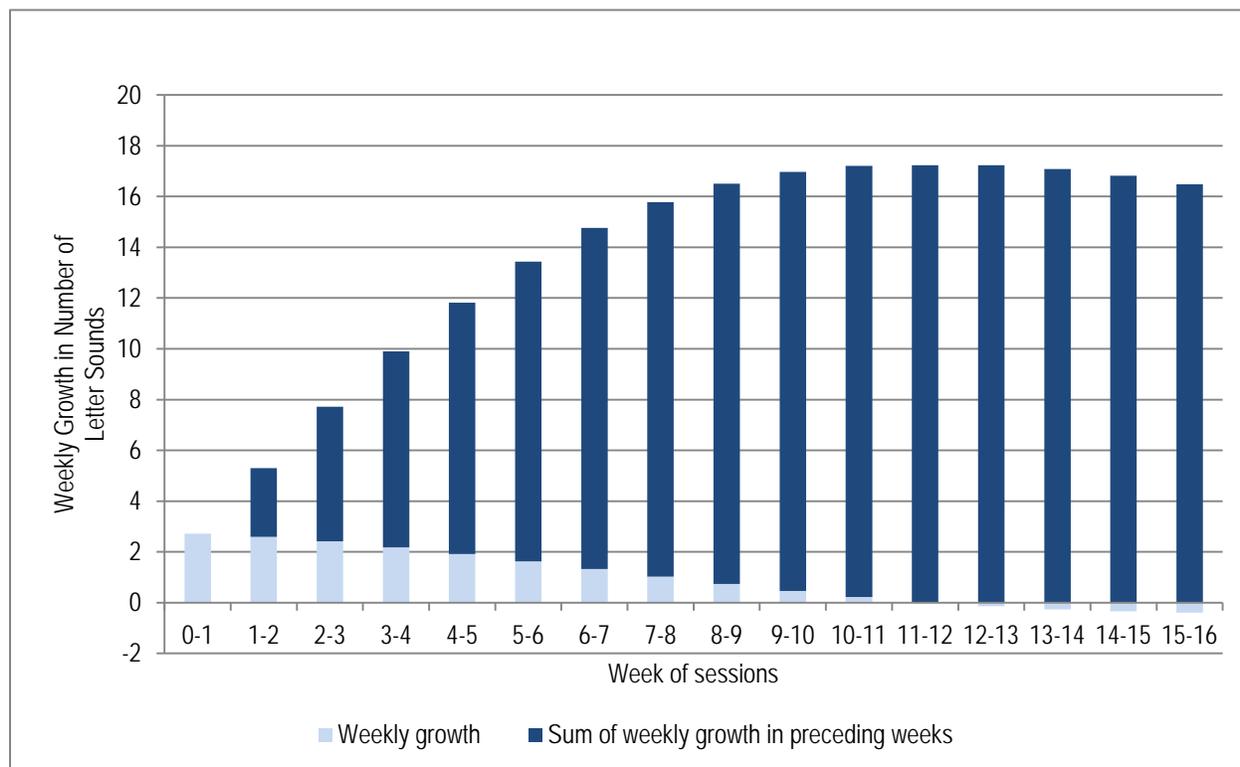
What is the week over week and cumulative effect of the MRC program on first grade students?
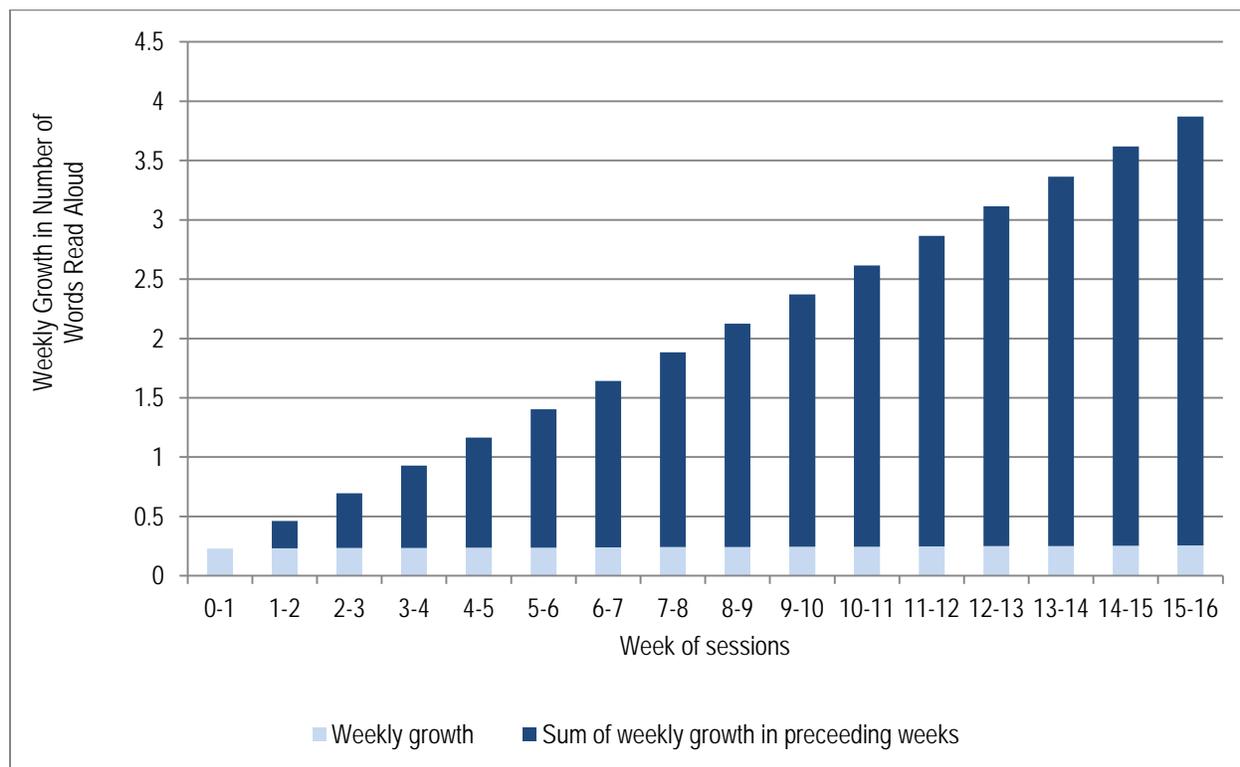
The average week over week gains in *nonsense word fluency* produced by first grade students receiving MRC tutoring, shown in **Figure V.2.**, remained relatively steady. First grade students produced average gains of between 0.5 and 0.6 more letter sounds embedded within non-real words in a one minute period throughout the 16 week period, resulting in continued gains in nonsense word fluency among first grade students each week. **Figure V.2** also shows that, despite slight differences in the amount of week over week gains, we continued to observe steady cumulative gains in *nonsense word fluency* among first grade students. By the end of the 16 week period, we observed cumulative gains (above expected average student growth) of more than 9 letter sounds embedded within non-real words in a one minute period among first grade students receiving MRC tutoring.

