



Evaluation Brief

July 2009

Evaluation Unit

Impact of *READ 180* on At-Risk Middle School Students' Literacy Outcomes: 2007–2008

Margaret H. Gheen and Shahpar Modarresi, Ph.D.

Executive Summary

This evaluation examines literacy achievement of middle school students at risk of low reading achievement who were enrolled and not enrolled in the *READ 180* program in academic year 2007–2008. *READ 180*, published by Scholastic Inc., is a literacy program designed to accelerate the reading achievement of low-performing students. Scholastic Inc. indicates the importance of maintaining high levels of program implementation for achieving optimal results.

End-of-year literacy scores on the Measures of Academic Progress-Reading (MAP-R) and the Maryland School Assessment (MSA) in reading were compared for the following three groups: a) students enrolled in *READ 180* classes with the intended higher levels of implementation, b) students enrolled in *READ 180* classes with lower levels of implementation, and c) students who were not enrolled in *READ 180* classes.

At-risk students in Grades 6, 7, and 8 who had scored basic on the 2007 MSA in reading made significant one-year gains on the MAP-R from spring 2007 to spring 2008, across *READ 180* and non-*READ 180* groups. Overall, average gains ranged from 4.7 to 9.0 scale score points, indicating substantial improvement. Nevertheless, many students with low reading achievement in spring 2007 continued to lag below the levels of performance desired for their grade level by spring 2008, regardless of *READ 180* participation. Therefore, despite substantial gains made in the 2007–2008 academic year, many of these students likely would benefit from continued literacy supports in subsequent years.

In analyzing which groups of *READ 180* and non-*READ 180* students had the highest literacy outcomes in spring 2008, findings across grade levels, literacy measures, and statistical methods yielded small and sometimes inconsistent patterns of differences among groups. However, overall, students enrolled in *READ 180* had slightly higher

end-of-year reading scores than nonparticipants. The biggest differences in reading outcomes among groups were found for Grade 6 students. For this grade level, 90 minutes of daily *READ 180* instruction and implementation of at least four of the program's five core instructional formats four or five days a week were found to have the greatest impact on reading scores.

Recommendations are as follows:

- Continue to implement *READ 180* to address the needs of secondary students at risk of low literacy performance.
- Because findings suggest that *READ 180* may be especially effective for at-risk students in Grade 6, consider including these students among the population prioritized for program enrollment. Montgomery County Public Schools (MCPS) leadership initially indicated students in Grades 7 through 10 were to be given enrollment priority.
- Expand the collection of districtwide data on implementation of *READ 180* and other literacy programs intended to accelerate student achievement. As part of this effort, identify ways to collect information on literacy achievement for students in Grades 9 through 12 so that analyses of targeted reading programs can include high school students.
- Given the patterns of positive outcomes for *READ 180* students in Grades 6 through 8, establish and monitor opportunities for implementing the critical features of *READ 180* (e.g., 90 minutes daily and instructional formats) with high fidelity and high quality.

Background

Goal 1 of the MCPS strategic plan, *Our Call to Action: Pursuit of Excellence* (MCPS, 2008), articulates the district's commitment to enabling all students to achieve or exceed state and local proficiency standards. As part of this commitment, in 2007–2008 the *READ 180* program, published by

Scholastic Inc., was expanded to 37 MCPS middle schools,¹ 25 high schools, and 2 special schools to accelerate the reading performance of students at risk of scoring below proficiency (basic) on the MSA in reading. The Office of Curriculum and Instructional Programs requested that OSA conduct an evaluation of the *READ 180* program.

Scholastic Inc. indicates the importance of maintaining high levels of implementation for achieving optimal results, stating—

The research is clear: We must invest sufficient time for instruction for students who are at risk of failure. Studies have conclusively shown that when schools implement and follow the 90-Minute Instructional Model, significant gains can be expected after one or two years of program participation.²

MCPS central office staff provided schools with guidelines for *READ 180* program implementation. A districtwide review of classes implementing the program found that students with a history of low performance on reading assessments were appropriately selected for enrollment; small class size of 15 students was maintained³; teachers reported they had access to key *READ 180* materials; and *READ 180* professional development had high teacher attendance rates. *READ 180* teachers gave high marks to the program's instructional materials and professional development as useful for effective literacy instruction.

MCPS leadership recommended that *READ 180* classes were to be scheduled for 90 minutes of daily instruction and that the program's core instructional formats (i.e., whole group instruction, three rotations—small group instruction with teacher, independent reading, and independent *READ 180* computer software—and whole-group wrap up) should be implemented with fidelity.

This brief presents findings from a study of student outcomes associated with levels of *READ 180* implementation. The specific evaluation questions are as follows:

- To what extent did literacy achievement improve for at-risk students enrolled and not enrolled in *READ 180* for 90 minutes daily or fewer?
- After controlling for preexisting differences among students, are there differences in reading performance between at-risk students receiving the *READ 180* program for 90 minutes daily, those receiving it fewer than 90 minutes daily, and those not enrolled?
- Among *READ 180* classes, is implementation of the program's recommended instructional formats associated with higher levels of reading achievement, after controlling for preexisting differences among students?

Methodology

Findings are more reliable if multiple analyses lead to similar results. Therefore, this evaluation seeks to triangulate outcomes from multiple measures of student literacy, analytical procedures, and features of program implementation. A short summary of the methodology is provided. For an extended technical description, refer to Appendices A and B.

Literacy Measures

Multiple measures were utilized to examine students' literacy achievement: 1) scale scores from spring 2007 and 2008 computer adaptive MAP-R, developed by the Northwest Evaluation Association (NWEA) and 2) scale scores from the 2007 and 2008 MSA in reading, published by the Maryland State Department of Education (MSDE). MAP-R scores are vertically equated, which means that an individual student's MAP-R score in 2007 can be directly compared with his/her score in 2008 to measure growth over time. However, MSA scores are not appropriate for measuring a student's gains or growth in performance from one grade to the next, because the assessments are not vertically equated. This means a student's MSA score in one year cannot be subtracted from a score earned in a subsequent year to measure change.

Program Implementation Measures

Level of *READ 180* implementation was defined in two ways. The first definition was based on students' scheduled *READ 180* class time as follows: a) students in classes meeting 90 minutes daily, b) students in classes meeting fewer than 90 minutes daily, and c) students not enrolled in *READ 180*.

If a *READ 180* class did not implement the program's core instructional formats regularly, students in that

¹ Herbert Hoover was the only MCPS middle school that did not offer *READ 180* in 2007–2008.

² See link "Why 90 Minutes? Learn More" at <http://teacher.scholastic.com/products/read180/overview/instrmodel.htm>.

³ Maximum class size was 15 students if one teacher was assigned, and 20 students if an additional co-teacher or paraeducator was assigned.

class did not experience the full program as designed. To investigate the impact of implementing the program's instructional formats, the second definition of implementation differentiated among *READ 180* classes based on teachers' reports of how frequently the program's core instructional formats were implemented. A *READ 180* class was identified as "near-full instructional model" if the teacher reported that a student in the class typically experienced at least whole group and all three rotations—with or without wrap up—four times a week. Average achievement for the study sample of students in these *READ 180* classes was compared with average achievement for the study sample in *READ 180* classes wherein these formats were implemented less frequently.

