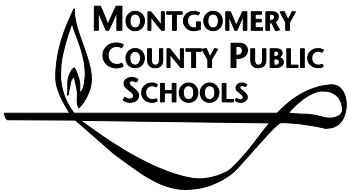


# Examination of the Impact of Year 1 of Monitoring Instructional Reading Levels Strategy on the Performance of Grades 3-5 Students on Measures of Academic Progress in Reading (MAP-R)

**Office of Shared Accountability**

**February 2017**

**Nyambura Susan Maina, Ph.D. and Natalie Wolanin**

Montgomery County Public Schools

**OFFICE OF SHARED ACCOUNTABILITY**

**850 Hungerford Drive**
**Rockville, Maryland 20850**
301-279-3553

**Dr. Jack R. Smith**
*Superintendent of Schools*

**Dr. Janet S. Wilson**
*Associate Superintendent*

# Table of Contents

# List of Tables



# List of Figures

# Executive Summary

Monitoring Instructional Reading Levels (MIRL) is a districtwide strategy being implemented in MCPS to improve reading instruction and increase reading performance for K–5 students. The objective of the MIRL strategy is to provide effective instruction by increasing systematic implementation of high quality guided reading, monthly documentation of students' instructional reading levels, analysis, and use of monthly reading data for instructional planning (Appendix A). This report is one in a series of four reports from the MIRL evaluation study. The purpose of this report is to examine effects of the MIRL strategy on Grades 3–5 students' performance in reading during the 2015–2016 school year by comparing the performance of students in Grades 3–5 on Measures of Academic Progress in Reading (MAP-R) to performance in 2014–2015, before the implementation of MIRL

MCPS students in Grades 3–5 from across all schools in the district comprised the study sample. Students in Grade 3 ($N$ = 11,587), Grade 4 ($N$ = 11,544), and Grade 5 ($N$ = 11,331) during the school year 2015−2016 were included in the MIRL Year 1 group. Students in Grade 3 ($N$ = 11,403), Grade 4 ($N$ = 10,632), and Grade 5 ($N$ = 11,259) during the school year 2014−2015 were included in the No MIRL group. Across the three grade levels and two groups, approximately 14% of the students were Asian, 21% were African American or Black, 30% were White, and 30% were Hispanic/Latino. Additionally, approximately 38% received the Free and Reduced-price Meal System (FARMS), 12% received Special Education services, 20-22% of Grade 3 students received English for Speakers of Other Languages (ESOL) services and 6-11% of Grade 4 and 5 students received ESOL services.

Two different analyses on the outcome measure were used in this study in order to provide alternative ways to examine the data: 1) the spring Measures of Academic Progress in Reading, (MAP-R) mean Rasch Unit (RIT) scores and 2) the proportion of students attaining a RIT score equivalent or above the 50$^{th}$ spring reading percentile based on 2015 MAP national norms. For each grade level, an analysis of covariance (ANCOVA) was conducted to assess the effect of MIRL Year 1 on spring RIT scores while simultaneously controlling for pre-existing achievement level (corresponding fall MAP-R scores), school attendance, and a propensity score to statistically balance for any demographic differences in students from the two groups. Further, differences in the proportions of students with scores equivalent or above the 50$^{th}$ spring reading percentile were examined using two-way contingency table analyses (Pearson's chi-square). Lastly, effect sizes for group differences were computed to examine whether the detected effects were practically meaningful in an educational setting ($d$>=0.15).

## Summary of Findings

### *Was there a difference in Grade 3−5 spring MAP-R RIT scores from 2015–2016 (MIRL) compared to 2014–2015 (no MIRL)?*

After controlling for reading performance in fall, demographic characteristics, and daily attendance, the analyses revealed statistically significant mean differences between MIRL Year 1 students and No MIRL Grade 3, 4, and 5 students, in favor of MIRL Year 1 students. While

statistically significant differences were found, the magnitude of the between group effect size was less than the threshold set (*d*=>0.15) for a practically meaningful effect in an educational setting. At the subgroup level, some adjusted mean MAP-R RIT scores for MIRL Year 1 group were significantly higher than their counterpart in the No MIRL group. However, effect sizes for the group differences were not large enough to be considered practically significant in an educational setting.

***Was there a difference in proportions of Grade 3−5 students scoring at or above the 50th percentile on the 2015–2016 spring MAP-R (MIRL Year 1) compared to 2014–2015 (no MIRL)?***

Overall, over 60% of Grades 3, 4, and 5 students attained a score equivalent to or above the respective grade-level national 50th spring reading achievement percentile score during both years (Figure 1a)..

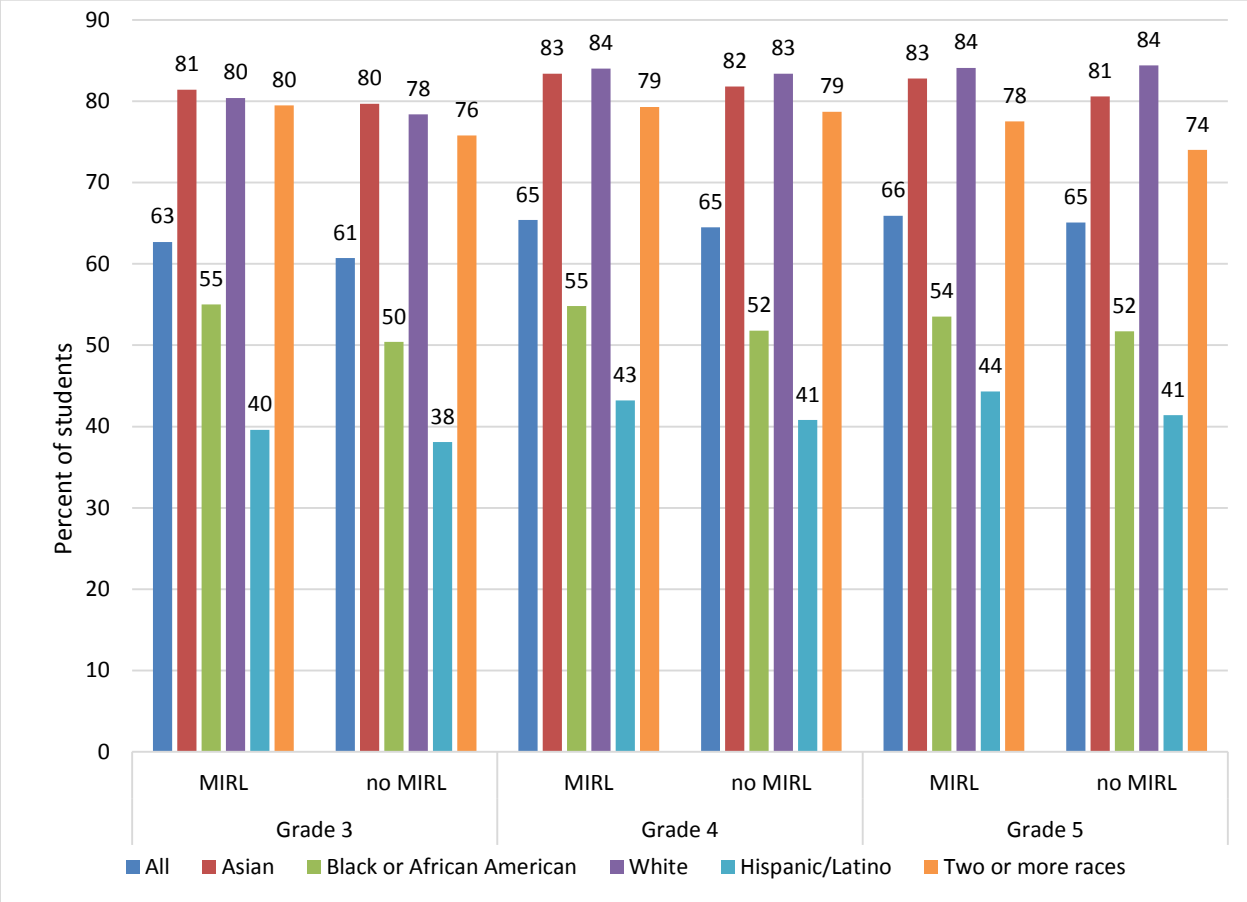


*Figure 1*a. Percent of students with scores equal or above the 50th spring percentile RIT score by race/ethnicity.