Figure V.2. Cumulative week over week growth in first grade nonsense word letter sounds for students receiving MRC tutoring in the Fall
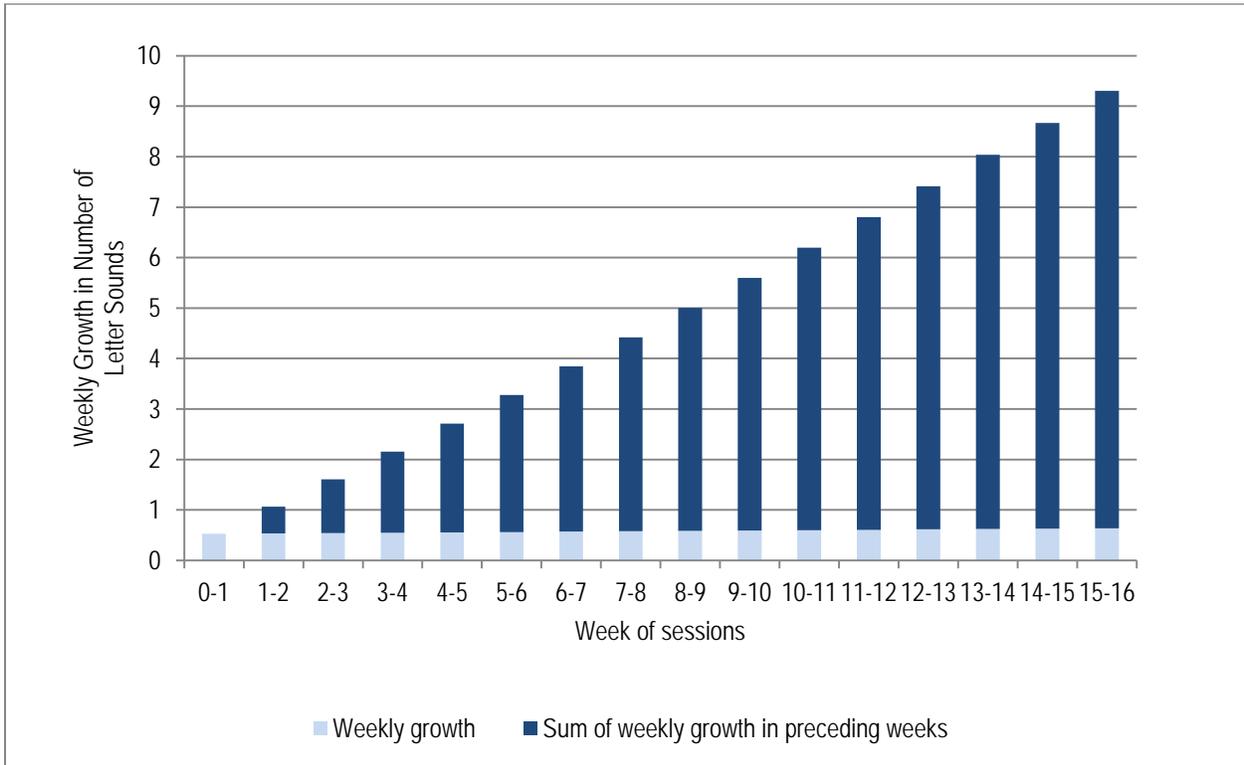
While *nonsense word fluency* is assessed on first grade students during the first semester of the school year, in the second semester, first grade students are assessed on *oral reading fluency*. **Figure V.3** shows the average week over week and cumulative growth in words read aloud by first grade students in the second semester. We observe small, but steady increases in the average rate of growth of about 0.2 words read aloud correctly from one week to the next across the entire 16 week period. These small week over week increases result in considerable cumulative growth over 16 weeks. Thus, we found that first grade students are making reasonable gains over time.
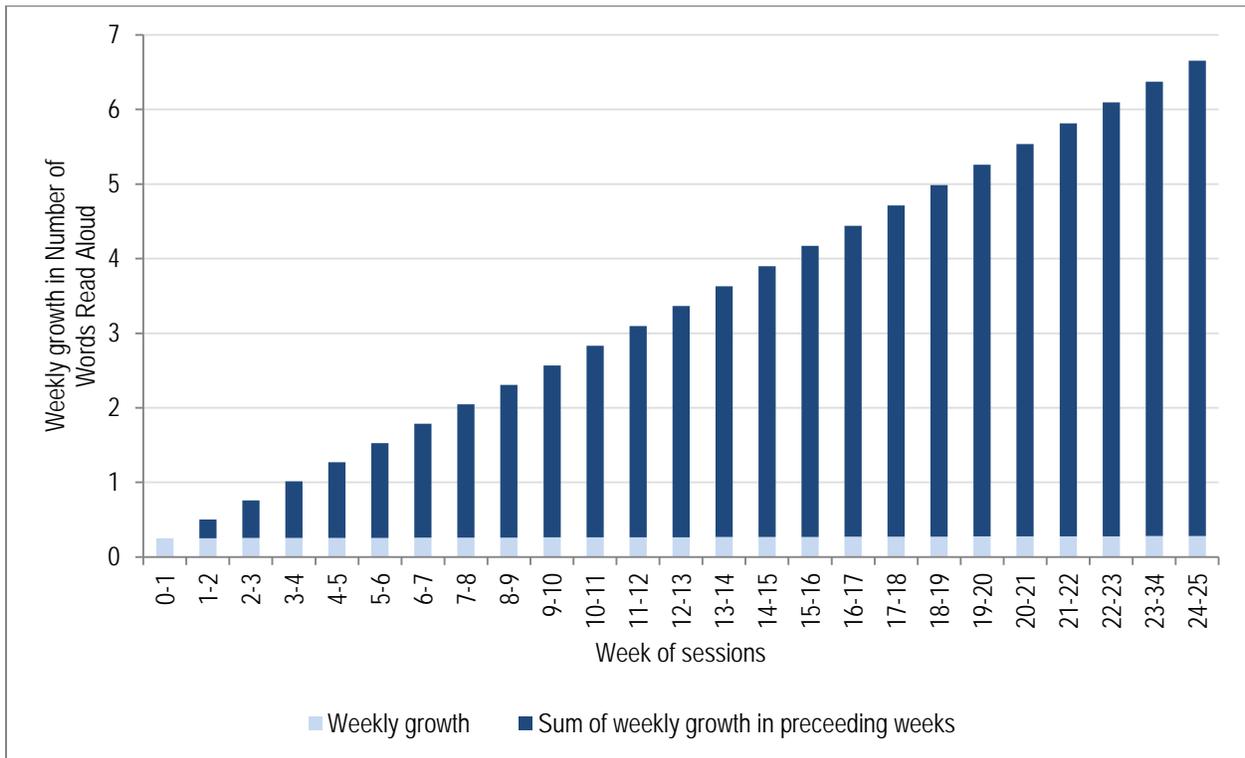
Figure V.3.   Cumulative week over week growth in first grade words read aloud for students receiving MRC tutoring in Spring