Study Sample

To facilitate making accurate comparisons across groups of students, the samples of students in each group were selected based on similar preenrollment academic characteristics. Students in Grades 6 through 8 who scored basic on the spring 2007 MSA in reading were included. See Appendix A Table A1 for a description of the sampling criteria used in this study and a summary of the sample's demographic profile.

Data Analyses

Descriptive analyses of average change in MAP-R scores from spring 2007 to 2008 and of spring 2008 MSA reading scores were conducted for students enrolled and not enrolled in *READ 180*. Findings are presented for student subgroups including gender; race/ethnicity; and receipt of special services such as English for Speakers of Other Languages (ESOL), Free and Reduced-price Meals System (FARMS), and special education.

In addition, the achievement of students experiencing varying levels of *READ 180* was compared using advanced statistical procedures which controlled for differences in students' prior achievement, demographics, and receipt of special services. First, analyses compared changes in spring 2007 and spring 2008 MAP-R scores for three groups experiencing different amounts of time in *READ 180*. Second, analyses compared spring 2008 MSA reading outcomes of these same three groups of students. Finally, analyses examined how spring 2008 MAP-R and MSA achievement compared for *READ 180* classes implementing full or near-full instructional

models with *READ 180* classes that implemented core instructional formats with lower fidelity.⁴

Effect sizes were used to judge whether the observed differences between groups were large enough to be of practical significance to educators (American Psychological Association, 2001).⁵ Effect sizes provide an estimate of the relative size of differences between groups, and interpretations of effect sizes are provided (<http://web.uccs.edu/lbecker/Psy590/es.htm#Cohen#Cohen>). An effect size up to .20 typically is considered small, .50 is medium, and .80 is large (Cohen, 1988). Effect sizes greater than .10 are presented, regardless of the level of statistical significance. An effect size of .20 or greater typically is considered an indication that the magnitude of the difference has practical significance.

Replication Analyses

The above analyses were repeated by the second author with a different statistical procedure. Specifically, performance of students who were enrolled in the *READ 180* program for 90 minutes daily was compared with performance of students enrolled for fewer than 90 minutes and also with performance of similar students who were not enrolled in *READ 180*.

Results

Results are presented by research question.

Question one: To what extent did literacy achievement improve for at-risk students enrolled and not enrolled in READ 180 for 90 minutes daily or fewer?

Descriptive analyses of gains in MAP-R scores indicate that, overall, literacy performance of at-risk students across grades, student subgroups, and *READ 180* groups improved substantially from spring 2007 to 2008 (see Table 1). Table A2 in Appendix A provides additional descriptive detail. These descriptive data are useful for acquiring an overall impression of student progress.

⁴ Approximately one third of *READ 180* classes in middle school enrolled students in more than one grade level. Of these mixed-grade classes, approximately one third enrolled students in Grade 6, and almost all enrolled students in Grades 7 or 8. Analyses did not differentiate between single-grade or mixed-grade classrooms.

⁵ Effect sizes were calculated as $(M_t - M_c)/SD$. The M_t and M_c are adjusted group means for the groups of students being compared, and SD is the standard deviation of the pooled spring 2008 scores.

Table 1
Mean Spring 2007 and 2008 MAP-R Scores and Gains for At-Risk Students by Time in *READ 180* and Student Subgroups

	<i>READ 180</i> 90 minutes daily			<i>READ 180</i> fewer than 90 minutes daily			No <i>READ 180</i>		
	2007	2008	Gain	2007	2008	Gain	2007	2008	Gain
Grade 6									
All	195.3	201.8	6.5	197.0	200.3	3.4	196.5	201.2	4.7
Male	194.4	200.9	6.4	196.0	199.0	2.9	195.5	200.0	4.5
Female	196.4	202.9	6.5	198.7	202.8	4.1	197.8	202.7	4.9
AfAm	194.9	202.1	7.2	198.4	199.6	1.2	196.4	200.2	3.9
AsAm	189.5	200.6	11.1	195.0	203.4	8.4	199.1	205.7	6.6
Hisp	196.2	202.3	6.1	196.3	198.9	2.6	194.8	200.1	5.3
White	193.9	194.9	1.0	194.6	205.7	11.1	199.3	203.6	4.3
ESOL	194.5	202.8	8.3	190.7	193.5	2.8	189.2	198.6	9.4
FARMS	195.5	201.7	6.2	197.0	198.9	1.9	195.2	200.1	4.9
SpEd	191.8	197.5	5.8	191.6	196.1	4.4	189.9	194.5	4.6
Grade 7									
All	194.9	203.9	9.0	196.9	204.3	7.4	196.0	202.4	6.4
Male	193.3	202.8	9.5	195.5	202.3	6.8	194.0	200.4	6.4
Female	197.3	205.5	8.2	198.5	206.8	8.3	198.6	205.1	6.5
AfAm	193.6	202.5	8.9	198.0	207.2	9.2	195.7	202.2	6.4
AsAm	192.5	203.1	10.6	200.3	204.5	4.2	231.0	236.0	5.0
Hisp	196.1	204.6	8.5	197.5	204.1	6.7	200.0	205.8	5.7
White	196.9	207.5	10.6	194.1	201.3	7.2	194.6	201.4	6.7
ESOL	188.2	197.3	9.2	189.9	202.7	12.8	198.1	204.0	5.9
FARMS	194.3	203.1	8.8	196.0	203.9	7.9	186.8	197.9	11.1
SpEd	192.3	201.4	9.1	191.9	199.6	7.7	195.0	201.7	6.7
Grade 8									
All	199.8	208.3	8.5	199.2	207.4	8.2	203.5	209.3	5.8
Male	198.4	207.4	9.0	198.0	207.7	9.7	202.2	208.1	6.0
Female	201.4	209.3	7.8	200.7	206.9	6.3	205.2	210.7	5.5
AfAm	198.7	207.2	8.5	198.5	206.6	8.2	203.7	208.5	4.8
AsAm	201.3	211.9	10.7	199.6	211.5	11.9	207.9	214.8	7.0
Hisp	199.4	207.8	8.5	198.9	207.1	8.2	201.0	207.2	6.2
White	203.1	210.3	7.2	202.0	207.0	5.0	206.5	212.8	6.3
ESOL	193.4	206.8	13.3	192.4	204.9	12.5	191.4	203.9	12.5
FARMS	199.2	206.0	6.8	198.4	207.1	8.7	201.7	207.5	5.9
SpEd	196.8	205.0	8.2	195.4	202.4	7.0	196.0	202.2	6.2

Note. Sample includes students who scored basic on 2007 MSA reading and had both 2007 and 2008 MAP-R scores. Due to small numbers, data for American Indian students are not presented, although they are included in totals. AfAm=African American, AsAm=Asian American, Hisp= Hispanic, SpEd=special education, ESOL=English for Speakers of Other Languages, FARMS=Free and Reduced-price Meals System.

