



**Impact Evaluation of *Read 180*
in Miami-Dade County Public
Schools: An Application of the
Regression Discontinuity
Approach**

June 2025

**Impact Evaluation of *Read 180* in
Miami-Dade County Public Schools:
An Application of the Regression Discontinuity Approach**

Douglas Ready
Rebecca Shmoys

Teachers College, Columbia University

June 2025

Table of Contents

Introduction.....	2
Background.....	3
Data and Methods.....	5
Results	10
Conclusion	17
References	18

Introduction

Many students struggle to develop literacy skills, even those in the upper grades. Nationally, only thirty percent of eighth graders scored at or above proficient levels on the most recent administration of the NAEP literacy assessment (U.S. Department of Education, 2025). Since the educational disruptions associated with COVID-19, differences in student literacy performance across racial/ethnic and socioeconomic groups have grown even more pronounced (Fahle et. al., 2023; Guryan & Ludwig, 2023; Peters et al., 2023). Understanding which curricular and instructional approaches best support striving readers is crucial for their academic success, as well as their longer-term social and economic prospects.

This paper describes a large-scale evaluation of *Read 180*, which the Miami-Dade County Public Schools used as a Tier 2 literacy intervention during the 2022-23 academic year. Students reading below grade level in the district were to participate in *Read 180*. Using data from this implementation, we conducted a series of analyses to explore the extent to which participation in *Read 180* improved student literacy outcomes. The sections below begin with a description of *Read 180* and the extant research on its efficacy. We then describe our data and analytic methods. The findings section provides both descriptive and causal results regarding the associations between *Read 180* participation and student literacy development. We close with a summary and discussion of our findings.

Background

The Read 180 Model

Read 180 employs a blended learning, station rotation approach that includes teacher-led instruction in whole- and small-group settings, as well as scaffolded practice via computer-based and independent reading rotations (see HMH, 2024). Each class session begins with teacher-led, whole-class instruction in reading skills and strategies, content-area and academic vocabulary, writing, conventions, and academic discussions. Teachers then organize student rotations between computer-based instruction stations, teacher-led small-group lessons, and independent reading stations, each of which lasts approximately 20 minutes. Classes conclude with a whole-group wrap-up that reinforces what students learned that day. For more details about *Read 180* and its components, see HMH, 2024.

Prior Research on *Read 180*

A large number of studies have explored associations between *Read 180* and student learning. Rather than describing this full body of work, we instead focus on rigorous studies that use analytic methods and data that produce causal estimates of *Read 180* impacts. These studies typically have research designs that compare outcomes among *Read 180* participants to those of comparable students who did not participate in *Read 180*. This is accomplished through experimental designs, including randomized controlled trials, or via quasi-experiment approaches, such as difference-in-differences designs. Studies that simply explore the associations between *Read 180* usage rates and student learning do not meet this requirement, given that students with higher levels of program engagement may differ from other students, and those differences may in turn be associated with student learning. Another important consideration in determining the rigor of individual studies is the extent to which they leverage student or school sample sizes that are large enough to support the detection of statistically significant program effects.

A U.S. Department of Education *What Works Clearinghouse* review (2016) identified studies of *Read 180* that met these requirements alongside a host of other technical demands. Three experimental studies met group-design standards without reservations. One study entailed a within-school randomization of 881 sixth- through ninth-grade students in five Milwaukee public schools. The authors reported a statistically significant, modest effect of *Read 180* on general literacy skills ($ES = 0.14$; Swanlund, et al., 2012).¹ The two remaining studies were not particularly salient to our current work, as they involved implementations within afterschool programs (Fitzgerald & Hartry, 2008; Kim et al., 2010).

The *WWC* review also identified three experimental or quasi-experimental studies that met group-design standards with reservations that identified significant, positive effects. The largest of these, which involved within-school comparisons of roughly 700 middle school students in Boston, Houston and Dallas, reported a relatively substantial effect on reading comprehension ($ES = 0.4$; Interactive, 2002). Another cluster RCT involving five Massachusetts high schools and 456 ninth graders identified a somewhat smaller effect on general literacy skills ($ES = 0.18$; Sprague et al., 2012). Lastly, researchers matched roughly 2,000 *Read 180* high school participants in Phoenix to 2,000 statistically comparable peers, reporting a modest impact on general literacy skills ($ES = 0.10$; White, 2006). As we describe below, our current study represents by far the largest causal

¹ We only include estimates, samples and outcomes that WWC deemed appropriate for inclusion in their own (often aggregate) estimates of program impacts.

study of *Read 180* to date, involving over 62,000 students and 160 schools in one of the Nation's most diverse school districts.

***Read 180* Implementation in Miami-Dade**

During the 2022-23 academic year, students in Miami-Dade County Public Schools with below-proficient scores on the Spring 2022 Florida ELA assessment were to participate in *Read 180*. The intervention allows for multiple implementation models. Miami-Dade schools that offered traditional class periods and schedules adopted a single-period implementation in which a full *Read 180* rotational model occurred over two days. On the first day, students started with whole-class instruction, then completed one station. On the second day, students rotated between the remaining two stations, without the whole-class introductory lesson. Both days ended with a 5-10 minute whole-class wrap-up. In schools with block schedules (double periods), students completed the full rotational model each day they were scheduled for *Read 180* (typically 2-3 days per week). These sessions began with 15-20 minutes of whole-class instruction, followed by three 20-minute rotations. Classes then ended with a 5-10 period whole-class wrap up.