Compared with the No MIRL group (2014−2015), a statistically higher proportion of MIRL Year 1 (2015−2016) Grade 3 students attained this performance level in 2015−2016 (63% vs. 61%);

however, the effect size for the group differences indicated that the differences were not large enough to be considered practically significant in an educational setting.

Further, the proportions of MIRL Year 1 students (2015−2016) with spring RIT scores at or exceeding their respective grade-level 50th percentile scores were significantly higher than the students in the previous year with No MIRL for several student subgroups. A moderate effect size (*d*=0.57) was detected only for MIRL Year 1 students receiving ESOL (Figure 1b) services in Grade 4 (13% vs. 5%), indicating that the difference in proportions was practically meaningful in an educational setting.

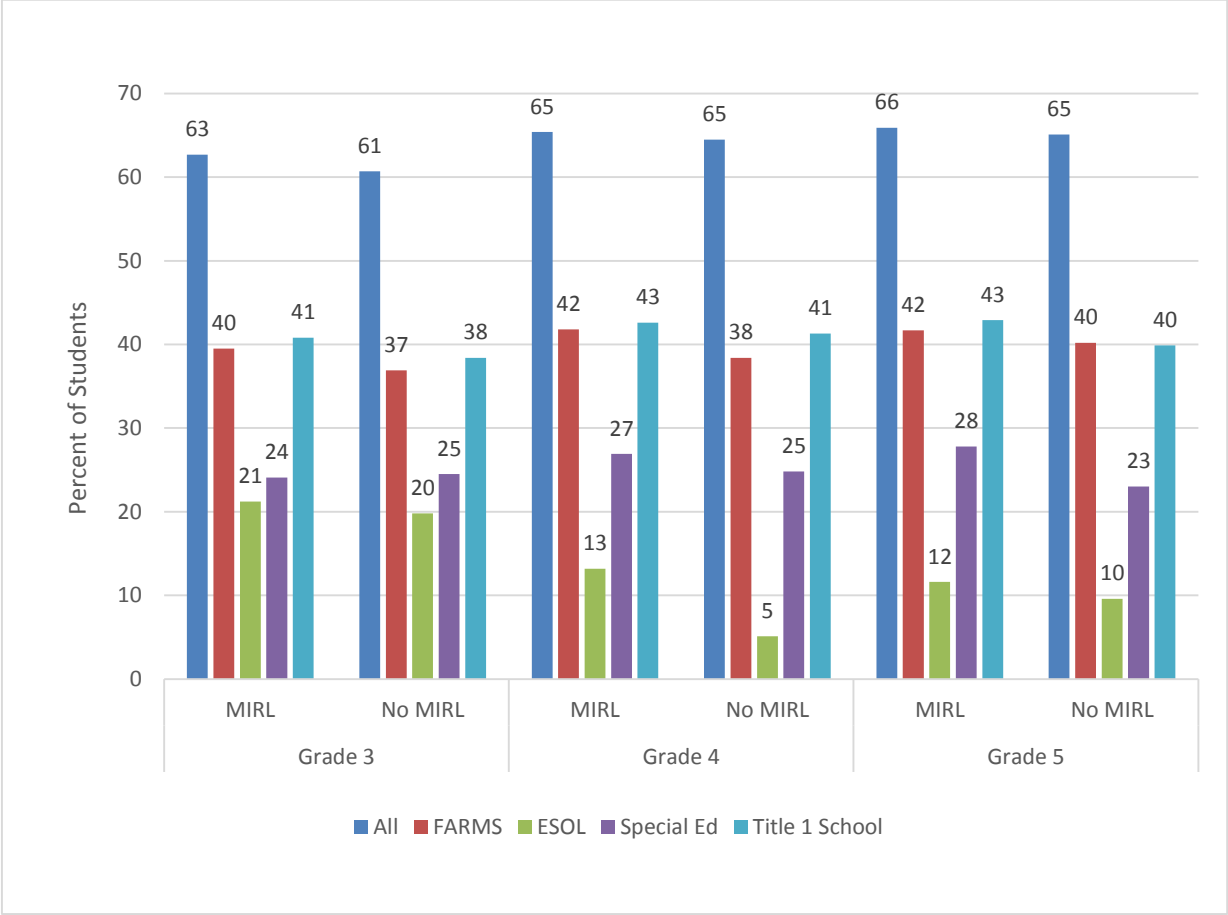


Figure 1b. Percent of students with score equal or above the 50th spring percentile by service receipt.

## Recommendations

- Continue to examine the effects of the MIRL strategy on student performance in reading with each successive year of implementation. The results of a survey of teachers' experiences with implementing MIRL (Maina & Wolanin, 2016) revealed that some aspects of the MIRL strategy were not fully implemented during the first year. As such, the results in this report may not reflect the eventual effects of MIRL; examining effects of MIRL after teachers and

students have had sufficient time to master the use of a strategy as intended, will be more definitive.

- Examine the extent to which MIRL practices are implemented as intended for a) subgroups of students and b) schools still performing at low levels in reading before and after MIRL Year 1.
- Extend the study of effects of the MIRL strategy to include growth in specific aspects of reading: 1) Literary Text, Informational Text, 2) Vocabulary, 3) Writing Expression, and 4) Knowledge and use of Language Conventions.

# Examination of the Impact of Year 1 of Monitoring Instructional Reading Levels Strategy on the Performance of Grades 3-5 Students on Measures of Academic Progress in Reading, (MAP-R)

Nyambura Susan Maina, Ph.D. and Natalie Wolanin

## Background

The overall goal of the implementation of the MIRL strategy is to increase student performance in reading. Through ongoing observation and documentation of students' reading progress during guided reading, it is expected that teachers will be able to diagnose students' immediate needs and then use the information to adjust instruction.

### Theoretical Perspective/Theory of Action for MIRL

MCPS emphasizes that an effective instructional program makes proper use of three types of assessments categorized as: a) assessments AS learning, b) assessments FOR learning, and c) assessments OF learning (MCPS, 2015a; MCPS, 2015b). MIRL is guided by the premise that "If we systematically monitor students' instructional reading levels (accuracy, fluency, and comprehension) and use these data to support and program for students, then classroom instruction will be more strategic and reading achievement will increase for all students" (MCPS, 2015b p. 11). As such, MIRL is a strategy which uses "assessment FOR learning" as a tool for diagnosing a student's immediate need so that instruction can be adjusted appropriately to improve student performance in reading.

Ultimately, the consistent and full implementation of these MIRL practices across schools and classrooms is expected to bring about 1) established use of monitoring of instructional levels during guided reading for all K–5 students, 2) improved reading performance for all K–5 students, and 3) progress toward reducing achievement gaps among student subgroups. The purpose of this evaluation was to examine effects of the MIRL strategy on students' achievement in reading during the 2015–2016 school year by comparing the performance of Grades 3–5 students to performance in 2014–2015, before the implementation of MIRL. Information related to student performance on the 2016 Assessment Program in Primary Reading (Kindergarten to Grade 2) is presented in a separate report (Wang, 2016).

### Components of MIRL in 2015–2016

Starting in fall 2015, all Grade K–5 teachers in all elementary schools were expected to implement the following specified MIRL practices (Appendix A; Maina & Wolanin, 2016):

- High quality guided reading

- Ongoing monitoring of instructional reading levels and monthly collecting and entering data into a monthly reading data collection tool. The monthly online data collection tool is a tool managed centrally by the Office of the Chief Technology Officer (OCTO) in MCPS;
- Use of monthly data to adjust instruction;
- Collaborative planning at the grade-level; and
- Activities and structures instituted to support the implementation of the above listed components.

## Selected Literature Highlights

*Challenges that limit quantification of impact of MIRL or similar practices.* Limited literature was found that conclusively demonstrates the impact of formative assessments as learning or a combination of the practices that make up the MIRL strategy on student achievement in reading. Indeed, several aspects of formative assessments make it challenging to reliably quantify the impact of such practices or interventions on student performance. The term formative assessment represents an evolving set of practices commonly described as a planned process in which assessment-elicited evidence is used: a) by teachers to adjust their ongoing teaching and learning procedures to improve students' achievement of intended instructional outcomes, and b) by students to adjust their current learning tactics (Bennett, 2011; McManus, 2008; Popham, 2008; Heritage, 2010; Griffin 2007). Assessment as learning or assessments administered in order to gauge what students do and do not know, so that teachers can modify their instruction accordingly, are generally accepted as a viable tool for improving student achievement. The weaknesses in methodological research to investigate formative assessment to date do not invalidate the promise of these studies' findings and do not refute that the studies support assertions of formative assessment's efficacy.