In general, the largest average gains were observed for students in *READ 180* classes that met for 90 minutes daily. Furthermore, the average one-year gains for students in *READ 180* classes that met for 90 minutes daily were almost two to three times larger than the average one-year gains reported for a national sample of students in Grades 6, 7, and 8 (Table 2).⁶ For some *READ 180* and non-*READ 180*

⁶ 2008 normative data available from NWEA at http://www.nwea.org/assets/downloads/980/Normative%20Data%20Sheet_v2.pdf.

student subgroups, the gains were even larger (see Table 1).

Although in each grade level *READ 180* and non-*READ 180* groups and student subgroups made substantial gains in MAP-R scores over the course of the academic year, average spring 2008 scores remained below the national grade-level score norms reported by NWEA (Table 2). Table A3 in Appendix A presents NWEA national spring MAP-R score norms for students in kindergarten through Grade 9, as well as increases in these mean scores from one grade level to the next. Note that in the national sample, average MAP-R scores for students in earlier grades tend to show large increases from year to year, and yearly gains tend to be much smaller for students in later grades.

Table 2
Average Spring MAP-R Scores and One-Year Gains for National (Norm) Sample and Students in *READ 180* for 90 Minutes Daily by Grade Level

	MAP-R mean spring scores for national (norm) sample			MAP-R mean spring scores for students in <i>READ 180</i> for 90 minutes daily		
	Prior grade	Sub- sequent grade	Gain	Prior grade (2007)	Sub- sequent grade (2008)	Gain
Grade 5 to 6	211.1	214.8	3.7	195.3	201.8	6.5
Grade 6 to 7	214.8	217.9	3.1	194.9	203.9	9.0
Grade 7 to 8	217.9	221.2	3.3	199.8	208.3	8.5

Note. MAP-R mean spring scores for the national sample of students are mean “end of year” scores reported by NWEA. “Gain” is the difference between spring mean scores from one grade level to the next. *READ 180* student sample included those who scored basic on the 2007 MSA in reading.

Analysis of student performance on spring 2008 MSA reading confirmed that many students who had scored basic in 2007 scored basic again in 2008. See Tables A4 and A5 in Appendix A for descriptive statistics. Between 46.9% (Grade 7, *READ 180* for fewer than 90 minutes daily) to 65.8% (Grade 8, *READ 180* for 90 minutes daily) of students who had scored basic on the 2007 MSA in reading scored basic again in 2008. These data indicate that although students made substantial progress, many likely would require continuing support in subsequent years to reach grade-level norms (as measured by MAP-R) and proficiency (as measured by MSAs) in reading.

Question two: After controlling for preexisting differences among students, are there differences in reading performance among at-risk students receiving the READ 180 program for 90 minutes daily, those receiving it fewer than 90 minutes daily, and those not enrolled?

To answer this question, data were analyzed in the following two ways: 1) change in MAP-R scores from spring 2007 to spring 2008 and 2) spring 2008 MSA reading scores. These analyses are different from the descriptive analyses presented for question one because statistical controls were used to adjust for preexisting differences between students. Findings are summarized in Table 3. A more extended description of results follows.

Table 3
Summary of Results for Question Two: Comparing Effects of Time in *READ 180 (R180)* by Grade Level

Grade	Change in MAP-R from spring 2007 to 2008	Spring 2008 MSA reading
6	Groups began with similar scores in 2007. By 2008, <i>R180</i> for 90 minutes daily had higher scores than <i>R180</i> for fewer than 90 minutes daily and non- <i>R180</i> .	<i>R180</i> for 90 minutes daily and non- <i>R180</i> had higher scores than <i>R180</i> for fewer than 90 minutes daily.
7	Groups began with similar scores in 2007. By 2008, <i>R180</i> for 90 minutes daily and <i>R180</i> for fewer than 90 minutes daily had higher scores than non- <i>R180</i> .	<i>R180</i> for fewer than 90 minutes daily had higher scores than non- <i>R180</i> .
8	In 2007, <i>R180</i> for 90 minutes daily and <i>R180</i> for fewer than 90 minutes daily had lower scores than non- <i>R180</i> , but caught up by 2008.	Non- <i>R180</i> had higher scores than <i>R180</i> for fewer than 90 minutes daily.

Effect of Time in READ 180 on Change in MAP-R Scores from Spring 2007 to 2008

In Grade 6, after making statistical adjustments, all three groups began with similar MAP-R scores in spring 2007. However, by spring of 2008, students in *READ 180* classes that met for 90 minutes daily attained higher scores compared with students in *READ 180* classes that met less frequently (effect size=.24). This difference is large enough to be of practical significance.⁷ Students in *READ 180* classes that met for 90 minutes daily had slightly higher scores compared with students not enrolled (effect size=.16). This difference is very small.

In Grade 7, after making statistical adjustments, all three groups began with similar MAP-R scores in

⁷ For an effect size of .24, a student who scored higher than 50% of students taking *READ 180* for 90 minutes daily scored higher than approximately 60% of students taking *READ 180* less frequently.

spring 2007. However, by spring 2008, students enrolled in *READ 180* classes that met for 90 minutes daily (effect size=.11) or fewer than 90 minutes daily (effect size=.13) attained slightly higher MAP-R scores compared with students not enrolled in *READ 180*. These differences between groups are very small.

After making statistical adjustments, Grade 8 students not enrolled in *READ 180* had slightly higher MAP-R scores in spring 2007 compared with students in classes that met for 90 minutes daily (effect size=.10) and compared with *READ 180* students who met less frequently (effect size=.15). By 2008, no differences between groups were observed, indicating that the MAP-R achievement of the two *READ 180* groups attained scores on par with those not in the program. Full statistical details for each grade level analysis are in Table A6 in Appendix A.

Effect of Time in READ 180 on Spring 2008 MSA Reading Scores

Analyses of spring 2008 MSA reading scores found small differences between *READ 180* and comparison groups. The largest differences were found for Grade 6 students. Full statistical details are in Table A7 in Appendix A.

In Grade 6, after making statistical adjustments, spring 2008 MSA reading scores for students in *READ 180* classes that met for 90 minutes daily (effect size=.19) and for nonparticipants (effect size=.18) were higher than for students enrolled in *READ 180* classes with less class time.⁸ These differences between groups are small but have practical significance.

After making statistical adjustments, Grade 7 students taking *READ 180* for fewer than 90 minutes daily had slightly higher scores than did students not taking *READ 180* (effect size=.13).⁹

The reverse pattern was observed for Grade 8. After making statistical adjustments, students taking

⁸ For effect sizes of .18 to .19, a student who scored higher than 50% of students in *READ 180* for 90 minutes daily and students not enrolled scored higher than approximately 57% of students in *READ 180* for fewer than 90 minutes daily.

⁹ The replication analysis presented in Appendix B found a small effect size in Grade 7 in favor of *READ 180* for 90 minutes daily compared with no *READ 180* (effect size=0.13), although this difference was only marginally statistically significant ($p=.07$).

READ 180 for fewer than 90 minutes daily had slightly lower MSA reading scores compared with students not enrolled in the program (effect size=-.15). These differences between groups' achievement outcomes were very small. Scores for Grade 8 students enrolled in *READ 180* for 90 minutes daily were not different from the other two groups.