What is the week over week and cumulative effect of the MRC program on second grade students?

Similar to the findings for first grade *oral reading fluency*, **Figure V.4** shows small and incremental average week over week gains in growth rates produced by second grade students over a 24 week observation period. Second grade students receiving MRC services produced average week over week gains of about 0.2 more words read aloud. These incremental gains compounded over time, producing sizable growth by the end of the 24 week observation period. Given the small and non-significant effects found in the Fall-Winter Experimental Study among second grade students after 16 weeks of tutoring sessions, this pattern of week over week gains suggests that with more time, the study could have detected greater, and possibly significant, impacts of the MRC program on second grade students' *oral reading fluency* .

Figure V.4.   Cumulative week over week growth in second grade words read aloud for students receiving MRC tutoring

What is the week over week and cumulative effect of the MRC program on third grade students?

Consistent with the findings for *oral reading fluency* in first and second grade, **Figure V.5** shows third grade students producing positive week to week gains across the 16 weeks of tutoring. Steady week over week increases of 0.2 words read aloud continued throughout the study period. As with the second grade student findings, this pattern of week over week gains suggest that the small effects of the MRC program found on third grade students' *oral reading fluency* in the Fall-Winter Experimental Study may have been more substantial with more time to accumulate.

Figure V.5. Cumulative week over week growth in third grade words read aloud for students receiving MRC tutoring
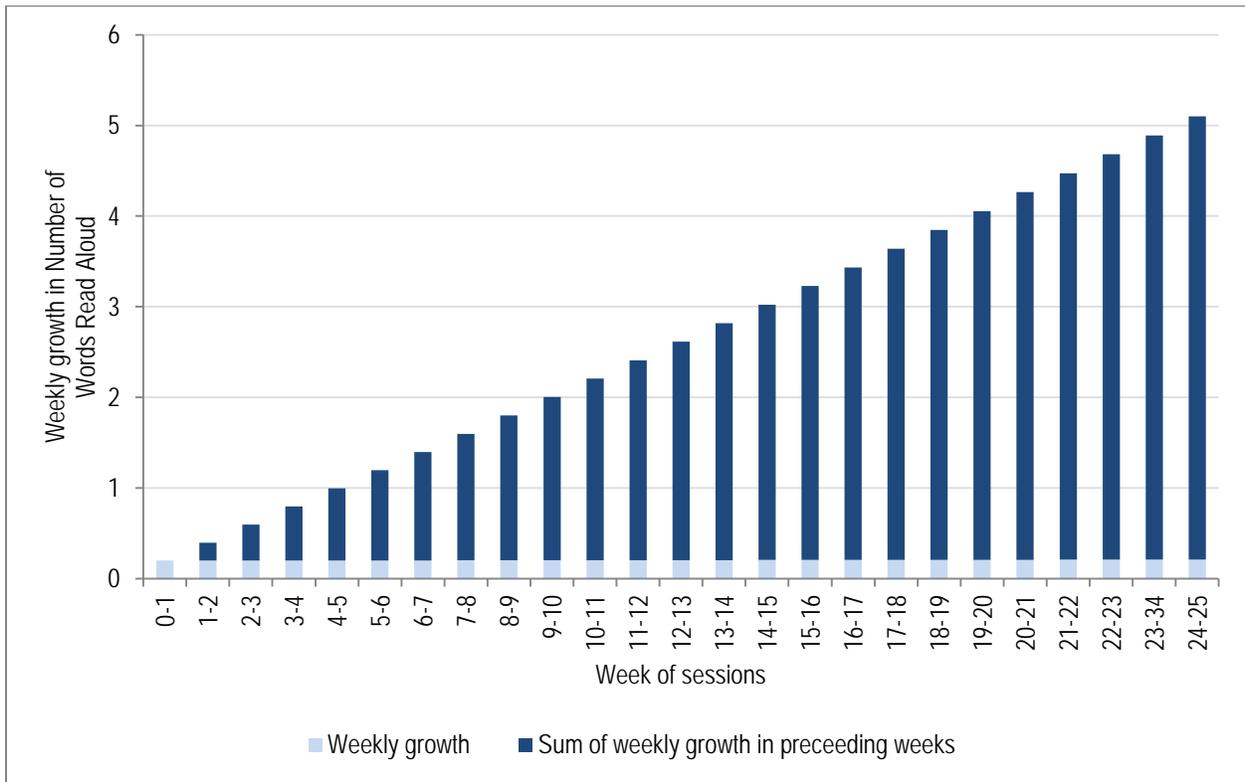
VI. Exploratory Analysis