*Impact of formative assessment.* One influential study conducted by Paul Black and Dylan Wiliam in 1998 is associated with the strongest claims for effectiveness of using formative assessments. This study was a review of 250 empirical studies on classroom-based assessment practices and their impact on a mixed set of student populations from a variety of academic settings and grade ranges. The findings showed that student-learning gains in classrooms where formative assessment was used were among the largest ever reported in educational interventions (Hanover Research, 2014). Notably, the largest gains were reportedly most evident for low achievers (Black & William, 1998; 2009; Bennett, 2011; Popham, 2008). Similar results were reported by other researchers who analyzed studies that were more explicitly aligned with K–12 forms of formative assessment (Kingston & Nash, 2009; Pinchok & Brandt, 2009). The researchers found practically meaningful effects that were "large enough to indicate formative assessment can be a significant and readily achievable source of improved student learning" (Bennett, 2011). A notable finding of Black and Wiliam's work is that formative assessment has a disproportionately beneficial impact on low achieving students. They stated that, "while formative assessment can help all pupils, it yields particularly good results with low achievers by concentrating on specific problems with their work and giving them a clear understanding of what is wrong and how to put it right." (Black & Wiliam, 1998).

*Trends in MCPS reading performance of Grade 3–5 students.* MCPS started using the Measures of Academic Progress in Reading (MAP-R) as a milestone for monitoring district-wide

performance in reading in Grades 3, 5, and 8 in 2013–2014 (MCPS, 2014; MCPS, 2015a). In 2013–2014 the majority of students in Grades 3 and 5 met or exceeded the end-of-year benchmarks (75% and 86% respectively) as measured by MAP-R. Nonetheless, performance gaps between ethnic/racial categories and for students receiving special services and those not receiving special services varied at each grade level. For example in Grade 3, the proportion of students who were proficient ranged from 56 percent for Hispanic/Latino students to 89 percent for White students. At the same time, 56 percent of Grade 3 students who received FARMS services, 39 percent of students who received special education services, and 57 percent of students receiving English for Speakers of Other Languages (ESOL) services met or exceeded established benchmarks.

### Evaluation questions

This study was guided by the question "To what extent did reading performance for all Grade 3–5 students improve in 2015–2016 compared with 2014–2015?" Specific questions were:

1. Was there a difference in Grade 3−5 spring MAP-R RIT scores from 2015–2016 (MIRL) compared to 2014–2015 (no MIRL)?
2. Was there a difference in proportions of Grade 3−5 students scoring at or above the 50$^{th}$ percentile on the 2015–2016 spring MAP-R (MIRL) compared to 2014–2015 (no MIRL)?

## Methodology

### Design

A quasi-experimental design that compared outcomes for students who had access to MIRL and with those who did not was applied. The purpose of this evaluation was to examine effects of the MIRL strategy on student achievement in reading during the 2015–2016 school year by comparing the performance of Grade 3–5 students to Grade 3−5 students from the prior year (2014−2015) who did not have access to MIRL.

### Study Samples

MCPS students in Grades 3, 4, and 5 during school years 2014−2015 and 2015−2016 from across all schools in the district comprised the study sample. To address the first evaluation question, students must have had both spring and fall MAP-R scores from 2015−2016 to be included in the MIRL Year 1 group ($N$ = 34,462) and have had both spring and fall MAP-R scores from 2014−2015 to be included in the non-MIRL group ($N$ = 33,294). As shown in Table 1a, the proportion of 2015-2016 Grade 3, 4, and 5 student demographics were similar to those of 2014−2015 with the exception of Grade 4 current ESOL students, where the percentage of students in 2015−2016 nearly twice (11%) that of 2014−2015 (6%).

Table 1a
Descriptive Statistics of the Analytic Sample for Evaluation Question 1

| Characteristic | Grade 3 | | | | Grade 4 | | | | Grade 5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MIRL Year 1 (2015−2016) (*N* =11,587 ) | | No MIRL (2014−2015) (*N* =11,403 ) | | MIRL Year 1 (2015−2016) (*N* =11,544 ) | | No MIRL (2014−2015) (*N* =10,632 ) | | MIRL Year 1 (2015−2016) (*N* =11,331) | | No MIRL (2014−2015) (*N* =11,259 ) | |
| | n | % | n | % | n | % | n | % | n | % | n | % |
| American Indian | 23 | 0.2 | 31 | 0.3 | 30 | 0.3 | 24 | 0.2 | 24 | 0.2 | 14 | 0.1 |
| Asian | 1,605 | 13.9 | 1,612 | 14.1 | 1,642 | 14.2 | 1,601 | 15.1 | 1,704 | 15.0 | 1,689 | 15.0 |
| African American or Black | 2,387 | 20.6 | 2,303 | 20.2 | 2,356 | 20.4 | 2,156 | 20.3 | 2,388 | 21.1 | 2,324 | 20.6 |
| Pacific Islander | 8 | 0.1 | 7 | 0.1 | 7 | 0.1 | 1 | 0.0 | 2 | 0.0 | 6 | 0.1 |
| White | 3,449 | 29.8 | 3,485 | 30.6 | 3,460 | 30.0 | 3,412 | 32.1 | 3,574 | 31.5 | 3,736 | 33.2 |
| Hispanic/Latino | 3,514 | 30.3 | 3,348 | 29.4 | 3,446 | 29.9 | 2,953 | 27.8 | 3,127 | 27.6 | 2,946 | 26.2 |
| Two or more races | 601 | 5.2 | 617 | 5.4 | 603 | 5.2 | 485 | 4.6 | 512 | 4.5 | 544 | 4.8 |
| Male | 5,960 | 51.4 | 5,858 | 51.4 | 5,940 | 51.5 | 5,388 | 50.7 | 5,766 | 50.9 | 5,723 | 50.8 |
| Female | 5,627 | 48.6 | 5,545 | 48.6 | 5,604 | 48.5 | 5,244 | 49.3 | 5,565 | 49.1 | 5,536 | 49.2 |
| Current FARMS | 4,437 | 38.3 | 4,365 | 38.3 | 4,381 | 38.0 | 3,764 | 35.4 | 4,073 | 35.9 | 3,946 | 35.0 |
| Current ESOL | 2,369 | 20.4 | 2,351 | 20.6 | 1,211 | 10.5 | 662 | 6.2 | 875 | 7.7 | 723 | 6.4 |
| Current Special Ed | 1,261 | 10.9 | 1,275 | 11.2 | 1,398 | 12.1 | 1,211 | 11.4 | 1,311 | 11.6 | 1,228 | 10.9 |
| Title 1 School | 2,189 | 18.9 | 2,133 | 18.7 | 2,093 | 18.1 | 1,796 | 16.9 | 1,860 | 16.4 | 1,879 | 16.7 |

To address the second evaluation question, students must have had spring MAP-R scores from 2015−2016 to be included in the MIRL Year 1 group (*N* =35,435). Students must have had spring MAP-R scores from 2014−2015 to be included in the sample for the non-MIRL group (*N* =34,292). The demographic profiles of 2015−2016 Grade 3, 4, and 5 student was similar to those of 2014−2015 with the exception of Grade 4 current ESOL students, where the percentage of students was higher in 2015−2016 (11%) compared to 2014−2015 (7%). Table 1b presents demographic information for the students in each of these two analytical samples by grade.