Replication of Results

Overall, the replication analyses found similar results as those presented above. After adjusting for preexisting differences among students, patterns of slightly higher performance were found for Grade 6 students as a result of high exposure to the *READ 180* program (90 minutes daily compared with fewer than 90 minutes daily), as measured by both spring 2008 MAP-R (effect size=0.23) and MSA reading (effect size=0.18). No differences between 90 minutes of daily *READ 180* instruction compared with fewer than 90 minutes daily were found for students in Grades 7 and 8.

Similar analyses of spring 2008 MAP-R indicated small differences for Grade 7 students (effect size=0.16) and very small differences for students in Grades 6 (effect size=0.12) and Grade 8 (effect size=0.11) in favor of students receiving 90 minutes of *READ 180* instruction compared with non-*READ 180* students. For Grade 7 students, slightly higher MSA performance (effect size=0.13) was found for those receiving *READ 180* for 90 minutes compared with performance of students not enrolled in the program. Detailed description of methodology and findings are in Appendix B.

Question three: Among READ 180 classes, is implementation of the program's recommended instructional formats associated with higher levels of reading achievement, after controlling for preexisting differences among students?

To answer this question, analyses focused exclusively on students in *READ 180* classes, and adjustments were made to account for preexisting differences among students. Analyses estimated the degree to which reading achievement outcomes were different in classes that implemented a full or near-full *READ 180* instructional model as compared with achievement outcomes in classes that had lower levels of *READ 180* instructional implementation. Table A8 in Appendix A presents full statistical details. A brief description of results for each grade level follows.

Grade 6. The instructional model implemented in *READ 180* classes had the greatest effects on the spring 2008 reading performance of students in Grade 6, especially for MAP-R scores. Students in classes that implemented at least the near-full instructional model (whole group instruction and all three small-group rotations—with or without wrap up—at least four days a week) attained MAP-R scores that were on average 4.2 points higher (effect size=.37) than did classes implementing the instructional model with lower fidelity.¹⁰ This effect size was the largest found in this study and is of practical significance.

The impact of implementing at least the near-full *READ 180* instructional model on MSA scores was not as large as for MAP-R scores but was significant. Students in *READ 180* classes that implemented whole group instruction and all three small-group rotations, with or without wrap up, at least four days a week attained MSA reading scores that were 3.9 points higher (effect size=.19) than scores in classes implementing this instructional model less frequently.¹¹ A difference of this magnitude has practical significance.

Grades 7 and 8. For students in Grades 7 and 8, the impact of implementing a near-full instructional model at least four days a week yielded average scores that were no different than scores in classes with lower levels of implementation.

Conclusions

Academic year 2007–2008 marked the first year that *READ 180* was available to almost all middle and high schools in MCPS. Gains in reading achievement scores and increases in the proportion of students meeting grade-level performance may grow as teachers gain experience with the program's assessment tools, materials, and instructional formats.

At-risk students in Grades 6, 7, and 8 who had scored basic on the 2007 MSA in reading made significant one-year gains on the MAP-R from spring 2007 to spring 2008, across *READ 180* and non-*READ 180*

¹⁰ For an effect size of .37, a student who scored higher than 50% of students in a class implementing at least whole group and three rotations four days a week scored higher than approximately 64% of students in classes implementing these formats less frequently.

¹¹ For an effect size of .19, a student who scored higher than 50% of students in a class implementing whole group and three rotations at least four days a week scored higher than 57% of students in classes implementing these formats less frequently.

groups. Overall, average gains ranged from 4.7 to 9.0 scale score points, indicating substantial improvement. Nevertheless, many students with low reading achievement in spring 2007 continued to lag below the levels of performance desired for their grade level by spring 2008, regardless of *READ 180* participation. Therefore, although substantial achievement gains were made in the 2007–2008 academic year, many of these students likely would benefit from continued literacy supports in subsequent years.

In analyzing which *READ 180* and non-*READ 180* groups had the highest literacy outcomes in spring 2008, findings across grade levels, literacy measures, and statistical methods yielded small and sometimes inconsistent patterns of differences. However, overall, students enrolled in *READ 180* appear to have slightly higher end-of-year reading scores than nonparticipants. The biggest differences in reading outcomes among groups were found for Grade 6 students. For this grade level, 90 minutes of daily *READ 180* instruction and implementation four or five days a week of at least four of the program’s five core instructional formats (i.e., whole group instruction, three rotations—small group instruction with teacher, independent reading, and independent *READ 180* computer software—with or without whole-group wrap up) were found to have the greatest impact on reading scores.

Strengths and Limitations of the Study

Although the findings obtained from this study were based on sound evaluation design, as well as appropriate analyses, it should be noted that causality cannot be directly inferred from the current study.

Information on the nature of targeted literacy programs experienced by students not enrolled in *READ 180* is limited, as this was not the focus of this evaluation. It is possible that the non-*READ 180* groups received literacy supports that were of similar intensity and quality as those provided in *READ 180* classes, resulting in small differences in outcomes between groups of students enrolled in *READ 180* and the group of students not enrolled.

Analyses did not account for student enrollment in *READ 180* before 2007–2008 because these data were not available. Some students in Grades 7 and 8 may have taken the course in prior middle school grades and may have benefited from earlier program exposure. This confounding variable in program participation may partly explain why differences in achievement between *READ 180* and non-*READ 180*

groups were smaller in Grades 7 and 8 than in Grade 6. Similarly, the larger differences observed between groups in Grade 6 may be partly attributed to the fact that *READ 180* and non-*READ 180* students in this grade level had no prior exposure to the program.

Information on the quality of instruction in non-*READ 180* and *READ 180* classes, independent of instructional time and implementation of the program’s core instructional formats, was not collected. The impact of any instructional endeavor likely depends on a teacher’s ability to effectively implement it.

To pursue these issues, a more comprehensive evaluation of literacy supports and programs in MCPS and careful analysis of quality of the implementation of program features would be required.

Recommendations

Recommendations are as follows:

- Continue to implement *READ 180* to address the needs of secondary students at risk of low literacy performance.
- Because findings suggest that *READ 180* may be especially effective for at-risk students in Grade 6, consider including these students among the population prioritized for program enrollment. MCPS leadership initially indicated students in Grades 7 through 10 were to be given enrollment priority.
- Expand the collection of districtwide data on implementation of *READ 180* and other literacy programs intended to accelerate student achievement. As part of this effort, identify ways to collect information on literacy achievement for students in Grades 9 through 12 so that analyses of targeted reading programs can include high school students.
- Given the patterns of positive outcomes for *READ 180* students in Grades 6 through 8, establish and monitor opportunities for implementing the critical features of *READ 180* (e.g., 90 minutes daily and instructional formats) with high fidelity and high quality.

References

- American Psychological Association. (2001). *Publication manual of the American Psychological Association (5th ed.)*. Washington, DC: Author.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum and Associates.
- Montgomery County Public Schools. (2008). *Our call to action: Pursuit of Excellence*. Rockville, MD: Montgomery County Public Schools.