To address the final research question, which focuses on the longer-term effects of the MRC program (i.e., end of the school year), the evaluation team conducted two exploratory analysis approaches using the full-year of assessment data to predict. The third research question (RQ3) focuses on the longer-term effects of MRC tutoring on students' proficiency levels by examining all students who received MRC services at any point throughout the school year, despite initial assignment to program or control groups.

In the first analysis, the evaluation team considered whether starting MRC tutoring earlier rather than later in the school year resulted in a higher likelihood of a consistent effect over time. It was possible for us to explore this question because program and control students tended to receive MRC services at different points in time (i.e., program students during the first semester vs. control students during the second semester). Therefore, for this analysis, we categorized students.[33] into three distinct groups according to their longer-term (i.e., end of school year) achievement trajectories:

1. Students showing no effect of the program for the entire school year (chronically below benchmark);
2. Students demonstrating a temporary program effect (moving above and below benchmark throughout the school year and ending below benchmark).[34]; or
3. Students demonstrating a final program effect (starting below benchmark, but eventually progressing and remaining above benchmark for the remainder of the school year).

The evaluation team then modeled the likelihood of falling into each of these groups as a function of program assignment and weeks of receiving MRC tutoring. After calculating the probabilities of group membership, we then simulated and compared two hypothetical students: one initially assigned to the program group and another initially assigned to the control group at the beginning of the year, but both receiving 10 weeks of sessions. In this analysis the only difference between the program and control students is the timing of receipt of MRC services (i.e., first vs. second semester). Therefore, the difference between the hypothetical program and control students' probabilities of being included in the "final effect group" represents the effect of early program intervention on the likelihood of a longer lasting effect.

For the second analysis, we focused on changes to the trajectory of student achievement prior to, during, and after receiving MRC tutoring, regardless of initial assignment to conditions (program or control group). A spline model was

---

[33] The trajectory analysis was limited to Kindergarten, second and third grade students because first grade students were assessed on two different measures during the Fall and Spring semesters.

[34] A very small number of cases (< 1%) did not fit the expected patterns. Four cases started above benchmark and either fell below or remained above the benchmark for the remainder of the school year.

employed, which allows for linear trajectory across the entire school year, but also allows for this baseline trajectory or slope to be adjusted based on two additional phases: during program and post-program. We limited the analysis to only 10 weekly tutoring sessions to remove the cross-semester (i.e., first vs. second semester) effects of low performing students who would have required a larger dosage of the MRC intervention.

Analysis of Probabilities of Group Membership

Table VI. 1 shows the findings from our analysis of longer-term program effects. Our findings show that the likelihood of being persistently below benchmark is much higher for those students initially assigned to the control group than for students initially assigned to the program group (67% and 22%, respectively). The evaluation team also estimated a higher likelihood of membership in the final program effect group among students initially assigned to the receive MRC tutoring (30%) compared to students assigned to the control group (14%) who would not have begun tutoring sessions until the second semester. These findings indicate that program group students who received tutoring early in the school year have more than twice the likelihood of both achieving and remaining above benchmark by the end of the school year compared to students assigned to the control group who received equal amounts of tutoring, but later in the school year. The probability of membership in the temporary program effect group was higher among the program group (49%) than the control group (20%); however, it is important to note that, because of sample size limitations, these estimates were calculated after only 10 weeks of tutoring. Thus, students in the temporary program effect group may eventually move above benchmark after additional weeks of tutoring. These findings support the conclusion that early intervention by the MRC program, controlling for dosage, may have a greater impact on students' longer-term (i.e., end of school year) literacy proficiency outcomes.

Table VI.1.   Estimated probabilities of effect patterns for students receiving 10 weeks of MRC tutoring by initial study group assignment (program or control group)

Group	Control		Program	
	Percent	SE	Percent	SE
No Program Effect	0.67	4.30	0.22	3.50
Temporary Program Effect	0.20	3.90	0.49	3.90
Final Program Effect	0.14	2.60	0.30	3.20

One caveat to this analytic approach is that the final impact (above or below benchmark) could only be measured in the Spring of the participating students' program year. While the number of weeks of receiving MRC tutoring included in the simulation model is the same for both program and control groups (i.e., 10 weeks), the amount of weeks of intervention data available to model the control group students' growth, on average, was fewer than for the program group. Control group students on average received treatment for only 3 or fewer weeks compared to 16 or more

weeks for the program group. This raises questions as to the robustness of the data for the control students used to create the simulation model.

Analysis Using Spline Models

In the second analysis, the evaluation team explored whether average student growth rates across the entire year could be tracked, regardless of initial assignment to treatment or control group. The goal of this model was to provide insight into the growth rates of students during one of three phases:

1. The time period prior to participation in the MRC program (i.e., baseline)
2. The time period during participation in the MRC program (i.e., program)
3. The time period after participation in the MRC program (i.e., post-program)

The evaluation team had the ability to conduct this type of analysis because we were in the unique position of having weekly progress monitoring data that had been collected on all students across the entire school year, regardless of when they actually began participation in the program. Therefore, the evaluation team was able to take advantage of this unique opportunity to fit a modified "spline" model to the data, which would provide estimates of weekly growth rates for all students across all three phases of the program, regardless of when they began MRC participation during the school year. The analysis includes students who received MRC services at any point throughout the school year, despite initial assignment to program or control groups.