Table 1b
Descriptive Statistics of the Analytic Sample for Evaluation Question 2

| Characteristics | Grade 3 | | | | Grade 4 | | | | Grade 5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MIRL Year 1 (2015−2016) (*N* =11,961) | | No MIRL (2014−2015) (*N* =11,807) | | MIRL Year 1 (2015−2016) (*N* =11,845) | | No MIRL (2014−2015) (*N* =10,921) | | MIRL Year 1 (2015−2016) (*N* =11,629) | | No MIRL (2014−2015) (*N* =11564) | |
| | n | % | n | % | n | % | n | % | n | % | n | % |
| American Indian | 23 | 0.2 | 31 | 0.3 | 30 | 0.3 | 25 | 0.2 | 26 | 0.2 | 16 | 0.1 |
| Asian | 1,663 | 13.9 | 1,666 | 14.1 | 1,689 | 14.3 | 1,640 | 15.0 | 1,752 | 15.1 | 1,734 | 15.0 |
| African American or Black | 2,482 | 20.8 | 2,405 | 20.4 | 2,430 | 20.5 | 2,240 | 20.5 | 2,464 | 21.2 | 2,404 | 20.8 |
| Pacific Islander | 9 | 0.1 | 7 | 0.1 | 7 | 0.1 | 1 | 0.0 | 2 | 0.0 | 6 | 0.1 |
| White | 3,512 | 29.4 | 3,563 | 30.2 | 3,517 | 29.7 | 3,471 | 31.8 | 3,622 | 31.1 | 3,793 | 32.8 |
| Hispanic/Latino | 3,656 | 30.6 | 3,506 | 29.7 | 3,554 | 30.0 | 3,055 | 28.0 | 3,238 | 27.8 | 3,061 | 26.5 |
| Two or more races | 616 | 5.2 | 629 | 5.3 | 618 | 5.2 | 489 | 4.5 | 525 | 4.5 | 549 | 4.7 |
| Male | 6,165 | 51.5 | 6,073 | 51.4 | 6,086 | 51.4 | 5,550 | 50.8 | 5,915 | 50.9 | 5,875 | 50.8 |
| Female | 5,796 | 48.5 | 5,734 | 48.6 | 5,759 | 48.6 | 5,371 | 49.2 | 5,714 | 49.1 | 5,688 | 49.2 |
| Current FARMS | 4,650 | 38.9 | 4,593 | 38.9 | 4,537 | 38.3 | 3,930 | 36.0 | 4,239 | 36.5 | 4,111 | 35.6 |
| Current ESOL | 2,574 | 21.5 | 2,557 | 21.7 | 1,356 | 11.4 | 785 | 7.2 | 1,004 | 8.6 | 865 | 7.5 |
| Current Special Ed | 1,303 | 10.9 | 1,329 | 11.3 | 1,429 | 12.1 | 1,255 | 11.5 | 1,337 | 11.5 | 1,249 | 10.8 |
| Title 1 School | 2,285 | 19.1 | 2,238 | 19.0 | 2,153 | 18.2 | 1,864 | 17.1 | 1,929 | 16.6 | 1,962 | 17.0 |

## Outcome Measures

Two different analyses on the outcome measure were used in this study in order to provide alternative ways to examine the data: 1) the spring Measures of Academic Progress in Reading, (MAP-R) mean Rasch Unit (RIT) scores and 2) the proportion of students attaining a RIT score equivalent or above the 50th spring reading percentile based on 2015 MAP national norms. The MAP assessment is an untimed, computer adaptive assessment which is aligned to the state standards. The RIT score is an equal-interval vertical scale and is independent of grade level; RIT scores range from about 150 to 300. Grade 3−5 spring RIT scores and the reading proficiency levels (scores at or above the 50th percentile), were used to examine differences in performance of students.

## Data Analysis Procedures.

The following analyses procedures were completed separately for each grade level and for subgroups within each grade level.

*Analyses of covariance*. The ANCOVA provides a way of statistically controlling for the effects of variables of concern that are not the independent variable(s) in the study. For that reason, ANCOVA for continuous outcome variables (MAP-R RIT scores) was conducted to detect any statistical differences in reading performance between MIRL Year 1 students (2015−2016) and students without MIRL (2014–2015) while simultaneously controlling for pre-existing achievement level (corresponding fall MAP-R scores), and school attendance. Further, for each year, propensity scores were computed using FARMS status, receipt of ESOL, and race/ethnicity and used as covariates to statistically control for preexisting differences between the two student populations as recommended by literature (Shadish, Luellen, & Clark, 2005).

*Descriptive Statistics.* Students in the study sample data file with RIT scores corresponding to or higher than the 50th spring reading achievement percentile score were flagged. The 50th percentile scores are as follows: Grade 3=199; Grade 4=206; and Grade 5=212 (Appendix B; Thum, & Hauser, 2015; Wang, Zhao, & Addison, (2016). Then, differences in the proportions of students with scores equivalent or exceeding the 50th percentile score were examined using two-way contingency table analyses (Pearson's chi-square), to determine whether differences in the percentages between MIRL Year 1 and No MIRL were statistically significant.

*Computation of effect sizes*. Effect sizes were computed[1] to a) assess the magnitude of the observed differences in mean RIT scores at the respective grade level; and b) assess the magnitude of differences in proportions of students with scores equivalent or higher than the 50th percentile score from 2014–2015 to 2015–2016. The effect sizes were used to judge the practical significance of the observed differences (American Psychological Association, 2010). An effect of 0.15 was considered an appropriate level for the threshold for practical significance. Most studies compare the overall program effect size to Cohen's (1988) definitions of a small effect within the behavioral sciences, $d$ = 0.20; a medium effect, $d$ = 0.50; and a large effect, $d$ = 0.80 (Wasik & Slavin, 1993;

[1] $ES = \frac{\overline{X}_{MIRL} - \overline{X}_{no\,MIRL}}{s_{pooled}}$ for the RIT scores; Effect sizes from differences in proportions (logit $d$) = ln (OR)/pi/sqrt 3.

Cohen, 1988). Across 346 comparisons of education programs for at-risk children, the average effect size, adjusted for methodological characteristics, was $d = 0.12$ (Borman, Hewes, Overman, & Brown, 2002). According to Lipsey et al. (2012), the mean effect size of interventions that focus on curriculum or broad instructional programs is 0.13 and the median effect size is 0.08. As such, in this study, an effect of 0.15 was considered an appropriate level for the threshold for practical significance.

### Strengths and Limitations

A limitation to this study is the lack of an experimental design; therefore, causality may not be inferred from this study because students were not randomly assigned to treatment groups. Additionally, the comparison group was from the prior year, before the initiative, and may consist of unknown influences or differences that may not exist in the year of the treatment.

A strength of this study was the use of ANCOVA to examine any differences between MAP-R RIT scores among students with access to MIRL (2015−2016) and students without MIRL (2014–2015) while simultaneously controlling for pre-existing ability (corresponding fall MAP-R scores), school attendance, and demographics. This strategy improved the internal validity of the study by controlling for selection bias. Also, the MIRL Year 1 and No MIRL groups were similar in their student demographic profiles as evidenced in Table 1 and Table 2 of Appendix B.

## Results

### *Evaluation Question 1. Was there a difference in Grade 3−5 spring MAP-R RIT scores from 2015–2016 (MIRLYear 1) compared to 2014–2015 (No MIRL)?*

An analysis of covariance (ANCOVA) was conducted on spring MAP-R RIT scores to evaluate the effect of implementing MIRL in 2015−2016. The results are presented by grade level; student subgroups are presented within each grade-level.

*Grade 3.* The adjusted means, mean group differences, and effect sizes for spring RIT scores of Grade 3 students are presented in Table 2a. The ANCOVA showed that the differences in adjusted mean spring RIT scores between MIRL Year 1 and No MIRL was statistically significant for Grade 3, $F(1, 22985)=5.38$, $p=0.02$, in favor of the MIRL group. However, no effect size for group differences was detected, indicating the observed differences in mean RIT scores were not practically significant.

At the subgroup level, statistically significant differences in the adjusted mean spring RIT scores between MIRL Year 1 and No MIRL were found for Black or African American students, $F$ (1, 4,685) =11.2, $p=0.00$, students receiving FARMS services, $F$ (1, 8797) =7.68, $p=0.01$, and students receiving special education services, $F$ (1, 2,531) =6.54, $p=0.01$ in favor of the MIRL Year 1 students. For every Grade 3 subgroup examined, the effect sizes for differences in performance between the two years were negligible or did not meet the threshold ($d<0.15$) set to indicate a practically meaningful impact of MIRL on the spring RIT scores.