Author Note. The authors thank Ms. Natalie Wolanin for her assistance in organizing data for analyses, Ms. Catherine A. Jasperse for her efforts to collect implementation data, and *READ 180* teachers who generously provided information on the activities in their classrooms.

Appendix A

Sample

Students were selected for the sample if they were enrolled in a school that offered *READ 180* and scored basic on the MSA reading assessment in spring 2007 (prior to the 2008 academic year). These selection criteria helped minimize preexisting differences among students that might bias results, although further statistical adjustments were required to more accurately compare groups of students. Students in Grades 9 through 12 were excluded from all analyses because a very large proportion were missing assessment data. Table A1 provides demographic information on students who met initial sampling criteria.

Table A1
Characteristics of Students in Sample by *READ 180* Enrollment Group

Characteristic	<i>READ 180</i> for 90 minutes daily (N=716)	<i>READ 180</i> for fewer than 90 minutes daily (N=548)	Not enrolled in <i>READ 180</i> (N=3,523)
	%	%	%
Male	56.4	58.0	56.3
Female	43.6	42.0	43.7
African American	37.8	38.5	36.8
American Indian	0.4	0.5	0.2
Asian American	5.7	8.6	8.0
Hispanic	46.1	37.2	39.4
White	9.9	15.1	15.6
ESOL	10.8	9.9	11.0
FARMS	59.4	55.8	55.9
Special Education	34.8	35.9	33.4
Grade 6	23.6	21.4	45.0
Grade 7	31.4	23.7	31.4
Grade 8	45.0	54.9	23.6

Note. Data most closely reflect status in the third academic quarter.

ESOL=English for Speakers of Other Languages. FARMS=Free and Reduced-price Meals System.

Table A2
Mean Spring 2007 and 2008 MAP-R Scores and Gains for At-Risk Students
by Time in *READ 180* and Student Subgroups

	<i>READ 180</i> for 90 minutes daily				<i>READ 180</i> fewer than 90 minutes daily				No <i>READ 180</i>			
	N	Mean			N	Mean			N	Mean		
		2007	2008	Gain		2007	2008	Gain		2007	2008	Gain
Grade 6												
All	166	195.3	201.8	6.5	108	197.0	200.3	3.4	1134	196.5	201.2	4.7
Male	93	194.4	200.9	6.4	70	196.0	199.0	2.9	620	195.5	200.0	4.5
Female	73	196.4	202.9	6.5	38	198.7	202.8	4.1	514	197.8	202.7	4.9
African American	58	194.9	202.1	7.2	51	198.4	199.6	1.2	419	196.4	200.2	3.9
Asian American	8	189.5	200.6	11.1	11	195.0	203.4	8.4	88	199.1	205.7	6.6
Hispanic	91	196.2	202.3	6.1	35	196.3	198.9	2.6	422	194.8	200.1	5.3
White	9	193.9	194.9	1.0	10	194.6	205.7	11.1	200	199.3	203.6	4.3
ESOL	22	194.5	202.8	8.3	10	190.7	193.5	2.8	114	189.2	198.6	9.4
FARMS	109	195.5	201.7	6.2	56	197.0	198.9	1.9	620	195.2	200.1	4.9
Special Education	45	191.8	197.5	5.8	25	191.6	196.1	4.4	346	189.9	194.5	4.6
Grade 7												
All	208	194.9	203.9	9.0	127	196.9	204.3	7.4	1043	196.0	202.4	6.4
Male	122	193.3	202.8	9.5	70	195.5	202.3	6.8	595	194.0	200.4	6.4
Female	86	197.3	205.5	8.2	57	198.5	206.8	8.3	448	198.6	205.1	6.5
African American	75	193.6	202.5	8.9	37	198.0	207.2	9.2	397	195.7	202.2	6.4
Asian American	16	192.5	203.1	10.6	6	200.3	204.5	4.2	80	200.0	205.8	5.7
Hispanic	102	196.1	204.6	8.5	51	197.5	204.1	6.7	417	194.6	201.4	6.7
White	15	196.9	207.5	10.6	32	194.1	201.3	7.2	148	198.1	204.0	5.9
ESOL	20	188.2	197.3	9.2	9	189.9	202.7	12.8	124	186.8	197.9	11.1
FARMS	123	194.3	203.1	8.8	74	196.0	203.9	7.9	592	195.0	201.7	6.7
Special Education	74	192.3	201.4	9.1	49	191.9	199.6	7.7	379	188.6	195.4	6.8
Grade 8												
All	296	199.8	208.3	8.5	282	199.2	207.4	8.2	1029	203.5	209.3	5.8
Male	162	198.4	207.4	9.0	158	198.0	207.7	9.7	576	202.2	208.1	6.0
Female	134	201.4	209.3	7.8	124	200.7	206.9	6.3	453	205.2	210.7	5.5
African American	117	198.7	207.2	8.5	112	198.5	206.6	8.2	379	203.7	208.5	4.8
Asian American	16	201.3	211.9	10.7	28	199.6	211.5	11.9	94	207.9	214.8	7.0
Hispanic	120	199.4	207.8	8.5	105	198.9	207.1	8.2	396	201.0	207.2	6.2
White	40	203.1	210.3	7.2	36	202.0	207.0	5.0	160	206.5	212.8	6.3
ESOL	30	193.4	206.8	13.3	28	192.4	204.9	12.5	78	191.4	203.9	12.5
FARMS	156	199.2	206.0	6.8	160	198.4	207.1	8.7	556	201.7	207.5	5.9
Special Education	108	196.8	205.0	8.2	110	195.4	202.4	7.0	321	196.0	202.2	6.2

Note. Sample includes students who scored basic on 2007 MSA reading, had 2007 and 2008 MAP-R scores, and were enrolled in schools that offered *READ 180*. Due to small numbers, data for American Indian students are not presented, although they are included in totals. With the exception of the initial sampling criteria, these data do not account for any preexisting differences among groups.

Table A3 presents information on NWEA’s national norms for students in kindergarten through Grade 9. These are the average MAP-R spring (“end of year”) scores for each grade level. Note that in the national sample, average MAP-R scores for students in earlier grades tend to show large score increases from year to year, and yearly increases tend to be much smaller for students in later grades.

Table A3
Mean Spring MAP-R Scores and Differences in Mean Spring
MAP-R Scores for National Student Sample

Grade levels	Spring prior grade	Spring subsequent grade	One-year increase
K to 1	156.3	171.9	15.6
1 to 2	171.9	189.6	17.7
2 to 3	189.6	199.0	9.4
3 to 4	199.0	205.8	6.8
4 to 5	205.8	211.1	5.3
5 to 6	211.1	214.8	3.7
6 to 7	214.8	217.9	3.1
7 to 8	217.9	221.2	3.3
8 to 9	221.2	222.6	1.4

Note. Data are 2008 mean score norms for a national sample of students as reported by NWEA at http://www.nwea.org/assets/downloads/980/Normative%20Data%20Sheet_v2.pdf.