Data and Methods

Miami-Dade County Public Schools and HMH provided student-level standardized assessment scores for both the 2021-22 and 2022-23 academic years, as well as data on student academic and demographic backgrounds, and measures of *Read 180* usage for students who participated in the program. Our initial sample included 81,356 students in seventh through tenth grades in schools implementing *Read 180*. We excluded students missing either Spring 2022 or Spring 2023 Florida ELA assessment scores, which produced a sample of 65,523 students. We then eliminated students who received both *Read 180* and *System 44*, HMH's Tier 3 literacy intervention. This produced a final analytic sample of 62,858 students. This is the sample used with our main causal analyses of *Read 180*. We fortunately had no missing data on student gender, race/ethnicity, free/reduced-price lunch, and special education and English as Second or Other Language (ESOL) status. The sample has an even gender distribution, and just under 60% of students were eligible for free/reduced-price lunch. Roughly 10% received special education and, separately, ESOL services. In terms of student race/ethnicity, almost 73% of the sample was identified as Hispanic, 19.2 percent Black, 6.2 percent white, 1.1 percent Asian/Pacific Islander, with the remaining 0.6% of students identified as Native American or multi-racial.

For our correlational analyses of the links between *Read 180* usage and student literacy learning, we further restricted the sample to *Read 180* students who had completed at least one *Read 180* segment ($n=15,165$). With the usage models that leveraged *Reading Inventory (RI)* Lexile scores as outcomes, we additionally restricted the sample to students with both the *RI* outcome (post-test) and baseline (pre-test) scores ($n=12,968$).

Measures

Outcomes

Our first set of analyses, which explore the link between *Read 180* usage rates and student literacy development among *Read 180* participants, leverage two sets of standardized assessment outcomes. The first includes student-level scores on the Spring 2023 Florida state ELA assessment. The second set of scores are from the *Reading Inventory (RI)*, a computer-adaptive assessment designed to measure how well students read literature and expository texts of varying difficulties. It focuses on a broad set of skills, including: identifying details in a passage; identifying cause-and-effect relationships and sequence of events; drawing conclusions; and making comparisons and generalizations. Performance on the *Reading Inventory* is reported as a Lexile (L) score. The higher a student's score, the more challenging material that student is likely to be able to read and understand. As we describe in more detail below, these usage models also incorporate same-assessment baseline scores from the Spring 2022 administrations.

Our second set of analyses, which involve a series of regression discontinuity (RD) models, rely exclusively on the Spring 2023 Florida ELA assessment scores as outcomes. As noted above, Spring 2022 scores served as the tool to decide which students would participate in *Read 180*. We provide more information on the substantive and technical aspects of the implementation and RD analyses as they relate to our use of the Spring 2022 state assessment scores below. With both the Florida ELA assessments and *RI* Lexile outcomes, we standardized (z-scored) scores within grades at each administration period.

Covariates

Both our usage and RD models include controls for student grade, sex (females=1, males=0), and race/ethnicity (a series of dummy variables indicating whether the student identified as Asian, Black, white, or other race/ethnicity with Hispanic students serving as the comparison group in the multivariate analyses). The models also account for free/reduced-lunch eligibility (yes=1, no=0), and include separate indicators of whether the student received special education or ESOL services (yes=1, no=0).

Analytic Approach

Read 180 Usage

We first explored associations between *Read 180* usage and literacy development among students who participated in *Read 180*. Our aim was to identify the extent to which additional *Read 180* segment completion was associated with increased literacy learning. The analytic models we constructed can be defined as:

$$(1) Y_{is} = B_0 + B_1 \text{Medium}_i + B_2 \text{High}_i + \gamma' X_i + \eta + \epsilon_i$$

in which Y_{is} represents the Spring 2023 literacy outcome for student i in school s . *Medium* and *High* are dummy variables representing 3-5 *Read 180* segments completed and 6+ segments completed, respectively. The low-usage category (only 1-2 segments completed) is omitted and serves as the uncoded comparison group. Recall that *Read 180* students who did not complete one segment were excluded from the usage analyses. X_i is then a vector of student-level background covariates, including baseline (Spring 2022) literacy assessment scores, race/ethnicity, English language learner and special education status, eligibility for free/reduced-price lunch, and grade level. School fixed effects—which allow us to compare outcomes among students attending the same school—are indicated by η , while ϵ_i indicates the student-level error term. We conduct this dosage analysis first using ELA state test scores as the outcome and then run the identical models using the *RI* Lexile scores as the outcome. For each outcome, we examine all *Read 180* participants collectively and then each grade level separately.

To improve the interpretability of these results, we calculate the average Spring 2023 literacy score for the typical low-, medium-, and high-use *Read 180* participant. In Miami-Dade, the typical *Read 180* participant is Hispanic, qualified for free- and reduced-price lunch (FRL), did not receive special education or language services, and scored nearly two-thirds of a standard deviation below the mean within their same-grade Miami-Dade cohort. For this “typical” student, we calculate the average spring literacy score under low-, medium-, and high-use conditions. Assuming a normal distribution, we convert this average outcome from standard deviations to percentile points.