Table 2a
Adjusted Mean Spring MAP-R RIT Scores and Effect Sizes for Differences in Performance Between MIRL Year 1 and No MIRL Grade 3 Students

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Subgroup | N | Adjusted mean | Std Error | N | Adjusted Mean | Std Error | Adjusted Mean Difference | | F | Sig (*p*) | Effect Size |
| All | 11,587 | 201.91 | .07 | 11,403 | 201.67 | .08 | .24 | * | 5.38 | .02 | .01 |
| Asian | 1,605 | 209.54 | .18 | 1612 | 209.43 | .18 | .11 | | .17 | .69 | .05 |
| Black or African American | 2,387 | 197.78 | .17 | 2,303 | 196.96 | .18 | . 82 | *** | 11.20 | .00 | .05 |
| White | 3,449 | 208.88 | .13 | 3,485 | 208.88 | .13 | .00 | | .00 | 1.00 | .00 |
| Hispanic/Latino | 3,514 | 193.16 | .14 | 3,348 | 192.79 | .14 | .37 | | 3.47 | .06 | .02 |
| Two or more races | 601 | 207.46 | .31 | 617 | 208.08 | .31 | -.62 | | 2.04 | .15 | -.04 |
| FARMS | 4,437 | 192.49 | .13 | 4,365 | 192.00 | .13 | .50 | ** | 7.68 | .01 | .03 |
| ESOL | 2,369 | 186.79 | .18 | 2,351 | 186.54 | .18 | .25 | | .95 | .33 | .02 |
| Special Ed | 1,261 | 184.07 | .27 | 1,275 | 183.10 | .27 | .97 | ** | 6.54 | .01 | .05 |

* *P* < .05, ** p < .01, *** p< .001

*Grade 4.* Table 2b presents the results of the ANCOVA for Grade 4 students. The ANCOVA showed that the difference in adjusted mean spring RIT scores between MIRL Year 1 and No MIRL students was statistically significant for Grade 4, $F(1, 22{,}171)=15.98$, $p=0.00$, in favor of the MIRL Year 1 students. However, the effect size was .02, indicating that the difference in scores for MIRL Year 1 and No MIRL students was not practically significant.

At the subgroup level, significant differences in the adjusted mean spring RIT scores for MIRL Year 1 and No MIRL students were found for the subgroups of Black or African American, $F(1, 4{,}507)=8.26$, $p=0.00$; Hispanic/Latino, $F(1, 6394)=20.80$, $p = .00$; students receiving FARMS services, $F(1, 8{,}140)=20.22$, $p=0.00$; students receiving ESOL services, $F(1, 1868)=10.36$, $p=.00$; and students receiving special education services, $F(1, 2{,}604)=10.21$, $p=0.00$ in favor of MIRL students. For every Grade 4 subgroup examined, the effect sizes for differences in performance between the two years were negligible or did not meet the threshold ($d<0.15$) to indicate a practically meaningful impact of MIRL on their performance on the spring MAP-R reading assessment.

Table 2b
Adjusted Mean Spring MAP-R RIT Scores and Effect Sizes for Differences in Performance Between MIRL Year 1 and No MIRL Grade 4 Students

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Subgroup | N | Adjusted mean | Std Error | N | Adjusted Mean | Std Error | Adjusted Mean Difference | | F | Sig (*p*) | Effect Size |
| All | 11,544 | 210.42 | .07 | 10,632 | 210.01 | .07 | .41 | *** | 15.98 | .00 | .02 |
| Asian | 1,642 | 217.29 | .17 | 1,601 | 217.49 | .17 | -.20 | | .75 | .39 | .06 |
| Black or African American | 2,356 | 205.61 | .17 | 2,156 | 204.91 | .18 | .70 | *** | 8.26 | .00 | .05 |
| White | 3,460 | 217.52 | .12 | 3,412 | 217.57 | .12 | -.06 | | .10 | .75 | .00 |
| Hispanic/Latino | 3,446 | 201.73 | .14 | 2,953 | 200.81 | .15 | .92 | *** | 20.80 | .00 | .06 |
| Two or more races | 603 | 215.94 | .28 | 485 | 215.56 | .32 | .38 | | .78 | .38 | .00 |
| FARMS | 4,381 | 200.94 | .12 | 3,764 | 200.13 | .13 | .81 | *** | 20.22 | .00 | .05 |
| ESOL | 1,211 | 187.51 | .27 | 662 | 186.04 | .37 | 1.47 | *** | 10.36 | .00 | .10 |
| Special Ed | 1,398 | 192.99 | .24 | 1,211 | 191.86 | .26 | 1.13 | *** | 10.21 | .00 | .06 |

*$p$ < .05, ** p < .01, *** p< .001

*Grade 5.* Table 2c presents the results of the ANCOVA for Grade 5 students. The ANCOVA showed that the difference in adjusted mean spring RIT scores between MIRL Year 1 and No MIRL was statistically significant for Grade 5, $F($1, 22,585)=22.61, $p$=.00, in favor of the MIRL group. However, the negligible effect size indicated that overall, the observed difference in performance between the two groups was not practically significant.

At the subgroup level, significant differences between the adjusted mean spring RIT scores were found for the subgroups of Black or African American, $F$(1,4,707)=12.15, $p$=.00; Hispanic/Latino, $F$(1, 6,068)=24.75, $p$ = .00; students receiving FARMS services, $F$(1, 8,014)=31.87, $p$=.00; students receiving ESOL services, $F$(1, 1,593)=6.10, $p$=.01; and students receiving special education services, $F$(1, 2,534)=8.67, $p$=.00 in favor of the MIRL Year 1 students. For all of the subgroups, the effect sizes for group differences were negligible or very small; indicating that the observed difference in performance for Grade 5 MIRL Year 1 and their No MIRL counterparts on the spring MAP-R did not reach the threshold of being practically meaningful in an educational setting.

Table 2c
Adjusted Mean Spring MAP-R RIT Scores and Effect Sizes for Differences in Performance Between MIRL Year 1 and No MIRL Grade 5 Students

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Subgroup | N | Adjusted mean | Std Error | N | Adjusted Mean | Std Error | Adjusted Mean Difference | | F | Sig (*p*) | Effect Size |
| All | 11,331 | 217.05 | .07 | 11,259 | 216.58 | .07 | .47 | *** | 22.61 | .00 | .03 |
| Asian | 1,704 | 223.88 | .16 | 1,689 | 223.70 | .16 | .18 | | .58 | .45 | .05 |
| Black or African American | 2,388 | 211.61 | .16 | 2,324 | 210.80 | .17 | .81 | *** | 12.15 | .00 | .05 |
| White | 3,574 | 223.96 | .12 | 3,736 | 224.01 | .12 | -.06 | | .11 | .74 | .00 |
| Hispanic/Latino | 3,127 | 208.44 | .15 | 2,946 | 207.44 | .14 | 1.00 | *** | 24.75 | .00 | .06 |
| Two or more races | 512 | 221.30 | .30 | 544 | 220.52 | .29 | .78 | | 3.55 | .06 | .05 |
| FARMS | 4,073 | 207.79 | .12 | 3,946 | 206.79 | .13 | .97 | *** | 31.87 | .00 | .07 |
| ESOL | 875 | 193.86 | .35 | 723 | 192.69 | .32 | 1.17 | ** | 6.10 | .01 | .08 |
| Special Ed | 1,311 | 199.65 | .24 | 1,228 | 198.63 | .25 | 1.02 | *** | 8.67 | .00 | .05 |

*p < .05, ** p < .01, *** p< .001

***Evaluation Question 2. Was there a difference in proportions of Grade 3−5 students scoring at or above the 50th percentile on the 2015–2016 spring MAP-R (MIRL Year 1) compared to 2014–2015 (No MIRL)?***

To address question two, the percentages of students from MIRL Year 1 and students in the previous year with No MIRL groups who had RIT scores at or higher than the 50th spring reading student achievement percentile score at their respective grade level were examined. A chi-square test of independence was performed to examine differences in proportion of MIRL Year 1 and No MIRL students. The results are presented by grade level; student subgroups are presented within each grade-level.

Regardless of grade level, over 70% of the MIRL Year 1 and No MIRL students from the Asian, White, and Two or more races subgroups had RIT scores equivalent to or above the 50th spring reading student achievement percentile score. Conversely, across the three grade-levels, less than one-half of Hispanic/Latino students, students receiving ESOL, FARMS, or special education services, and students from Title I schools, attained a RIT score equivalent to or above the 50th spring reading student achievement percentile score for both years

Table 3a presents the percentages of Grade 3 students from the MIRL Year 1 and No MIRL groups with scores at or exceeding the 50th percentile score on the spring MAP-R assessment. Overall, over 60% of Grade 3 students from both groups attained the 50th percentile RIT score. Compared with the No MIRL group (61%), a statistically higher proportion of MIRL Year 1 students (63%) attained the 50th percentile, $x^2$ (1, $N$ = 23,768) = 10.34, $p$=.001. However, the negligible effect size ($d$=0.05) indicated the difference did not reach the threshold of being practically meaningful in an educational setting (Table 3a).