Even though the data allowed this approach, there were several natural limitations to these analyses. First, the amount of time spent in the baseline period (i.e., prior to participation in the MRC program) was non-existent for half of the student sample because students who were assigned to the treatment group began receiving MRC services almost immediately at the beginning of the Fall semester. Second, since control group students were not eligible to receive MRC tutoring until the Spring semester and the program encourages schools to tutor students closest to benchmark first, the sample was by nature biased in that lower-performing students in the control group were less likely to receive MRC services and those that did were more likely to receive them later in the year. Therefore, any observed effects may be confounded with this practice. We limited the analysis to only 10 weekly tutoring sessions to address the cross-semester (i.e., first vs. second semester) effects of low performing students who would have required a larger dosage of the MRC intervention. Despite these limitations providing the results of the spline model still provides some exploratory information about the program, and serves as a model for follow-up research. Finally, this method is relatively new in multi-level models. While our preliminary simulations indicate that these methods provide reliable estimates of true underlying effects, the literature has yet to fully understand the estimation properties.

Table VI. 2 shows the findings from our analysis using spline models. The table presents values for Kindergarten, second and third grades.[35] At Kindergarten, growth rates showed significant increases over the baseline growth rate in both the program and post-program phases, implying that the MRC program may provide longer-lasting effects beyond the program period. Specifically, our models indicate that Kindergarten students had a growth rate of about 0.9 words per week prior to the MRC program. During the program this rate increased to about 1.9 words per week (0.9 + 1.0 = 1.9). This growth rate was maintained after student's exited the program, showing a rate of about 1.9 words per week as well.

No significant increases were found in second grade, and a significant but small decrease was found during the program phase in third grade. Again, these results should be interpreted with caution and are meant to illustrate a model for potential future research on the MRC program.

Table VI.2. Parameter estimates of linear growth (slope) and adjustments across baseline, program and post-program phases.

Grade	Baseline Slope	Slope Adjustments	
		Program phase	Post-program phase
Kindergarten (N=359)	0.908***	0.256***	0.446***
2nd Grade (N=409)	1.043***	0.061	0.020
3rd Grade (N=308)	1.067***	-0.110*	-0.088

***<.001
**<.01
*<.05

---

[35] As for the trajectory analysis, first grade had to be omitted from this analysis because students were assessed on two different measures during the Fall and Spring semesters.

VII. Conclusions

The findings from the MRC K-3 impact evaluation provide important evidence for addressing the studies' key research questions (presented in Chapter III). Below, the evaluation team offers our conclusions based on these findings and organizes them by the three major research questions. Following our assessment of the questions is a discussion on the implications of our findings for the MRC program.

1. What is the impact of the MRC program on student literacy outcomes?

The results of the Fall-Winter Experimental Study showed that Kindergarten, first and third grade students who received MRC tutoring achieved significantly higher literacy assessment scores by the end of the first semester than did control students who did not participate in MRC tutoring (see Chapter IV for more detail on these findings). The magnitude of MRC tutoring effects differed by grade, with the largest effects found among the youngest students (i.e., Kindergarten and first grade students), and the smallest effects among the oldest students (i.e., third grade students). Significant effects were not found for second grade students.

Kindergarten students who participated in the MRC program produced more than twice as many correct letter sounds by the end of the first semester than did students in the control condition. This large effect among Kindergarten students was not unexpected given that students participating in the program knew few if any letter sounds at the beginning of the school year and received a large amount of targeted intervention (twice the dosage of students in the later grades). In Kindergarten, many students may qualify to receive MRC services (i.e., know few letter sounds) because of a lack of exposure to or instruction in letter sound correspondence at home. Previously unexposed Kindergarten students who are explicitly taught can quickly learn the correspondence between letters and their sounds.[36] Kindergarten students in the MRC program are provided with at least 100 minutes a week of tutoring focused explicitly on mastering letter sound correspondence and related skills (e.g., phonological awareness). This intensive, targeted intervention produced large gains in letter sound knowledge, such that on average MRC students not only achieved the Winter benchmark expectation by the 12th week of intervention, but could produce over double the number of letter sounds as control group students by the end of the first semester. These findings indicate that the MRC program had more than achieved its goal of accelerating Tier 2 Kindergarten students' *letter sound fluency*, setting them on track to exceed grade level expectations by the end of the first semester.

---

[36] National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* National Institute of Child Health and Human Development, National Institutes of Health

Similarly, first grade students participating in MRC tutoring posted significantly higher *nonsense word fluency* scores than students in the control group. Although smaller than the program effect for Kindergartners, first grade students demonstrated both statistically significant and sizeable effects. By the end of the first semester, students who received MRC tutoring were able to produce more letter sound segments in nonsense words than students in the control group. Not only did MRC students exhibit greater *nonsense word fluency*, but the rate at which their fluency improved increased over the course of the first semester. This increased rate of growth enabled MRC students to reach the first grade Winter benchmark by the end of the first semester and set them on a trajectory to stay on track for grade level literacy proficiency.

In contrast to the findings for Kindergarten and first grade students, the effect of the MRC program was significant but small for third grade students and not statistically significant for second grade students. Also unlike in Kindergarten and first grade, neither second nor third grade students on average achieved the Winter benchmark level of achievement by the end of the first semester. As mentioned in the limitations section of Chapter III, the RCT ended after one semester, making it problematic to examine the differential rate of growth over the remainder of the school year. However, the Chapter V analyses of the Full-Year data suggests that third grade students' growth trajectories had indeed changed (i.e., accelerated) as a result of MRC tutoring and that their *oral reading fluency* continued to improve in subsequent weeks as a result. Therefore, an important conclusion for the evaluation is that the MRC program has an effect on third grade students' oral reading scores; however, these improvements take longer to detect among older students.