Regression Discontinuity Models

Although the dosage analyses provide important insights into *Read 180*’s efficacy, it cannot establish a causal link between *Read 180* and student literacy outcomes, given that *Read 180* students who completed more sessions differed in many ways from *Read 180* students who completed fewer sessions. This suggests that the usage findings may conflate *Read 180* effects with effects stemming from the types of students who have higher participation rates, such as those with increased attendance rates, motivation levels, or enhanced instructional supports.

Fortunately, the implementation of *Read 180* in Miami-Dade lends itself well to a second strategy, a regression discontinuity design, which can provide causal evidence of *Read 180*'s impact on student literacy development. The regression discontinuity design leverages the highly non-linear relationship between baseline student test scores and probability of receiving the *Read 180* intervention. Miami-Dade County Public Schools assigned which students received *Read 180* in the 2022-23 school year using their Spring 2022 literacy scores: students who scored below the state-determined proficiency cut-score qualified for *Read 180* intervention, while those at or above the cut-score did not.

We organize students into four groups based on their *Read 180* eligibility and actual participation (see Table 1). The first two groups, labeled “compliers,” include: 1) students eligible for *Read 180* who received it, and; 2) students not eligible who did not receive it. These students represent over 85% of the sample, suggesting a relatively clean implementation that followed the assignment guidelines. In addition to these compliers, we also identified a small number of “non-compliers,” including: 3) students who were eligible for *Read 180* but did not receive it (14.5% of the sample), and; 4) a very small number of students who were not eligible but did participate (0.4% of the sample). As expected, students who received *Read 180* had academic and social backgrounds that differed considerably from their peers who did not participate. On average, and again by design, students who received *Read 180* had much weaker baseline literacy skills ($ES = -1.23$). Participants were also more likely to be male, eligible for free/reduced-price lunch, Black, and receive special education and ESOL services. Over 60% of eligible students who were denied participation were receiving ESOL services, suggesting that schools did not feel that these students were ready for *Read 180* given their nascent English language skills.

Table 1. *Read 180* Eligibility and Participation by Student Background

	Compliers		Non-Compliers		
	<i>Read 180</i> Correctly Received (<i>n</i> =19,591)	<i>Read 180</i> Correctly Denied (<i>n</i> =33,941)	<i>Read 180</i> Incorrectly Received (<i>n</i> =224)	<i>Read 180</i> Incorrectly Denied (<i>n</i> =9,102)	Total Sample (<i>n</i> =62,858)
% Female	45.4	53.0	42.4	46.1	49.6
% Free/Red. Lunch	71.4	52.8	68.8	50.9	58.4
% Special Education	20.5	5.0	17.0	9.5	10.5
% Eng. Other Lang.	3.6	0.1	3.1	60.8	10.0
% Asian/Pac. Is.	0.5	1.5	0.0	0.7	1.1
% Black	29.9	13.8	40.2	15.8	19.2
% Hispanic	66.0	75.3	54.0	79.7	72.9
% White	3.1	8.6	5.8	3.6	6.2
% Other Race/Eth.	0.4	0.8	0.0	0.2	0.6
Baseline ELA Score	-0.617	0.608	0.156	-0.971	0.000
SD	0.721	0.732	0.664	0.805	1.000

With our regression discontinuity analyses, we use the exogenous variation around the cut-score to identify a causal effect, defining our model as:

$$(2) Y_i = \lambda'X_i + \pi Below_i + f(Baseline_i) + \epsilon_i$$

Here, Y_i is the Spring 2023 state ELA test score.² $Below$ is a dummy variable indicating whether the student scored below the ELA assessment proficiency cut-score and $f(Baseline_i)$ is a non-parametric function of the Spring 2022 literacy score, which serves as our running variable. We use a data-driven approach to determine the appropriate bandwidth—that is, to determine the range of observations around the cut-score to use in the model. There is a tradeoff here between precision and bias: a wider range of observations increases the precision of the estimate but also increases the risk of bias (Murnane & Willett, 2011). We select an equal bandwidth on either side of the cut-score that minimizes the mean-square-error (MSE; Calonico et al., 2014). In the model above, π represents the causal effect of qualifying for *Read 180*, or the intent-to-treat (ITT) estimate, for students right at the cut-score at baseline.

However, we are interested not just in the effect of *qualifying* for *Read 180*, but also in the effect of *receiving* *Read 180*. We therefore employ a fuzzy regression discontinuity design, using qualifying for *Read 180* as an instrument for *Read 180* participation. We leverage a two-stage least squares approach, defining our first stage as:

$$(3) R180_i = \mu'X_i + \rho Below_i + f(Baseline_i) + \epsilon_i$$

in which $R180$ is a dummy variable indicating whether student i participated in *Read 180*. We conservatively define *Read 180* participation as any *Read 180* usage, as measured by HMMH-provided backend data. This conservative definition increases the plausibility that the exclusion restriction assumption is met—that is, the adjusted effect of scoring below versus above the cut-score can be explained entirely through *Read 180* participation. We then substitute our first stage estimates into the second stage, which we define as follows:

$$(4) Y_i = \nu'X_i + \sigma R180_i + f(Baseline_i) + \epsilon_i$$

Here, σ provides an unbiased estimate of the effect of receiving *Read 180*, or the treatment on the treated (ToT), for students who score right at the cutoff at baseline. We conduct these analyses both for the whole sample and separately for each grade level. We then compare these effects to the average gap at baseline between students above and below the cut-score in Miami-Dade. While this comparison provides useful context, it should be interpreted with caution given that regression discontinuity results apply only to students who score right at the cutoff, but the comparison uses all students in Miami-Dade.