Compared to their peers in the No MIRL group, statistically higher proportions of MIRL Year 1 Black or African American, White, and FARMS students had a score at or above the 50th percentile score (Table 3a): Black or African American, $x^2$ (1, $N$ = 4,887) = 10.19, p=.001; White, $x^2$ (1, $N$ = 7075) = 4.42, p=.036; and FARMS, $x^2$ (1, $N$ = 9243) = 7.07, p = .008 students. However, for all

the subgroups, none of effects size for the group differences met the threshold ($d$=>0.15) for being practically significant.

Table 3a
Number and Percent of Grade 3 Students with Scores at Above the 50th Percentile Spring MAP-R Score

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | Sig. ($p$) | Effect Size |
|---|---|---|---|---|---|---|---|---|
| Student subgroup | *N* | n | % | *N* | n | % | | |
| All | 11,961 | 7,503 | 62.7 | 11,807 | 7167 | 60.7 | .001 | .05 |
| Asian | 1,663 | 1,354 | 81.4 | 1,666 | 1328 | 79.7 | .213 | .06 |
| Black or African American | 2,482 | 1,365 | 55.0 | 2,405 | 1213 | 50.4 | .001 | .10 |
| White | 3,512 | 2,824 | 80.4 | 3,563 | 2793 | 78.4 | .036 | .07 |
| Hispanic/Latino | 3,656 | 1,449 | 39.6 | 3,506 | 1335 | 38.1 | .177 | .04 |
| Two or more races | 616 | 490 | 79.5 | 629 | 477 | 75.8 | .116 | .12 |
| FARMS | 4,650 | 1,839 | 39.5 | 4,593 | 1693 | 36.9 | .008 | .06 |
| ESOL | 2,574 | 546 | 21.2 | 2,557 | 507 | 19.8 | .171 | .05 |
| Special Ed | 1,303 | 314 | 24.1 | 1,329 | 326 | 24.5 | .796 | -.01 |
| Title 1 School | 2,285 | 933 | 40.8 | 2,238 | 860 | 38.4 | .098 | .06 |

Table 3b presents the percent of Grade 4 students from the MIRL Year 1 group who met the 50th percentile on the spring MAP-R compared to the No MIRL group. The proportion of MIRL participants who performed at this level was comparable to the No MIRL participants (65% for both).

Compared to their peers in the No MIRL group, statistically higher proportions of the following MIRL Year 1 Grade 4 student subgroups had a score at or above higher than the 50th percentile score (Table 2b): Black or African American, $x^2$ (1, $N$ = 4,670) = 4.18, p=.041; Hispanic/Latino, $x^2$ (1, $N$ = 6,609) = 3.99, p=.046; FARMS, $x^2$ (1, $N$ =8,467) = 10.37, p = .001; and ESOL, $x^2$ (1, $N$ = 2,141) = 35.57, p=.000. The moderate effect size among the ESOL subgroup ($d$ = .57) was practically significant. This indicated that the differences in the proportions of Grade 4 ESOL MIRL Year 1 and No MIRL ESOL students was substantial and meaningful in an educational setting.

Table 3b
Number and Percent of Grade 4 Students with Scores at or Above the 50th Percentile Spring MAP-R Score

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | Sig. (*p*) | Effect Size |
|---|---|---|---|---|---|---|---|---|
| Student subgroup | *N* | n | % | *N* | n | % | | |
| All | 11,845 | 7,747 | 65.4 | 10,921 | 7,039 | 64.5 | .134 | .02 |
| Asian | 1,689 | 1,408 | 83.4 | 1,640 | 1,342 | 81.8 | .243 | .06 |
| Black or African American | 2,430 | 1,331 | 54.8 | 2,240 | 1,160 | 51.8 | .041 | .07 |
| White | 3,517 | 2,955 | 84.0 | 3,471 | 2,895 | 83.4 | .486 | .03 |
| Hispanic/Latino | 3,554 | 1,536 | 43.2 | 3,055 | 1,246 | 40.8 | .046 | .06 |
| Two or more races | 618 | 490 | 79.3 | 489 | 385 | 78.7 | .824 | .02 |
| FARMS | 4,537 | 1,897 | 41.8 | 3,930 | 1,508 | 38.4 | .001 | .08 |
| ESOL | 1,356 | 179 | 13.2 | 785 | 40 | 5.1 | .000 | .57 |
| Special Ed | 1,429 | 385 | 26.9 | 1,255 | 311 | 24.8 | .202 | .06 |
| Title 1 School | 2,153 | 917 | 42.6 | 1,864 | 769 | 41.3 | .392 | .03 |

*Note.* American Indian category is not reported due to small sample size.

Table 3c presents the results of the percent of Grade 5 students from the MIRL Year 1 group who met the 50th percentile on the spring MAP-R compared to the No MIRL group. A chi square analysis showed the percent of MIRL participants (66%) meeting the 50th percentile was comparable to the non MIRL participants (65%) with a negligible effect size.

Compared to their peers in the No MIRL group, statistically higher proportions of two subgroups of Grade 5 MIRL Year 1 had a score at or above the 50th percentile score (Table 3c): Hispanic/Latino, $x^2$ (1, $N$ = 6299) = 5.27, $p$=.022; and Special Education, $x^2$ (1, $N$ = 2586) = 7.98, $p$=.005. However the effect sizes for these groups and the other subgroups did not meet the threshold ($d$<=0.15) for being practically significant.

Table 3c
Number and Percent of Grade 5 Students with Scores at or Above the 50th Percentile Spring MAP-R Score

| | MIRL Year 1 (2015–2016) | | | No MIRL (2014–2015) | | | Sig. (*p*) | Effect Size |
|---|---|---|---|---|---|---|---|---|
| Student subgroup | *N* | n | % | *N* | n | % | | |
| All | 11,629 | 7,668 | 65.9 | 11,564 | 7,530 | 65.1 | .187 | .02 |
| Asian | 1,752 | 1,450 | 82.8 | 1,734 | 1,397 | 80.6 | .094 | .08 |
| Black or African American | 2,464 | 1,319 | 53.5 | 2,404 | 1,243 | 51.7 | .202 | .04 |
| White | 3,622 | 3,047 | 84.1 | 3,793 | 3,200 | 84.4 | .776 | -.01 |
| Hispanic/Latino | 3,238 | 1,433 | 44.3 | 3,061 | 1,267 | 41.4 | .022 | .06 |
| Two or more races | 525 | 407 | 77.5 | 549 | 406 | 74.0 | .173 | .11 |
| FARMS | 4,239 | 1,767 | 41.7 | 4,111 | 1,654 | 40.2 | .178 | .03 |
| ESOL | 1,004 | 116 | 11.6 | 865 | 83 | 9.6 | .220 | .11 |
| Special Ed | 1,337 | 372 | 27.8 | 1,249 | 287 | 23.0 | .005 | .14 |
| Title 1 School | 1,929 | 828 | 42.9 | 1,962 | 783 | 39.9 | .056 | .07 |

## Summary

This study was designed to assess the impact of the MIRL strategy on students' achievement in reading during the 2015–2016 school year by comparing the performance of Grades 3–5 students to performance of students from the prior year (2014−2015) who did not have access to MIRL. On average, the students in the MIRL Year 1 group attained statistically significantly higher spring RIT score than students in the No MIRL group; however the negligible effect sizes for group differences were not large enough to be considered practically significant in an educational setting.

For all three grades, the average spring MAP-R RIT scores for students in the MIRL Year 1 groups were significantly higher for the subgroups of Black or African Americans, students receiving FARMS and students receiving Special Education services. In addition, the spring MAP-R RIT scores for MIRL Year 1 for Grade 4 and 5 Hispanic/Latino students and students receiving ESOL services were significantly higher compared to their No MIRL counterparts. However, effect sizes for group differences were not large enough to be considered practically significant in an educational setting.

Grade 3, 4, and 5 students from the MIRL Year 1 group had significantly higher proportion of students with scores equal to or exceeding the 50th spring reading student achievement percentile score; although, the negligible effect sizes indicated the differences were not practically meaningful. Similarly, many of the subgroups of students from the MIRL Year 1 group were more likely than their peers from the No MIRL group to score at or above the 50th spring percentile: Black or African American and FARMS students from Grades 3 and 4; Hispanic/Latino students from Grades 4 and 5; White students from Grade 3; ESOL students from Grade 4 and Special

Education students from Grade 5. However, a moderate, practically significant effect size for group differences was detected only for Grade 4 students receiving ESOL services.