In second grade, we found no statistically significant differences between students who received MRC tutoring and control students who did not. In fact, over the course of the first semester, the oral reading fluency scores of second grade students in both the MRC program and control groups grew at similar rates and in a similar pattern. At the very end of the first semester, we began to see a slight increase in the *oral reading fluency* rate among second grade students in both the MRC program and control groups. Again, given that the RCT ended after 16 weeks, it was not possible to conclude whether more time was necessary to detect differences between the program and control groups. However, similar to the third grade students, the results of the Full-Year non-experimental analysis for second grade students indicate that more time may have been necessary to detect a difference between the second grade program and control groups in *oral reading fluency* scores.

In interpreting the Fall-Winter Experimental Study results, it is important to consider two factors that vary by grade: patterns in student development and grade-appropriate literacy skills. As students transition to higher grade levels, the reasons why students are eligible for MRC services (i.e., defined as Tier 2) are likely to change. For example, many Kindergarten students, particularly those from under-privileged homes, arrive academically unprepared for

school due to insufficient exposure to emergent literacy concepts..[37] Their lack of school readiness often results from insufficient exposure to academic language, books, and print. High-quality intensive instruction can quickly shift unexposed children to become on track for catching up with grade level expectations. However, as students progress from Kindergarten into later grades, those eligible for MRC tutoring are less likely to qualify because of lack of exposure. Instead, they are more likely to be eligible because they are struggling to acquire or integrate requisite skills. When a student struggles to learn a concept or skill, it takes significantly more in-depth intervention and time to change that student's learning trajectory so as to set them on a path to achieve grade level benchmark expectations. Whereas lack of exposure in young children can be relatively quickly remedied by intensive and explicit instruction, learning challenges in older children can take longer to overcome..[38]

Furthermore, in later grades (i.e., second and third) the literacy skills that older students are attempting to master become more complex; building upon skills learned in earlier grades (Pre-Kindergarten, Kindergarten and first grade). By second and third grade, the *oral reading fluency* measure used to assess students actually demonstrates the ability to read aloud increasingly complex connected texts with increasing fluency and requires comprehension. This task requires students to employ their previously learned skills to decode individual letters, string together letter sounds within words, and translate these strings of sounds into words, phrases and sentences. Further, the cognitive load of the oral reading fluency task is more demanding for older students because the concepts within the texts and the language become more difficult. More complicated text structures and concepts that may or may not be familiar to students, requiring more background information and prior experience with oral language in order to comprehend, can add to the difficulty of the oral reading fluency task. Thus, reading aloud is a complex skill because it requires mastery of earlier skills such as letter sound correspondence, concepts about print, phonological awareness, and vocabulary.

When we consider that oral reading is a more challenging skill to acquire and that older children who are eligible for MRC services are also more likely to have experienced challenges mastering prerequisite skills, it is not surprising that it takes longer to detect significant effects of the MRC program in second and third grade students. Perhaps the more focused, intense, and regular dialog a student experiences with an MRC tutor on an individualized basis provides increased engagement in language tasks required to understand and be more fluent over time.

[37] National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* National Institute of Child Health and Human Development, National Institutes of Health.

[38] National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* National Institute of Child Health and Human Development, National Institutes of Health.

In sum, the results of the Fall-Winter Experimental Study suggest that with respect to benchmark expectations the MRC program produces the largest effects most quickly with the youngest students, particularly Kindergarten students. The intensive one-on-one exposure to MRC tutoring produces large increases in students' *letter sound fluency*. In later grades (i.e., second and third), when students begin the more complex task of reading connected text, the MRC program takes longer to produce significant changes in student *oral reading fluency* outcomes in comparison to controls. While it was not possible to *experimentally* examine the full-year impact of the program on student outcomes in this study, the non-experimental analyses using findings from only those students who received MRC tutoring suggest that over the course of several weeks, the MRC program changes second and third grade students' growth trajectories towards increasing *oral reading fluency*.

1a. Does the impact vary by student characteristics/demographics?

After establishing the overall positive impact of the MRC program on Kindergarten, first and third grade students' literacy outcomes, we conducted analyses to examine whether differential effects of the program existed for specific subgroups of students based on the following student characteristics: gender (male/female), race (White, Black, Asian; White, non-White), Dual Language Learner (DLL) status (yes/no), and Free and Reduced Price Lunch eligibility (yes/no). A statistically significant impact of MRC tutoring was detected among Kindergarten and first grade students despite gender, minority status, DLL status, or FRPL eligibility. For each of these characteristics, students who received MRC tutoring significantly outperformed control students who did not receive tutoring on grade-specific literacy assessments.

Among third grade students, an impact was not detected within all subgroups. Third grade White, native English speaking (i.e., non-DLL), and eligible for FRPL students all produced significant differences between program and control group students. In contrast, among third grade Black and Asian students and third grade DLL students a statistically significant finding was not found between students who received MRC tutoring and control group students who did not participate in the program. Thus, the MRC program does not appear to have been effective among third grade minority students and third grade DLL students. However, it is important to recall that a large majority of Asian students (85%) were also Dual Language Learners. Therefore, it is not possible to determine whether Asian race or DLL status was independently responsible for the lack of statistically significant program impacts among these students. Furthermore, since the overall effect of the program was much smaller at third grade (and thus closer to the cutoff for statistical significance), it is not surprising that there would be greater variation in meeting this cutoff across several subgroups in third grade.

To interpret this finding, again it is important to consider qualitative differences among students who are eligible for MRC services at third grade as opposed to earlier grades. Third grade students classified as ELLs have still not yet mastered the English language. Some of these students may be new immigrants who have strong home language

skills, but lack English proficiency. Minnesota has a large immigrant Asian (majority Hmong) population, which is likely well represented in the 85% of Asian students who are ELLs in the third grade study sample. Such students could require significant remediation to produce improvements with a cognitively complex task such as English *oral reading fluency*. Also, since several schools provide specialized core instruction to third grade DLL students, it is possible that the MRC program did not have a demonstrable effect above and beyond the already intensive, DLL-specific curriculum the schools may already be providing to these students. While the MRC program may benefit DLL students in their reading fluency, the interventions used by MRC were not designed specifically to improve DLL students' English proficiency.