Table A4
Number and Percentage of Students Scoring Basic, Proficient, or Advanced on 2008 MSA in Reading
by Grade and *READ 180* Enrollment

<i>READ 180</i> enrollment group	N	Proficiency level on 2008 MSA in reading		
		Basic %	Proficient %	Advanced %
Grade 6				
<i>READ 180</i> for 90 minutes daily	168	57.7	39.6	2.7
<i>READ 180</i> fewer than 90 minutes daily	117	53.8	44.4	1.8
No <i>READ 180</i>	1199	48.7	47.6	3.7
Grade 7				
<i>READ 180</i> for 90 minutes daily	224	51.3	45.5	3.1
<i>READ 180</i> fewer than 90 minutes daily	130	46.9	49.2	3.8
No <i>READ 180</i>	1123	52.1	43.9	4.0
Grade 8				
<i>READ 180</i> for 90 minutes daily	322	65.8	33.2	1.0
<i>READ 180</i> fewer than 90 minutes daily	300	65.7	33.0	1.3
No <i>READ 180</i>	1159	56.9	39.4	3.7

Note. Sample includes students who scored basic on 2007 MSA reading, had 2007 and 2008 MSA scores, and were enrolled in schools that offered *READ 180*. With the exception of the initial sampling criteria, these data do not account for any preexisting differences among groups. Percentages may sum to 0.1 percentage points above or below 100.0 due to rounding. MSDE identified the following MSA reading scale scores as the ranges for basic, proficient, and advanced performance at each grade level: Grade 6 basic=240 to 380, proficient=381 to 420, advanced=421 to 650. Grade 7 basic=240 to 384, proficient=385 to 424, advanced=425 to 650. Grade 8 basic=240 to 390, proficient=391 to 424, advanced=425 to 650.

Table A5
Descriptive Statistics for Spring 2008 MSA Reading Scores by Grade Level and *READ 180* Enrollment

	<i>READ 180</i> for 90 minutes daily			<i>READ 180</i> for fewer than 90 minutes daily			Not enrolled in <i>READ 180</i>		
	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>N</i>	<i>Mean</i>	<i>SD</i>
Grade 6									
All students	168	377.1	19.7	117	375.0	20.4	1199	379.3	22.9
Male	95	376.2	19.3	77	373.4	21.2	663	377.5	23.2
Female	73	378.3	20.4	40	378.1	18.8	536	381.6	22.4
African American	59	377.6	18.6	54	375.6	21.8	436	379.2	22.6
Asian American	8	364.9	23.5	12	370.8	21.1	93	384.7	20.4
Hispanic	92	379.1	19.3	38	375.4	15.5	458	375.7	22.5
White	9	364.9	23.7	12	374.2	29.0	207	385.3	24.3
ESOL	22	380.2	21.4	12	377.3	14.3	144	371.2	22.4
FARMS	110	378.0	20.6	62	375.4	20.8	667	376.7	22.6
Special Education	48	369.0	16.9	28	362.9	22.1	374	370.6	22.8
Grade 7									
All students	224	383.2	21.4	130	385.3	22.2	1123	381.1	25.6
Male	129	382.0	22.4	71	382.5	21.3	644	377.6	25.0
Female	95	385.0	20.0	59	388.7	23.0	479	385.8	25.7
African American	82	382.7	20.8	40	388.6	17.7	424	381.6	24.3
Asian American	16	375.5	27.8	6	378.8	19.5	83	380.7	27.7
Hispanic	109	384.2	21.5	51	380.2	21.3	453	378.3	25.7
White	17	387.2	16.4	32	392.2	25.6	161	387.2	26.5
ESOL	21	376.4	21.9	9	373.0	18.5	140	370.9	26.9
FARMS	137	383.6	19.3	75	382.3	21.5	645	379.3	23.9
Special Education	82	381.0	22.6	50	380.4	21.2	414	373.2	25.8
Grade 8									
All students	322	383.1	20.1	300	381.9	21.3	1159	387.6	23.4
Male	178	382.0	20.6	169	382.8	20.4	648	386.0	23.6
Female	144	384.5	19.5	131	380.7	22.4	511	389.7	23.0
African American	129	380.9	20.2	116	379.5	19.6	421	386.9	21.7
Asian American	17	386.8	14.0	29	388.7	20.4	102	396.0	24.1
Hispanic	129	383.5	18.9	115	381.4	22.9	458	384.0	22.6
White	44	386.5	24.8	39	385.0	21.2	178	394.0	26.4
ESOL	34	378.3	18.6	33	377.9	23.9	98	375.5	24.5
FARMS	176	380.9	19.1	168	382.2	22.4	630	384.0	22.3
Special Education	119	379.5	20.7	119	376.0	21.1	368	378.6	23.2

Note. Sample includes students who scored basic on 2007 MSA reading, had 2007 and 2008 MSA scores, and were enrolled in schools that offered *READ 180*. Due to small numbers, data for American Indian students are not presented, although they are included in totals. With the exception of the initial sampling criteria, these data do not account for any preexisting differences among groups. MSDE identified the following MSA reading scale scores as the ranges for basic, proficient, and advanced performance at each grade level: Grade 6 basic=240 to 380, proficient=381 to 420, advanced=421 to 650. Grade 7 basic=240 to 384, proficient=385 to 424, advanced=425 to 650. Grade 8 basic=240 to 390, proficient=391 to 424, advanced=425 to 650.

Table A6
Change in MAP-R Scores from Spring 2007 to 2008 by Time in *READ 180* for Each Grade Level
Using Repeated Measures Analysis of Covariance (ANCOVA) General Linear Model (GLM)

Grade and MAP-R assessment year	Adjusted means			Interaction Time-by-Group		Within-year pairwise comparisons		
	Group A <i>READ 180</i> for 90 minutes daily	Group B <i>READ 180</i> for fewer than 90 minutes daily	Group C Not enrolled in <i>READ 180</i>	F	p	Groups compared	p	Effect size
Grade 6								
Spring 2007	193.2	193.1	192.9	2.54	.08	A – B	ns	--
						A – C	ns	--
						B – C	ns	--
Spring 2008	201.3	198.2	199.3			A – B	.02	.24
						A – C	.03	.16
						B – C	ns	--
Grade 7								
Spring 2007	194.1	195.8	195.3	5.26	.01	A – B	ns	--
						A – C	ns	--
						B – C	ns	--
Spring 2008	205.2	205.4	203.7			A – B	ns	--
						A – C	.08	.11
						B – C	.11	.13
Grade 8								
Spring 2007	197.8	197.1	199.2	7.30	<.01	A – B	ns	--
						A – C	.06	-.10
						B – C	.01	-.15
Spring 2008	210.6	209.6	209.6			A – B	ns	--
						A – C	ns	--
						B – C	ns	--

Note. Sample includes students who scored basic on spring 2007 MSA in reading and who were enrolled in schools that offered *READ 180*. Repeated measures analyses were conducted separately for each grade level. Spring 2007 MSA reading score was the covariate. Demographic and special service variables were included as independent main effects. Time-by-group was included as an interaction term. Comparison effect sizes and *p* values are shown only for effect sizes greater than or equal to .10.