Finally, we conduct a series of robustness checks. We examine the extent to which our results are consistent across various functional forms and bandwidths, testing linear and quadratic functional forms, as well as mean-square-error optimal bandwidths that are equal on either side of the cut-score and that vary on either side of the cut-score (Cattaneo et al., 2019; Murnane & Willett, 2011). We also conduct a falsification test, running our analyses at various locations *other than* the cut-score. Because these false cut scores are *not* the point at which

² Recall that only *Read 180* participants have *RI* Lexile scores, meaning they are unavailable for use as an outcome in the regression discontinuity models.

there is a highly non-linear relationship between baseline student test scores and probability of receiving *Read 180*, we would not expect to observe any significant effects (Murnane & Willet, 2011).

Results

We begin with our correlational analyses of the links between *Read 180* usage rates and student literacy development. We find that *Read 180* students with higher levels of program engagement gained more literacy skills (see Table 2). As indicated in the far-left column, students who completed six or more *Read 180* segments in the 2022-23 school year had a 0.16 standard deviation developmental advantage on the ELA state test over otherwise similar peers in the same school who completed only one or two *Read 180* segments ($p<.001$).

Table 2. *Read 180* Segment Completion and Student Learning (*State Test Scores*) in Miami-Dade County Public Schools (2022-23)

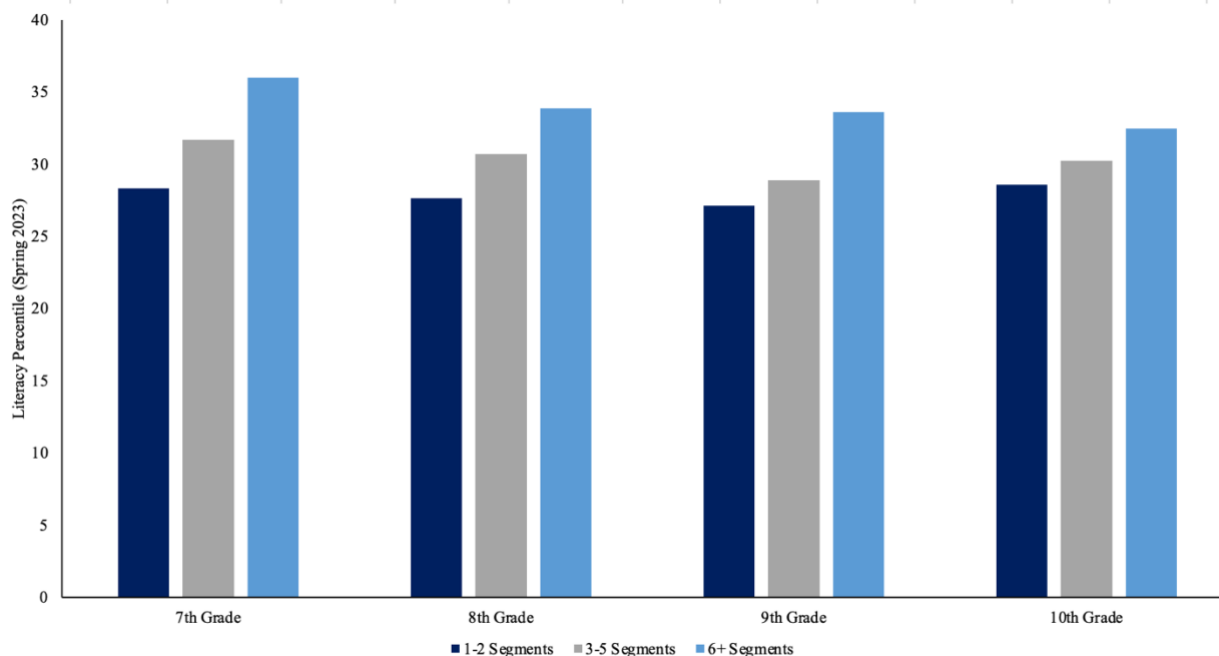
	(1) All Grades	(2) 7 th Grade	(3) 8 th Grade	(4) 9 th Grade	(5) 10 th Grade
3-5 Segments	0.066*** (0.012)	0.097*** (0.028)	0.089*** (0.024)	0.053* (0.023)	0.047 (0.025)
6+ Segments	0.161*** (0.014)	0.215*** (0.034)	0.178*** (0.030)	0.187*** (0.026)	0.110*** (0.029)
ELL	-0.056* (0.029)	-0.017 (0.057)	-0.033 (0.057)	-0.020 (0.055)	-0.176** (0.063)
FRL	-0.026* (0.011)	-0.027 (0.025)	0.008 (0.022)	-0.030 (0.020)	-0.054* (0.022)
IEP	-0.122*** (0.013)	-0.106*** (0.029)	-0.095*** (0.026)	-0.094*** (0.023)	-0.196*** (0.028)
Asian	-0.012 (0.067)	-0.119 (0.148)	0.022 (0.147)	-0.032 (0.114)	0.093 (0.135)
Black	-0.037* (0.015)	-0.055 (0.034)	-0.017 (0.029)	-0.035 (0.026)	-0.036 (0.031)
Other	-0.018 (0.078)	0.071 (0.148)	0.000 (0.133)	0.028 (0.166)	-0.237 (0.190)
White	-0.022 (0.028)	-0.098 (0.058)	0.084 (0.060)	0.027 (0.051)	-0.097 (0.059)
Seventh Grade	0.130 (0.074)				
Eighth Grade	0.103 (0.074)				
Ninth Grade	-0.006 (0.014)				
Baseline Score	0.659*** (0.010)	0.619*** (0.023)	0.650*** (0.020)	0.684*** (0.018)	0.668*** (0.020)
Constant	-0.173*** (0.037)	-0.139*** (0.032)	-0.174*** (0.028)	-0.130*** (0.028)	-0.072* (0.028)
Observations	15,165	3,141	3,645	4,608	3,771
R-squared	0.283	0.240	0.279	0.312	0.295
Schools	160	112	111	59	56