An important conclusion for the evaluation is that the program does appear to have an effect on students regardless of individual characteristics, with the only exception to this finding among third grade DLL students (who also happen to be predominately Asian, in the state of Minnesota). As discussed above, since the MRC program was not designed to improve English proficiency among DLL students, it is not surprising that we do not observe an impact of MRC tutoring on older DLL students' English *oral reading fluency*.

1b. Do assessment scores vary by AmeriCorps member characteristics/demographics?

Assessment scores did not vary by AmeriCorps member characteristics (i.e., gender, race, age, education, full/part time status, or years of education) nor by the specific school at which the tutoring occurred. These results support the conclusion that the MRC program is replicable in multiple school settings using AmeriCorps members with diverse backgrounds. This finding was particularly promising given the size and scope of the MRC program. The program currently recruits over 1,100 members each year to provide tutoring services in over 650 schools across the state of Minnesota. In order to maintain this level of engagement, the program recruits a diverse group of AmeriCorps members to serve as literacy tutors. Many MRC members have no previous experience working in schools, with students, or in the domain of literacy. In the *Process Assessment of the Minnesota Reading Corps,*[39] we concluded that the MRC program's high-quality training regime, research-based scripted interventions, regular objective assessment, ongoing on-site coaching, and multi-layered supervisory structure resulted in high levels of fidelity of program implementation and positive impacts on student literacy outcomes. The results of the Fall-Winter Experimental Study showed quantitatively that these critical program supports indeed reduced variability in the interventions delivered by AmeriCorps members within diverse school settings, such that the impact of member characteristics and individual school effects on first, second, and third grade students was near zero. Further, the larger but statistically insignificant school-level effects at Kindergarten suggest a possible impact of the K-Focus program, which is worth exploring in future research. The lack of member-level and school-level effects on student

---

[39] Hafford, C., Markovitz, C., Hernandez, M, et al. (February 2013). *Process Assessment of the Minnesota Reading Corps Program.* (Prepared under contract to the Corporation for National and Community Service). Chicago, IL: NORC at the University of Chicago.

outcomes validates MRC's approach to training, coaching and supervision, as well as their intentional recruitment of members with diverse backgrounds who do not necessarily have formal training or experience in education or literacy instruction.

2. **Does the impact of the program vary week to week? Does the number of weeks of intervention (i.e., dosage) impact student literacy outcomes?**

While the Fall-Winter Experimental Study examined the overall impact of the program during the first semester, the full-year analysis examined estimates of week over week impacts to identify patterns in student growth by grade. Although the Full-year non-experimental analysis is designed to provide a more complete understanding of the MRC program's impact over time, the study findings should be interpreted with caution, and any conclusions drawn should be carefully considered since the findings are based on a non-experimental design (i.e., control group students became eligible for MRC tutoring services after the Winter benchmarking).

Our conclusion from the examination of week over week patterns of impacts is that the MRC program appears to be effective with K-3 students; however, the patterns in week over week growth vary by grade. Kindergarten students showed immediate and large gains, the largest of which occurred in the first few weeks of tutoring. First, second, and third grade students showed small, but steady week to week gains throughout the entire period of analysis. The patterns of week over week effects for Kindergarten students in the MRC program are consistent with the findings in the experimental impact analysis, which showed large gains in *letter sound fluency* during the first semester. Given our discussion above, the growth trajectories provided in the week over week analysis are consistent with expectations for younger students receiving effective literacy interventions.

An important consideration when interpreting these findings is that roughly half of the schools in our study sample had K-Focus members, who spent a majority of their time providing Kindergarten students with two daily sessions of the MRC interventions. While the typical MRC K-3 program provides students with one 20-minute session per day, in the K-Focus program, each Kindergarten student participates in two (20-minute) sessions daily, for a total of 40 minutes of literacy-focused instruction. Although we are unable to disentangle the added effect of the K-Focus program on Kindergarten students' outcomes, it is reasonable to consider that the more intensive, higher dosage intervention for Kindergarten students in some schools may have contributed to producing the large effects we observed in the findings.

In contrast to the findings for Kindergarten students, the findings for first, second and third grade students show small, incremental gains week over week. This pattern of gains among students in these grades continued to build throughout the study period. These findings are not unexpected, given the earlier discussion that struggling students exhibit different patterns of literacy development depending on age (i.e., older students are likely to need more

intensive literacy intervention and practice, while younger students can benefit from increased exposure to literacy activities and time on task).

The findings in both second and third grade students indicate that the small effects found in the Fall-Winter Experimental Study may have been more substantial if it had been possible to follow students for a longer period of time beyond 16 weeks. Thus, if the timeline for the experimental study could have been lengthened to allow observation of differences in scores between the program and control group of students over an entire school year, our finding for second and third grade students may have been more substantial. The full-year findings support our earlier conclusion that the MRC program has an effect on second and third grade students' *oral reading fluency* scores; however, improvements in this more complex literacy task may take more time to accumulate in older students. Thus, the only possible method for confirming potentially larger effects of the program on second and third grade students would be to conduct a longer RCT study over an entire school year.

3. **Does participation in MRC have a longer-term impact on student literacy outcomes as measured at the end of the school year?**

The evaluation team found evidence that participation in the MRC program results in stronger end-of-year effects on literacy outcomes when interventions begin earlier in the school year. When MRC interventions are implemented later in the school year (i.e., second semester), the probability of progressing and staying above benchmark through the end of the school year decreases, while the likelihood of remaining chronically behind increases substantially. The findings showed that program group students who received tutoring assistance early in the school year have more than twice the likelihood of remaining above benchmark for the remainder of the school year compared to students assigned to the control group who received equal amounts of tutoring assistance, but later in the school year (29.5% compared to 11.8%). The higher likelihood for program group students indicates that early intervention by the MRC program, controlling for dosage, may have a greater impact on students' longer-term literacy proficiency outcomes. Thus, a key conclusion from our analysis is that early intervention from the MRC program (i.e., in the first semester) for struggling students results in a higher likelihood of positive outcomes by the end of the school year.