Table A7
Spring 2008 MSA Reading Scores by Time in *READ 180* by Grade Level
Using Analysis of Covariance (ANCOVA) General Linear Model (GLM)

Grade	Adjusted means			Main effect of <i>READ 180</i> group		Pairwise comparisons of adjusted means		
	Group A <i>READ 180</i> for 90 minutes daily	Group B <i>READ 180</i> for fewer than 90 minutes daily	Group C Not enrolled in <i>READ 180</i>	F	p	Groups compared	p	Effect size
6	376.6	372.2	376.3	1.69	.18	A – B	.06	.19
						A – C	ns	--
						B – C	.03	-.18
7	381.1	382.1	378.9	2.80	.06	A – B	ns	--
						A – C	ns	--
						B – C	.09	.13
8	384.9	383.2	386.6	5.00	<.01	A – B	ns	--
						A – C	ns	--
						B – C	<.01	-.15

Note. ANCOVA analyses were conducted separately for each grade level. Spring 2007 MSA reading score was the covariate and demographic and special service variables were included as independent main effects. Comparison effect sizes and *p* values are shown only for effect sizes greater than or equal to .10.

Table A8
 Comparing Spring 2008 MAP-R and MSA Scores for *READ 180* Classes Implementing At Least Near-Full
READ 180 Instructional Model at Least Four Days a Week Versus Fewer for Each Grade Level
 Using Hierarchical Linear Modeling (HLM)

	MAP-R models		MSA models	
	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)
Grade 6				
Intercept	200.85	199.41	376.13	374.74
Class-level effect ^a				
Instructional model implemented at least four days a week versus fewer (1, 0)	<i>N/A</i>	4.17*	<i>N/A</i>	3.88 ⁺
Student-level fixed effects ^b				
Spring 2007 MAP-R	0.39*	0.40*	0.59*	0.60*
Male (1, 0)	-2.19 ⁺	-2.18 ⁺	-0.71	-0.61
African American (1, 0)	0.06	-0.25	0.06	-0.01
Asian American (1, 0)	3.24	3.26	-4.86	-4.53
Hispanic (1, 0)	0.36	-0.19	-1.19	-1.45
ESOL (1, 0)	-0.76	-1.52	4.11	3.38
FARMS (1, 0)	-1.66	-1.79	0.18	-0.10
Special Education (1, 0)	-3.31*	-3.68*	-9.17*	-9.32*
Between-class variance				
Variance component	5.19	1.03	6.82	4.88
Chi-square	50.28	37.76	36.02	32.43
<i>p</i>	.07	.39	>.50	>.50
Between-student variance				
Variance component	97.55	97.86	298.46	298.15
Effect size of instructional model versus less	<i>N/A</i>	.37	<i>N/A</i>	.19
<i>N</i> classes implementing model	<i>N/A</i>	10	<i>N/A</i>	10
<i>N</i> classes implementing less than model	<i>N/A</i>	28	<i>N/A</i>	28

(continued)

(continued)

Table A8
 Comparing Spring 2008 MAP-R and MSA Scores for *READ 180* Classes Implementing At Least Near-Full
READ 180 Instructional Model At Least Four Days a Week Versus Fewer for Each Grade Level
 Using Hierarchical Linear Modeling (HLM)

	MAP-R models		MSA models	
	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)
Grade 7				
Intercept	204.23	204.28	384.30	384.50
Class-level effect ^a				
Instructional model implemented at least four days a week versus fewer (1, 0)	<i>N/A</i>	-0.10	<i>N/A</i>	-1.97
Student-level fixed effects ^b				
Spring 2007 MAP-R	0.49*	0.49*	0.60*	0.60*
Male (1, 0)	-1.40	-1.41	-2.69	-2.73
African American (1, 0)	0.47	0.45	-1.24	-0.98
Asian American (1, 0)	0.49	0.49	-7.36	-6.90
Hispanic (1, 0)	0.43	0.43	-4.24	-3.86
ESOL (1, 0)	-2.96	-2.98	-8.32*	-8.54*
FARMS (1, 0)	-1.15	-1.16	-0.04	-0.03
Special Education (1, 0)	-2.88*	-2.89*	-5.23*	-5.36*
Between-class variance				
Variance component	7.59	8.18	18.76	17.18
Chi-square	96.16	96.35	85.87	84.03
<i>p</i>	.01	.01	.04	.05
Between-student variance				
Variance component	80.42	80.32	373.00	374.60
Effect size of instructional model versus less	<i>N/A</i>	.01	<i>N/A</i>	-.09
<i>N</i> classes implementing model	<i>N/A</i>	21	<i>N/A</i>	21
<i>N</i> classes implementing less than model	<i>N/A</i>	45	<i>N/A</i>	45

(continued)

(continued)

Table A8
 Comparing Spring 2008 MAP-R and MSA Scores for *READ 180* Classes Implementing At Least Near-Full *READ 180* Instructional Model At Least Four Days a Week Versus Fewer for Each Grade Level
 Using Hierarchical Linear Modeling (HLM)

	MAP-R models		MSA models	
	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)	Baseline analysis: No class-level comparisons	Whole group and three rotations (with or without wrap up)
Grade 8				
Intercept	207.71	207.63	382.51	382.23
Class-level effect ^a				
Instructional model implemented at least four days a week versus fewer (1, 0)	<i>N/A</i>	0.29	<i>N/A</i>	1.10
Student-level fixed effects ^b				
Spring 2007 MAP-R	0.43*	0.43*	0.43*	0.43*
Male (1, 0)	0.75	0.78	0.40	0.45
African American (1, 0)	-1.13	-1.16	-3.63	-3.68
Asian American (1, 0)	2.04	2.05	2.63	2.70
Hispanic (1, 0)	-1.09	-1.12	-3.09	-3.12
ESOL (1, 0)	-0.83	-0.85	-3.05	-3.08
FARMS (1, 0)	-1.94*	-1.95	-1.75	-1.76
Special Education (1, 0)	-4.54*	-4.53*	-6.59*	-6.53*
Between-class variance				
Variance component	0.37	0.28	14.77	15.43
Chi-square	74.27	73.94	102.72	102.20
<i>p</i>	>.50	>.50	.04	.04
Between-student variance				
Variance component	81.94	82.14	337.69	337.65
Effect size of instructional model versus less	<i>N/A</i>	.03	<i>N/A</i>	.05
<i>N</i> classes implementing model	<i>N/A</i>	18	<i>N/A</i>	18
<i>N</i> classes implementing less than model	<i>N/A</i>	63	<i>N/A</i>	63

Note. *N/A* indicates parameters that are not applicable.

* $p < .01$, + $p < .10$

^a Class level effect is the difference in scale scores after adjusting for students' prior achievement, demographics, and special service receipt.

^b All student-level variables were grand-mean centered and between-class variance (slopes) were fixed (nonvarying).

Appendix B

This section presents findings from a replication of results.

Evaluation Questions

The following questions are addressed in the replication of analyses:

1. Are there differences in reading performance of students enrolled in *READ 180* for 90 minutes daily compared with students enrolled for fewer than 90 minutes daily, after adjusting for students' prior performance, demographics, and service receipt?
2. Are there differences in reading performance of students enrolled in *READ 180* for 90 minutes daily compared with similar students not enrolled in the program, after adjusting for students' prior performance, demographics, and service receipt?