## Discussion

The findings from this study reflect the trend in the achievement levels for MCPS students which vary widely by student subgroups; a trend that MIRL aims to reverse. While spring RIT scores for 2015–2016 Grades 3, 4, 5 students were significantly higher than spring RIT scores in 2014−2015, the observed differences were not practically significant. Despite the increases in performance levels for most student subgroups in MIRL Year 1, still, less than one-half of Hispanic/Latino students, students receiving ESOL, FARMS, or special education services, and students from Title I schools attained a RIT score equal to or above the 50$^{th}$ percentile score for both years.

The results of a teachers' survey on experiences with MIRL provided evidence that implementing MIRL had influenced many instructional practices as intended: implementing guided reading, using formative assessments to inform guided reading instruction, planning collaboratively for reading instruction, and consistently monitoring student performance (Maina & Wolanin, 2016;). Recent studies on MIRL (Maina & Wolanin, 2016; Maina, 2016; Maina, 2017) also revealed that some aspects of the MIRL strategy were not fully implemented during the first year. In particular, findings from a survey of classroom teachers as well as the observation of instructional planning meetings revealed the following:

- Less than one third of classroom teacher survey respondents reported a) coordinating additional support for students not meeting monthly grade-level targets, b) coordinating instruction with special education staff in a typical instructional planning meeting, or c) using the module pathways developed by OCIP to address instructional needs of English language learners, students with disabilities or highly able students.
- Language development data were rarely used or discussed by the sample of instructional planning teams observed.
- Limited instances of discussions of strategies and resources for differentiating instruction for students receiving special services were observed.
- Compared to grade level peers, the proportions of ESOL Level 1 and 2 students with complete MIRL information related to accuracy, fluency, comprehension, and reading instructional levels students were lower.

As such, the results in this report may not reflect the eventual effects of MIRL. One can surmise that in subsequent years, when fully implemented, MIRL could be expected to yield greater impact on reading performance.

## Recommendations

- Continue to examine the effects of the MIRL strategy on student performance in reading with each successive year of implementation. The results of a survey of teachers' experiences with implementing MIRL revealed that some aspects of the MIRL strategy were not fully implemented during the first year. As such, the results in this report may not reflect the eventual effects of MIRL; examining effects of MIRL after teachers and students have had sufficient time to master the use of a practice or strategy as intended, will be more definitive.
- Examine the extent to which MIRL practices are implemented as intended for a) subgroups of students and b) schools still performing at low levels in reading before and after MIRL Year 1.
- Extend the study of effects of the MIRL strategy to include growth in specific aspects of reading: 1) Literary Text, Informational Text, 2) Vocabulary, 3) Writing Expression, and 4) Knowledge and use of Language Conventions.

## Acknowledgements

The appreciation is extended to Ms. Laura Evans, Supervisor, Elementary Integrated Curriculum (EIC) Reading and her team of instructional specialists' valuable collaboration on the planning of this study; Dr. Shahpar Modarresi, Dr. Cara Jackson, and Ms. Julie Wade, of the Program Evaluation (PE) Unit for the valuable feedback in the development of the report; and Ms. Marilyn Powell, Applied Research Unit for her guidance on the utilization of the MAP-R data.

# References

Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice, 18,* 5 – 25.

Black, P. J., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80,* 139–48.

Black, P., & Wiliam, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation and Accountability,* 21, 5–31.

Borman, Hewes, Overman, & Brown (2002). *Comprehensive School Reform and Achievement: A Meta-Analysis.* Review of Educational Research Summer 2003, Vol. 73, No. 2, pp. 125–230.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Fuch, L. S., & Fuch, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children, 53,* 199-208.

Griffin, P. (2007). The comfort of competence and the uncertainty of assessment. *Studies in Educational Evaluation, 33,* 87–99.

Hanover Research, 2014. The Impact of Formative Assessment and Learning Intentions on Student Achievement. Retrieved December 20, 2016 from http://www.hanoverresearch.com/media/ The- Impact-of-Formative-Assessment-and-Learning-Intentions-on-Student-Achievement.pdf.

Heritage, M. (2010). *Formative Assessment and Next-Generation Assessment Systems: Are we Losing an Opportunity?* Council of Chief State School Officers, Washington, DC.

Kingston, N., & Nash, B. (2009). *The efficacy of formative assessment: A meta-analysis.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., and Busick, M.D. (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms. (NCSER, 2013–3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.

Maina, S. & Wolanin, N. (2016). *Status of Implementation of the Monitoring Instructional Reading Levels (MIRL) Strategy during 2015−2016 School Year: Findings from 2016 Classroom Teacher Survey.* Rockville, MD. Montgomery County Public Schools.

McManus, S. (2008). *Attributes of effective formative assessment.* North Carolina Department of Public Instruction, for the Formative Assessment for Students and Teachers Collaborative. Council of Chief School Officers. Retrieved from October 11, 2015 from http://www.ccsso.org/Documents/2008/Attributes_ of_Effective_2008.pdf.

MCPS (2014). *2014 Annual Report to the Community*. Rockville, MD. Montgomery County Public Schools.

MCPS (2015a). *Districtwide Assessment Strategy: Elementary Assessments*. Retrieved September 8, 2015 from http://www.boarddocs.com/mabe/mcpsmd/ Board.nsf/files/ 9ZX24L 76095D/$file/Policy%20IKA%20Tentative%20Action.pdf.

MCPS (2015b). *Elementary Principal Curriculum Update Meeting April 30 and May 1, 2015.* Rockville, MD. Montgomery County Public Schools.

Northwest Evaluation Association. (2011). RIT scale norms: For use with Measures of Academic Progress (MAP®) and MAP® for Primary Grades. Portland, OR.

Pinchok, N., & Brandt, W. C. (2009). *Connecting formative assessment research to practice: An introductory guide for educators.* Washington, DC: Learning Point Associates. Retrieved from http://www.learningpt.org/pdfs/FormativeAssessment.pdf.

Popham, W. J. (2008). *Transformative assessment.* Alexandria, VA: ASCD.

Thum Y. M., & Hauser, C. H. (2015). *NWEA 2015 MAP Norms for Student and School Achievement Status and Growth. NWEA Research Report*. Portland, OR: NWEA.

Wasik, B., & Slavin, R. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. Reading Research Quarterly, 28(2), 179-200.

Underwood, V. (2010). *The effect of guided reading instruction on reading achievemen*t. Retrieved December 2016 from http://pqdtopen.proquest.com/doc/577393098.html?FMT=ABS.

Shadish W. R., Luellen J. K., & Clark M. H. (2005). Propensity scores: An introduction and experimental test. Evaluation Review, 29, 530-558. doi:10.1177/0193841X0575596.

Wang, H. *Student performance on the 2016 assessment program in primary reading (Kindergarten to Grade 2).* Rockville, MD. Montgomery County Public Schools.

Wang, H., Zhao, H., & Addison, K. (2016). Linking PARCC and MAP Assessments for Students in Montgomery County Public Schools. Rockville, MD. Montgomery County Public Schools.

# Appendix A. Logic Model for MIRL: Activities, Expected Results, and Anticipated Outcomes for 2015–2016

| Needs and Issues (Rationale for MIRL) | Inputs | Outputs/Results | | Outcomes | |
|---|---|---|---|---|---|
| | (Resources and Structures Instituted) | Activities | Participation Metrics | Expected Short Term Changes | Expected Lasting Changes |
| • Decreasing performance in reading (K–5) and variability in performance levels as measured by MCPS AP-PR and MAP-R<br>• Limited observation, documentation, analysis, and use of formative data monitoring) of reading performance/of reading levels at specified intervals throughout the year | • Professional development (PD) sessions<br>• 2015–2016 MCPS Elementary Literacy Plan<br>• Online monthly reading data collection tool (OCTO)<br>• School level common team planning structures<br>• Monthly principal curriculum updates<br>• Elementary Literacy Instructional Core Team (ICT) | • Clarify and articulate Elementary Literacy Plan and vision for reading instruction at school level<br>• Facilitate ongoing PD to school staff reflecting on results and best practices<br>• Regularly assess and document reading levels during guided reading<br>• Introduction of Monthly Reading Data Collection Tool<br>• Discuss reading data at regular intervals (monthly collaborative teams and principal curriculum updates)<br>• Ongoing strategic use of formative reading data to adjust instruction | • PD sessions and Modules /Topics covered during PD for teachers<br>• Number and frequency of PD sessions for school leaders<br>• % Teachers attending PD<br>• % Administrators attending PD<br>• Types of structures and processes in place at school level and who is involved<br>• Extent of use of monthly reading data collection tool/ periodic online reports<br>• Frequency and structure of school level team meetings related to use of formative reading data to plan instruction<br>• Frequency and attendance at principal's curriculum update meetings | • Increased familiarity of teachers and school leaders with literacy plan, data collection tools, and monitoring of reading performance<br>• Initiating and formalizing processes and structures for collecting, entering, and using reading data<br>• Increased monitoring (observation and documentation) of instructional levels during guided reading<br>• Consistent use of monthly reading data collection tool to document instructional reading levels<br>• Ongoing coordinated analyses and use of reading data to inform instructional practices and support student learning<br>• Increased use of monitoring data to adjust instruction | • Established use of monitoring of instructional levels during guided reading levels for all K–5 students<br>• Improved reading performance for all K–5 students<br>• Progress toward reducing achievement gaps |