Program Implications from Conclusions

Overall, the results of the evaluation showed that the MRC program positively impacts Kindergarten through third grade students' literacy outcomes. The magnitude of the effect of the MRC program varies by grade, such that younger students achieve larger gains in a shorter period of time, whereas older students post smaller gains that accumulate over a longer period of time.

The large program effects found among Kindergarten students indicate that MRC tutoring can rapidly and effectively shift eligible Kindergarten students' learning trajectories so they become on track to meet or exceed grade level expectations within the first semester of school. The large and significant impact of the MRC program's early, intensive intervention suggests that MRC tutoring helps students who are not Kindergarten-ready catch up to their peers on the critical emergent literacy skill of *letter sound fluency*. Smaller but still moderate sized effects of participating in MRC tutoring were also found among first grade students who were extending letter sound knowledge to the more complicated task of sounding out nonsense words. MRC tutoring produced effects over a more extended period of time among second and third grade students who were assessed on *oral reading fluency*. While the evaluation results suggest that a lengthier period of time for the experimental study (i.e., a full school year) may have helped us detect more substantial improvement on this more complex literacy skill, the non-experimental results show that the cumulative impact of over a semester's worth of MRC tutoring can result in increases in students' proficiency in this skill.

The *Process Assessment of the MRC Program* (2013) provides important qualitative insights into how the program produced these significant impacts on student literacy outcomes. First, the MRC program uses objective data based decisions to determine student eligibility for services, to track students' learning progress, and to make adjustments to interventions when needed. Second, the program uses a small but effective set of highly scripted research based interventions to tutor students. These interventions are research validated as effective, and are scripted so that members can learn the interventions well and implement them with fidelity. Third, MRC provides comprehensive, ongoing training to and supervision of AmeriCorps members. The program utilizes a multi-layered supervisory structure, which supports members' on-site implementation of the assessment and tutoring interventions. Both centralized (i.e., Master Coach and Program Coordinator) and on-site (i.e., Internal Coach) support are provided for assuring the proper identification of students, implementation of interventions, and use of data-driven decision-making.

One of the most critical findings for program replication is MRC's successful deployment of AmeriCorps members lacking any specialized background in education or literacy. The results of the member analysis revealed no significant differences in student impacts due to the characteristics of the members providing the tutoring. The lack of member effects suggests that the combination of MRC training, supervisory/coaching supports, use of objective data, and scripted, proven interventions can allow an extremely diverse range of volunteers without education backgrounds to produce significant impacts on student literacy.

As supported by the findings and conclusions from the *Process Assessment*, the MRC program appears to be highly replicable. If similar program-based infrastructure and resources are provided and specialized interventions are accurately implemented and closely monitored, members with diverse backgrounds can serve without possessing

any specialized prerequisite technical skill. The combination of MRC program elements that resulted in positive impacts on student literacy outcomes can be considered an effective model for the development of other successful reading intervention programs for K-3 students.

Recommendations for Future Research

In our analysis, we identified some topics that may be useful to explore in future studies. A major conclusion from the Full Year Non-Experimental Study was that a longer embargo period for the Fall-Winter Experimental Study may have allowed us to detect even greater impacts of the MRC program on second and third grade students' oral reading fluency. The Full-Year non-experimental analysis suggested that more than 16 weeks is needed to detect significant impacts of the MRC program on oral reading fluency. Therefore, if feasible, a future study of the MRC program should examine a lengthier period of time beyond 16 weeks to allow more time for impacts to develop in second and third grade students. Alternatively, the smaller impacts found among older students may be more closely examined by studying the impact of specific MRC interventions on *oral reading fluency*, as well as the number and timing of changes in the use of these interventions, to more fully explore which interventions may be more effective with older students.

In addition, the subgroup analysis based on proximity to Fall benchmark (baseline) showed that students farthest away from benchmark on average experienced the largest impacts of the program. Additionally, the subgroup analysis showed that control group students closest to Fall benchmark on average reached Winter benchmark by the end of the first semester without MRC intervention. However, it is important that any potential reasons for differences in average outcomes by baseline benchmark score be fully explored in future research before a recommendation can be made as to how MRC identifies and prioritizes students to receive interventions.

Finally, in order to assess the long-term impact of the MRC program on student literacy outcomes, we recommend following the same group of randomized students into the future. A longitudinal study of students' literacy and academic outcomes on state proficiency tests, graduations rates, and other more distal educational and economic outcomes, would provide important information about the impact of the MRC program on students learning and life course trajectories.

Impact Evaluation of the Minnesota Reading Corps K-3 Program

March 2014

The Corporation for National and Community Service

The mission of the Corporation for National and Community Service (CNCS) is to improve lives, strengthen communities, and foster civic engagement through service and volunteering. CNCS, a federal agency, engages more than five million Americans in service through AmeriCorps, Senior Corps, the Social Innovation Fund, the Volunteer Generation Fund, and other programs, and leads the President's national call to service initiative, United We Serve. For more information, visit NationalService.gov.

Minnesota Reading Corps

Minnesota Reading Corps is a statewide literacy initiative of ServeMinnesota that blends the people power of AmeriCorps members with the science of how children learn to read. Trained AmeriCorps members provide individualized tutoring and proven interventions to those children who are at risk for not reading at grade level. Since 2003, Minnesota Reading Corps has helped more than 100,000 struggling readers, age 3 to grade 3, progress toward reading proficiency and the program has expanded into seven additional states and Washington, D.C. For more information, please visit MinnesotaReadingCorps.org.

NORC at the University of Chicago

NORC at the University of Chicago is an independent research organization headquartered in downtown Chicago with additional offices in the University of Chicago campus, the Washington, D.C. metro area, Atlanta, Boston, and San Francisco. With clients throughout the world, NORC collaborates with government agencies, foundations, education institutions, nonprofit organizations, and businesses to provide data and analysis that support informed decision making in key areas including health, education, crime, justice, and energy.

Corporation for
NATIONAL & COMMUNITY SERVICE ★★★▬