* $p<.05$; ** $p<.01$; *** $p<.001$. Usage categories compared to students completing 1 or 2 segments. All racial/ethnic categories compared to Hispanic students. Grades compared to tenth grade. Outcome is standardized (z-scored) within grade. All models include school fixed effects.

Students who completed three to five segments also developed somewhat more literacy skills compared to students who completed only one or two segments ($ES = 0.066$; $p < .001$). Importantly, these estimates are adjusted for a host of student academic and sociodemographic characteristics associated with both usage rates and literacy development.

This pattern is relatively constant across individual grade levels, with high-use students outperforming their low-use peers. The high-use advantage is strongest for younger students, with the high-low gap roughly twice as large in seventh grade as compared to tenth grade. In practical terms, this means that the typical high-use seventh grader ends the year just over the 35th percentile in Miami-Dade, while their low use peers—who started the year at the same baseline level—remain under the 30th percentile (see Figure 1). Although the gap is not as extreme in older grades, even in tenth grade the typical high-use student ends the year at roughly the 32nd percentile, while their low-use peers remain under the 30th percentile.

Figure 1. *Read 180* Segment Completion and Literacy Skills for the Typical *Read 180* Student in Miami-Dade County Public Schools (2022-23)



We check the robustness of these dosage results using a second outcome, *RI* Lexile scores, which span multiple grade level standards, and therefore are less likely to suffer from floor and ceiling effects. We find nearly identical results using the *RI* Lexile scores (see Table 3). Overall, students who completed six or more *Read 180* segments gained more skills than their peers who completed only one or two segments ($ES = 0.159$; $p < .001$), as did students who completed between three and five segments ($ES = 0.084$; $p < .001$). Across all grades, students who completed six or more segments gained between 0.144 and 0.214 standard deviations more than students who completed only one or two segments, and students completing three to five segments gained just under one-tenth of a standard deviation more compared to students completing only one or two segments. Again, the advantage is more pronounced in lower grades.

Although these usage models account for measured differences between students with varied levels of segment completion, and estimates are from students attending the same school, unmeasured differences between groups may remain for which we have not accounted. We therefore turn to our second approach, a regression discontinuity design, which allows us to isolate *Read 180*'s causal effect on student literacy development.

Table 3. *Read 180* Segment Completion and Student Learning (*State Test Scores*) in Miami-Dade County Public Schools (2022-23)