Logic Model for MIRL 2015–2016

# Appendix B. MAP Spring Reading Student Achievement National Percentiles

Table C.1.6: Spring Reading Student Achievement Percentiles , *continued*

| Pct | K | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | Pct |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *50* | 158 | 178 | 189 | 199 | 206 | 212 | 216 | 218 | 220 | 222 | 221 | 222 | *50* |
| *51* | 158 | 178 | 189 | 199 | 206 | 212 | 216 | 219 | 220 | 222 | 222 | 223 | *51* |
| *52* | 159 | 178 | 189 | 199 | 207 | 213 | 216 | 219 | 221 | 223 | 222 | 223 | *52* |
| *53* | 159 | 179 | 190 | 200 | 207 | 213 | 217 | 219 | 221 | 223 | 222 | 224 | *53* |
| *54* | 159 | 179 | 190 | 200 | 207 | 213 | 217 | 220 | 222 | 223 | 223 | 224 | *54* |
| *55* | 160 | 179 | 191 | 201 | 208 | 214 | 218 | 220 | 222 | 224 | 223 | 225 | *55* |
| *56* | 160 | 180 | 191 | 201 | 208 | 214 | 218 | 220 | 222 | 224 | 224 | 225 | *56* |
| *57* | 160 | 180 | 191 | 201 | 209 | 214 | 218 | 221 | 223 | 225 | 224 | 225 | *57* |
| *58* | 161 | 180 | 192 | 202 | 209 | 215 | 219 | 221 | 223 | 225 | 225 | 226 | *58* |
| *59* | 161 | 181 | 192 | 202 | 209 | 215 | 219 | 222 | 224 | 226 | 225 | 226 | *59* |
| *60* | 161 | 181 | 193 | 202 | 210 | 216 | 219 | 222 | 224 | 226 | 226 | 227 | *60* |
| *61* | 162 | 182 | 193 | 203 | 210 | 216 | 220 | 222 | 224 | 226 | 226 | 227 | *61* |
| *62* | 162 | 182 | 193 | 203 | 210 | 216 | 220 | 223 | 225 | 227 | 227 | 228 | *62* |
| *63* | 162 | 182 | 194 | 204 | 211 | 217 | 221 | 223 | 225 | 227 | 227 | 228 | *63* |
| *64* | 163 | 183 | 194 | 204 | 211 | 217 | 221 | 224 | 226 | 228 | 227 | 229 | *64* |
| *65* | 163 | 183 | 195 | 204 | 212 | 217 | 221 | 224 | 226 | 228 | 228 | 229 | *65* |
| *66* | 163 | 184 | 195 | 205 | 212 | 218 | 222 | 224 | 227 | 229 | 228 | 230 | *66* |
| *67* | 164 | 184 | 195 | 205 | 212 | 218 | 222 | 225 | 227 | 229 | 229 | 230 | *67* |
| *68* | 164 | 184 | 196 | 206 | 213 | 219 | 223 | 225 | 227 | 229 | 229 | 231 | *68* |
| *69* | 165 | 185 | 196 | 206 | 213 | 219 | 223 | 226 | 228 | 230 | 230 | 231 | *69* |
| *70* | 165 | 185 | 197 | 207 | 214 | 220 | 223 | 226 | 228 | 230 | 230 | 232 | *70* |
| *71* | 165 | 186 | 197 | 207 | 214 | 220 | 224 | 227 | 229 | 231 | 231 | 232 | *71* |
| *72* | 166 | 186 | 198 | 207 | 215 | 220 | 224 | 227 | 229 | 231 | 231 | 233 | *72* |
| *73* | 166 | 186 | 198 | 208 | 215 | 221 | 225 | 227 | 230 | 232 | 232 | 233 | *73* |
| *74* | 166 | 187 | 198 | 208 | 216 | 221 | 225 | 228 | 230 | 232 | 232 | 234 | *74* |
| *75* | 167 | 187 | 199 | 209 | 216 | 222 | 226 | 228 | 231 | 233 | 233 | 234 | *75* |
| *76* | 167 | 188 | 199 | 209 | 216 | 222 | 226 | 229 | 231 | 233 | 234 | 235 | *76* |
| *77* | 168 | 188 | 200 | 210 | 217 | 223 | 227 | 229 | 232 | 234 | 234 | 235 | *77* |
| *78* | 168 | 189 | 200 | 210 | 217 | 223 | 227 | 230 | 232 | 234 | 235 | 236 | *78* |
| *79* | 169 | 189 | 201 | 211 | 218 | 224 | 228 | 230 | 233 | 235 | 235 | 237 | *79* |
| *80* | 169 | 190 | 201 | 211 | 218 | 224 | 228 | 231 | 233 | 236 | 236 | 237 | *80* |
| *81* | 169 | 190 | 202 | 212 | 219 | 225 | 229 | 231 | 234 | 236 | 237 | 238 | *81* |
| *82* | 170 | 191 | 203 | 212 | 220 | 225 | 229 | 232 | 234 | 237 | 237 | 239 | *82* |
| *83* | 170 | 191 | 203 | 213 | 220 | 226 | 230 | 233 | 235 | 237 | 238 | 239 | *83* |
| *84* | 171 | 192 | 204 | 214 | 221 | 226 | 230 | 233 | 236 | 238 | 239 | 240 | *84* |
| *85* | 171 | 193 | 204 | 214 | 221 | 227 | 231 | 234 | 236 | 239 | 239 | 241 | *85* |
| *86* | 172 | 193 | 205 | 215 | 222 | 228 | 232 | 235 | 237 | 239 | 240 | 241 | *86* |
| *87* | 173 | 194 | 206 | 216 | 223 | 228 | 232 | 235 | 238 | 240 | 241 | 242 | *87* |
| *88* | 173 | 195 | 207 | 216 | 223 | 229 | 233 | 236 | 239 | 241 | 242 | 243 | *88* |
| *89* | 174 | 195 | 207 | 217 | 224 | 230 | 234 | 237 | 239 | 242 | 243 | 244 | *89* |
| *90* | 175 | 196 | 208 | 218 | 225 | 231 | 235 | 238 | 240 | 243 | 244 | 245 | *90* |
| *91* | 175 | 197 | 209 | 219 | 226 | 232 | 235 | 238 | 241 | 244 | 245 | 246 | *91* |
| *92* | 176 | 198 | 210 | 220 | 227 | 232 | 236 | 239 | 242 | 245 | 246 | 247 | *92* |
| *93* | 177 | 199 | 211 | 221 | 228 | 234 | 237 | 241 | 243 | 246 | 247 | 248 | *93* |
| *94* | 178 | 200 | 212 | 222 | 229 | 235 | 239 | 242 | 245 | 247 | 248 | 250 | *94* |
| *95* | 179 | 201 | 214 | 223 | 230 | 236 | 240 | 243 | 246 | 249 | 250 | 251 | *95* |
| *96* | 181 | 203 | 215 | 225 | 232 | 238 | 241 | 245 | 248 | 250 | 252 | 253 | *96* |
| *97* | 182 | 205 | 217 | 227 | 234 | 239 | 243 | 247 | 250 | 252 | 254 | 256 | *97* |
| *98* | 185 | 207 | 220 | 230 | 237 | 242 | 246 | 249 | 252 | 255 | 257 | 259 | *98* |
| *99* | 188 | 211 | 224 | 234 | 241 | 246 | 250 | 253 | 257 | 260 | 262 | 263 | *99* |

*Source.* Thum Y. M., & Hauser, C. H. (2015). *NWEA 2015 MAP Norms for Student and School Achievement Status and Growth. NWEA Research Report*. Portland, OR: NWEA.