Methodology

Study Samples

Three groups of students were identified according to time in *READ 180* instruction per week: a) students in classes meeting 90 minutes daily, b) students in classes meeting fewer than 90 minutes daily, and c) students not enrolled in *READ 180*. Fifty percent of students not enrolled in *READ 180* were randomly selected to maintain a more balanced design in analyses.

Outcome Measures

The measures included the following assessments for Grades 6, 7, and 8: a) the scale scores from spring 2008 MAP-R and b) the scale scores from spring 2008 MSA reading. Only students who had complete assessment data for both pretest scores (spring 2007 MAP-R and MSA) and posttest scores (spring 2008 MAP-R and MSA) were included in the analyses.

Analysis Procedures

The analysis of covariance was used to compare mean scale scores on the outcome measures for the *READ 180* and non-*READ 180* groups under investigation. To balance the nonequivalent groups, information on measured covariates was incorporated into estimation of the treatment effect through the use of propensity scores. Propensity scores (based on students' prior performance, demographics, and service receipt measures) were computed using a logistic regression model as recommended by previous researchers (Luellen, Shadish, & Clark, 2005). Propensity scores represent "enrollment bias" or the likelihood that a student with certain preexisting characteristics would be enrolled in one of the *READ 180* or comparison groups. The propensity score was divided into five categories and used as a categorical covariate in each of the statistical models. Students' demographics and service receipt measures also were included in the analysis of covariance models to reduce the residual variability of the outcome measures. To test for nonparallelism or interaction (homogeneity of regression slopes), the product term between pretest scores and group variable was included in each of the analytical models. In addition to utilizing propensity scores, the grade-level analyses controlled for students' prior performance (i.e., MSA reading and MAP-R scores from the previous year); demographics; and receipt of special services (FARMS, special education, and/or ESOL).

Findings

Question one: Are there differences in reading performance of students enrolled in READ 180 for 90 minutes daily compared with students enrolled for fewer than 90 minutes daily, after adjusting for students' prior performance, demographics, and service receipt?

Spring 2008 MAP-R

Grade level analyses of MAP-R performance comparing students enrolled in the *READ 180* program for 90 minutes daily and those enrolled for fewer than 90 minutes daily indicate that after factoring out the effects of students' prior performance, demographics, and service receipt, those who were enrolled for 90 minutes daily outperformed those who were enrolled for fewer than 90 minutes daily in Grade 6 (Table B1). Specifically, in Grade 6 there were statistically ($p < 0.05$) and practically (effect size=0.23) significant differences between the two groups of students on the 2008 Grade 6 MAP-R scores. No statistically or practically significant differences were observed for Grades 7 or 8.

Table B1
Adjusted Means and Effect Size for 2008 Reading MAP-R Scores for
READ 180 for 90 Minutes Daily versus Fewer than 90 Minutes Daily

Grade	<i>READ 180</i> for 90 minutes daily		<i>READ 180</i> for fewer than 90 minutes daily		Difference in adjusted means	Effect size
	<i>N</i>	Adjusted mean scale score	<i>N</i>	Adjusted mean scale score		
6	166	201.46	108	198.82	2.65*	0.23
7	208	204.40	127	203.80	0.60**	0.05
8	296	210.56	282	210.04	0.52***	0.05

* $F=4.23$; $p=0.04$

** $F=0.25$; $p=0.61$

*** $F=0.17$; $p=0.68$

Spring 2008 MSA Reading

Grade level analyses of the two groups of *READ 180* students revealed that on average, there were no statistically significant differences between the MSA reading scores of students enrolled in the program for 90 minutes daily and those students enrolled for fewer than 90 minutes daily, after factoring out the effects of students' prior performance, demographics, and service receipt (Table B2). The analyses of the effect sizes found a small difference between the two groups of students in Grade 6 in favor of those enrolled in *READ 180* for 90 minutes daily (effect size=0.18).

Table B2
Adjusted Means and Effect Size for 2008 MSA Reading Scores for
READ 180 for 90 Minutes Daily versus Fewer than 90 Minutes Daily

Grade	<i>READ 180</i> for 90 minutes daily		<i>READ 180</i> for fewer than 90 minutes daily		Difference in adjusted means	Effect size
	<i>N</i>	Adjusted mean scale score	<i>N</i>	Adjusted mean scale score		
6	168	378.69	117	375.02	3.67*	0.18
7	224	373.30	130	375.18	-1.88**	-0.09
8	322	385.44	300	383.88	1.56***	0.08

* $F=2.54$; $p=0.11$

** $F=0.77$; $p=0.38$

*** $F=0.61$; $p=0.44$

Question two: Are there differences in reading performance of students enrolled in *READ 180* for 90 minutes daily compared with similar students not enrolled in the program, after adjusting for students' prior performance, demographics, and service receipt?

Spring 2008 MAP-R

Grade level analyses indicate that Grade 7 and 8 students enrolled in *READ 180* for 90 minutes daily scored significantly higher on their MAP-R test than a group of similar students who were not enrolled in the program, after controlling for students' prior performance, demographics, and service receipt (Table B3). However, the calculated effect sizes reveal that the significant differences between the two groups of students was only large enough in Grade 7 to be useful in an educational setting (effect size=0.16). The analyses also found very small effects of the program on MAP-R performance for students in Grade 6 (effect size=0.12) and Grade 8 (effect size=0.11).

Table B3
Adjusted Means and Effect Size for 2008 Reading MAP-R Scores for
READ 180 for 90 Minutes Daily versus No *READ 180*

Grade	<i>READ 180</i> for 90 minutes daily		No <i>READ 180</i>		Difference in adjusted means	Effect size
	<i>N</i>	Adjusted mean scale score	<i>N</i>	Adjusted mean scale score		
6	166	201.54	569	200.01	1.52*	0.12
7	208	210.30	505	208.10	2.20**	0.16
8	296	211.70	528	210.27	1.44***	0.11

* $F=3.53$; $p=0.06$
 ** $F=6.31$; $p=0.01$
 *** $F=7.73$; $p=0.006$

Spring 2008 MSA Reading

On average, there were no statistically significant differences in MSA performance between the group of students enrolled in *READ 180* for 90 minutes and a similar group of students who were not enrolled in the program, after controlling for students' prior performance, demographics, and service receipt in all of the three comparisons made—Grades 6, 7, and 8 (Table B4). The analyses of effect sizes indicate a very small effect of the program for Grade 7 (effect size=0.13) only.

Table B4
Adjusted Means and Effect Size for 2008 Reading MSA Scores for
READ 180 for 90 Minutes Daily versus No *READ 180*

Grade	<i>READ 180</i> for 90 minutes daily		No <i>READ 180</i>		Difference in adjusted means	Effect size
	<i>N</i>	Adjusted mean scale score	<i>N</i>	Adjusted mean scale score		
6	168	374.08	599	372.70	1.39*	0.06
7	224	385.62	545	382.48	3.14**	0.13
8	322	387.66	586	388.71	-1.06***	-0.05

* $F=0.62$; $p=0.43$
 ** $F=3.23$; $p=0.07$
 *** $F=0.37$; $p=0.55$

References

Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29 (6), 530–558.