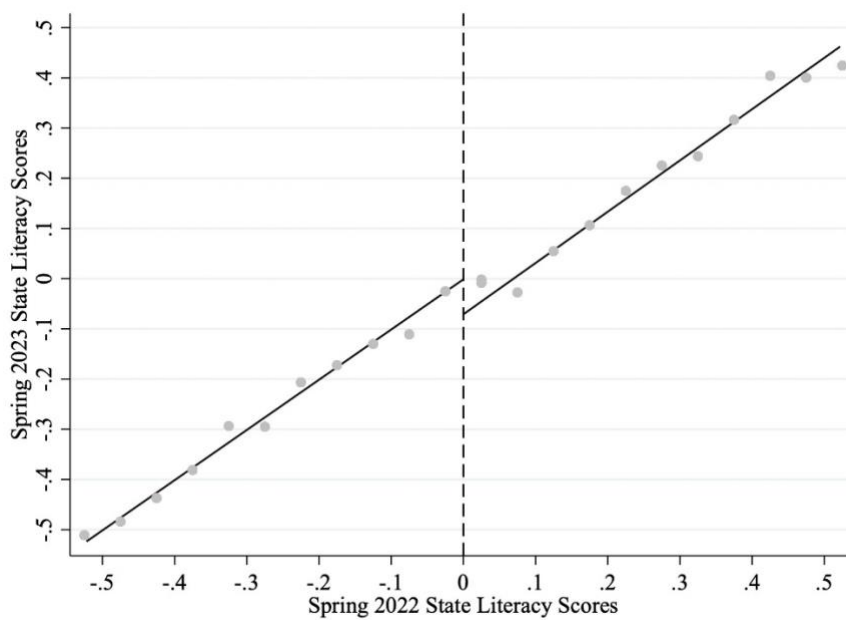
	(1) All Grades	(2) 7 th Grade	(3) 8 th Grade	(4) 9 th Grade	(5) 10 th Grade
3-5 Segments	0.084*** (0.013)	0.082** (0.029)	0.087*** (0.026)	0.079*** (0.024)	0.076** (0.024)
6+ Segments	0.159*** (0.014)	0.214*** (0.036)	0.168*** (0.032)	0.144*** (0.026)	0.157*** (0.028)
ELL	-0.196*** (0.030)	-0.174** (0.057)	-0.133* (0.062)	-0.149* (0.065)	-0.316*** (0.063)
FRL	-0.031** (0.011)	-0.048 (0.026)	-0.049* (0.024)	0.003 (0.020)	-0.045* (0.022)
IEP	-0.145*** (0.013)	-0.128*** (0.029)	-0.132*** (0.027)	-0.137*** (0.023)	-0.182*** (0.027)
Asian	-0.013 (0.066)	0.075 (0.150)	-0.224 (0.154)	-0.074 (0.113)	0.101 (0.124)
Black	-0.021 (0.015)	0.020 (0.036)	-0.033 (0.032)	-0.014 (0.026)	-0.043 (0.030)
Other	0.053 (0.079)	0.227 (0.145)	0.012 (0.146)	0.013 (0.170)	-0.168 (0.179)
White	0.028 (0.029)	0.018 (0.061)	-0.017 (0.065)	0.032 (0.052)	0.060 (0.059)
Seventh Grade	-0.039 (0.076)				
Eighth Grade	-0.007 (0.076)				
Ninth Grade	-0.067*** (0.014)				
Baseline Score	0.751*** (0.005)	0.726*** (0.013)	0.757*** (0.012)	0.748*** (0.010)	0.761*** (0.010)
Constant	0.053 (0.037)	-0.044 (0.031)	0.032 (0.028)	-0.008 (0.025)	0.094*** (0.026)
Observations	12,968	2,642	3,064	4,001	3,261
R-squared	0.630	0.567	0.595	0.633	0.684
Schools	156	106	101	55	54

* $p < .05$; ** $p < .01$; *** $p < .001$. Usage categories compared to students completing 1 or 2 segments. All racial/ethnic categories compared to Hispanic students. Grades compared to tenth grade. Outcome is standardized (z-scored) within grade. All models include school fixed effects.

Regression Discontinuity Analysis

We begin this section with a simple visual of the intent-to-treat (ITT) model (see Figure 2). We find, as expected, a linear positive relationship between Spring 2022 and Spring 2023 ELA test scores—students with low scores on the baseline assessment typically had low scores on the follow-up assessment, while students with high baseline scores typically had high follow-up scores. But the important finding here is located in the middle of the figure. The dotted vertical line represents the cut-score below which students were to receive *Read 180*. Note the *discontinuity* or “jump” in the relationship between baseline and follow-up test scores, with students just below the cut-score (who were eligible for *Read 180*) outperforming students just above the cut-score (who were not eligible).

Figure 2. Literacy Outcomes Relative to Proficiency/Assignment Cut-score for Students in Miami-Dade County Public Schools (2022-23)



As indicated in the top panel of Table 4, which quantifies this “jump,” we find that, on average, qualifying to participate in *Read 180* slightly improves literacy skills ($ES = 0.039$; $p < .01$). Recall that this treatment estimate is for all students who were eligible for participation, regardless of whether (or the extent to which) they actually did so. As we described above, not all students who were supposed to receive *Read 180* actually did, and those who did participate did so to varying degrees—hence the term, “*intent to treat*.”

We examined these ITT effects by grade level as well. The results for seventh, eighth, and ninth grades are consistent with the overall pattern. Across these three grades, we find small but significant effects, with the ITT estimate ranging from 0.065 to 0.074 *SDs*. Interestingly, though, we find no effect for tenth-grade students. Among the several potential explanations for this, one quite plausible hypothesis is that students who continue to receive a supplemental reading intervention in tenth grade may differ in unmeasured ways from intervention students in lower grades and may require additional supports beyond literacy. For example, our data did not include information on student attendance, which may have been less consistent among tenth graders.

Table 4. *Read 180* Effects on Student Literacy Learning in Miami-Dade County Public Schools (2022-23)

	(1) All Grades	(3) 7th Grade	(5) 8th Grade	(7) 9th Grade	(9) 10 th Grade
RD-ITT					
R180 Eligibility	0.039** (0.015)	0.070* (0.028)	0.065* (0.028)	0.074** (0.027)	-0.027 (0.029)
RD-First Stage					
R180 Eligibility	0.824*** (0.003)	0.858*** (0.007)	0.806*** (0.007)	0.821*** (0.006)	0.815*** (0.006)
F-Statistic	62,711.56	16,036.90	14,760.28	16,162.67	15,969.75
RD-TOT					
R180 Use	0.093** (0.031)	0.133* (0.055)	0.167* (0.074)	0.143* (0.057)	-0.051 (0.054)
Observations	62,858	13,945	15,670	16,522	16,721

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

Next, we turn to our treatment-on-the-treated (ToT) analyses, which recognize the differing participation rates among *Read 180* students. A reliable ToT estimate requires a strong first stage, meaning that status immediately above versus below the cut-score is highly predictive of receiving *Read 180*. Fortunately, we find a roughly 80 percentage point difference in likelihood of receiving *Read 180* services just at the cut-score, and F-statistics over 10,000 across all five models (see Table 4), well exceeding the standard for sufficient predictive power (Stock et al., 2002). These first stage estimates give us confidence to proceed with estimating the ToT, which suggests a significant effect of *Read 180* for students at the cut-score who participated in the program ($ES = 0.093$; $p < .01$). We find roughly similar ToT results for seventh through ninth grades, with effects ranging from 0.133 to 0.167 SDs. But again, we find no effects of *Read 180* among tenth graders.

We contextualize these effects from the regression discontinuity analysis, comparing them to the typical grade-level gap between qualifying and non-qualifying student (see Figure 3). In seventh through ninth grade, *Read 180* participation closed roughly 10% of the average baseline gap, meaningfully supporting students who start the year below grade level. However, it is important to interpret these estimates cautiously. Our estimates apply only to students right at the cutoff who ultimately participated in *Read 180*. The same effect would not necessarily hold for students who either started the year with lower baseline scores or, despite qualifying for *Read 180*, did not participate.

We conclude with a series of robustness checks. For ease of comparison, we reproduce our main ITT findings in the first set of rows of Table 5. In each subsequent set of rows, we examine our results under different specifications. In the second specification, we rely on a quadratic rather than linear functional form. Our results are directionally similar, with the overall effect and seventh- through ninth-grade estimates remaining positive. However, only the 9th grade effect is significant ($p < .05$). The third specification allows the bandwidth to vary above and below the cut-score. In this case, our findings exactly parallel the original model, with small but statistically significant positive effects overall and for seventh through ninth grades. The fourth specification combines the previous two, both using a quadratic functional form and allowing the bandwidth to vary above and below the cut-score. Again, our results are directionally consistent, with positive estimates overall and for seventh, eighth, and ninth grades. Similar to the prior quadratic specification though, not all four of these

estimates are significant: only the overall and ninth-grade estimates are significant in this case ($p < .05$). Given that our results are directionally consistent across all four models and local linear regression both generally tends to be the best fit (Cattaneo et al., 2019) and appears to most closely match our data (see Figure 2), these results provide moderate support for our initial findings.

Figure 3. Proportion of Average Literacy Gap Closed by *Read 180* Participation

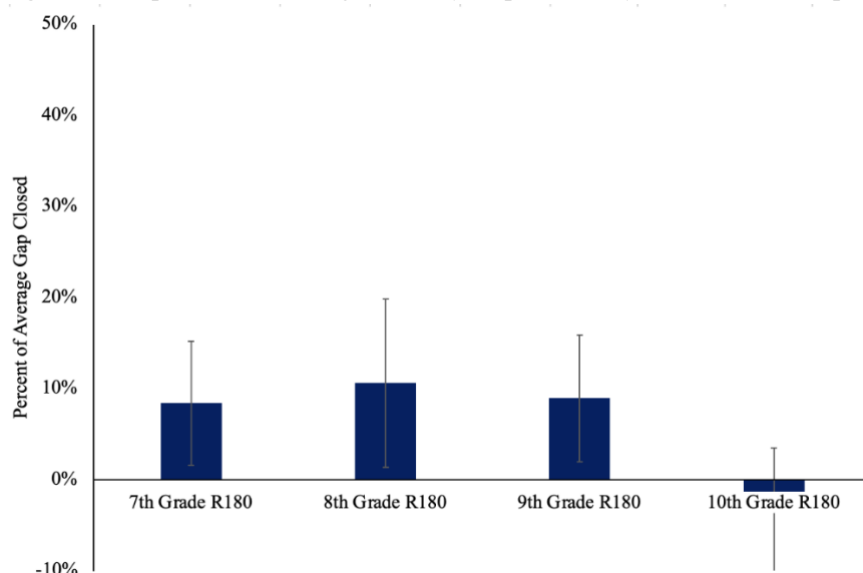


Table 5. Robustness Test of *Read 180* Effects on Student Literacy Learning in Miami-Dade County Public Schools (2022-23)

	(1) All Grades	(3) 7th Grade	(5) 8th Grade	(7) 9th Grade	(9) 10 th Grade
Linear & MSERD					
R180 Eligibility	0.039** (0.015)	0.070* (0.028)	0.065* (0.028)	0.074** (0.027)	-0.027 (0.029)
Quadratic & MSERD					
R180 Eligibility	0.027 (0.020)	0.056 (0.038)	0.055 (0.038)	0.077* (0.031)	-0.027 (0.033)
Linear & MSETWO					
R180 Use	0.034* (0.016)	0.067* (0.028)	0.065* (0.029)	0.068* (0.028)	-0.036 (0.030)
Quadratic & MSETWO					
R180 Use	0.037* (0.018)	0.041 (0.040)	0.060 (0.036)	0.075* (0.033)	-0.034 (0.037)
Observations	62,858	13,945	15,670	16,522	16,721

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

Our final robustness check tests for an “effect” at false cut-scores. Because there is no jump in *Read 180* usage at these false cut-scores, any significant effects would suggest bias in the model. We perform the same ITT analyses described above but exchange the actual cut-score for false cut-scores at one-half and one-quarter standard deviations above and below the real cut-score. Across all five ITT models, we find null effects at each false cut-score (Table 6), bolstering our confidence in the results.

Table 6. Falsification Test of *Read 180* Effects on Student Literacy Learning in Miami-Dade County Public Schools (2022-23)

	(1) RD-ITT All Grades	(2) RD-ITT 7th Grade	(3) RD-ITT 8th Grade	(4) RD-ITT 9th Grade	(5) RD-ITT 10 th Grade
False Cut at -0.5 SD	-0.005 (0.021)	-0.062 (0.037)	-0.041 (0.031)	0.021 (0.043)	0.015 (0.035)
False Cut at -0.25 SD	-0.030 (0.017)	-0.035 (0.035)	-0.029 (0.034)	-0.038 (0.027)	-0.022 (0.032)
False Cut at 0.25 SD	-0.023 (0.014)	-0.045 (0.030)	-0.035 (0.025)	-0.006 (0.027)	-0.022 (0.028)
False Cut at 0.5 SD	0.017 (0.015)	0.031 (0.027)	0.032 (0.029)	-0.005 (0.026)	0.010 (0.023)
Observations	62,858	13,945	15,670	16,522	16,721

No falsification effects are significant at the $p < .05$ level.

Conclusion

In this paper we shared results from a large-scale implementation of *Read 180* in Miami-Dade County Public Schools during the 2022-23 academic year. We found that descriptively, *Read 180* students who completed more segments also developed stronger literacy skills. This was true even after accounting for a host of student academic and sociodemographic characteristics associated with both program usage and student learning. Although both important and suggestive, those results remain descriptive. Fortunately, Miami-Dade implemented *Read 180* in a way that supported a far more rigorous, causal technique—the regression discontinuity approach. Using this technique, we compared students just above and just below the assignment cut-score. The assumption here is that the very small test score difference separating these students is random noise, meaning that differences in outcomes between groups at the end of the academic year can be attributed to *Read 180*. And indeed, we identified statistically significant positive effects of *Read 180* assignment on literacy development among seventh through ninth graders.

Implementations and corresponding studies that leverage regression discontinuity techniques are associated with both advantages and limitations. First, assuming an instructional intervention has a clear assignment mechanism, all students who require the intervention can receive it and will not have to wait for the conclusion of the study as some would in a randomized controlled trial. Second, schools can implement the intervention exactly as they would in the absence of the study; students and teachers are not asked to alter instructional approaches or classroom activities in service of the study beyond the adoption of *Read 180*. The third benefit flows from the first two, in that schools and districts are more likely to participate in the study when they know that students who need it will receive the program and that no other instructional or staffing changes will be required. This is particularly important in the current climate, given the increasing difficulty researchers are facing in identifying schools and districts willing to participate in large-scale evaluation efforts.

But the regression discontinuity approach—and thus our current study—is also accompanied by a set of limitations. Foremost is perhaps the fact that RD estimates are valid only for students situated around the assignment cut-score. An important question in the context of our current study is the extent to which *Read 180* effects vary depending on students' initial literacy skills. A second limitation is that the RD technique cannot isolate effects at various dosage levels. For instance, in this study, we capture effects of *Read 180* assignment and *Read 180* participation. We cannot, however, identify effects for students who participated at high- (6+ segment) or medium (3-5 segment) levels. Our dosage analyses provide suggestive evidence that higher usage might lead to more growth, but confirming these findings would require a multi-arm study, with sites randomly assigned to various dosage levels. Despite these limitations, the findings are promising, pointing to *Read 180*'s efficacy in boosting literacy achievement among striving readers.

References

- Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6), 2295-2326.
- Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.
- Fahle, E. M., Kane, T. J., Patterson, T., Reardon, S. F., Staiger, D. O., & Stuart, E. A. (2023). *School district and community factors associated with learning loss during the COVID- 19 pandemic*. Center for Education Policy Research at Harvard University: Cambridge, MA.
- Fitzgerald, R., & Hartry, A. (2008). *What works in afterschool programs: The impact of a reading intervention on student achievement in the Brockton Public Schools (phase II)*. Berkeley, CA: MPR Associates, Inc. and the National Partnership for Quality Afterschool Learning at SEDL.
- Guryan, J., & Ludwig, J. (2023). *Overcoming pandemic-induced learning loss*. Aspen Institute.
- HMH (2024). *Research evidence base: HMH Read 180*. HMH.
- Interactive, Inc. (2002). *An efficacy study of READ 180: A print and electronic adaptive intervention program, grades 4 and above*. Ashland, VA: Interactive, Inc.
- Kim, J. S., Samson, J. F., Fitzgerald, R., & Hartry, A. (2010). A randomized experiment of a mixed-methods literacy intervention for struggling readers in grades 4–6: Effects on word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. *Reading and Writing: An Interdisciplinary Journal*, 23(1), 1109–1129.
- Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. Oxford University Press.
- Peters, S. J., Langi, M., Kuhfeld, M., & Lewis, K. (2023). *Unequal learning loss: How the COVID-19 pandemic influenced the academic growth of learners at the tails of the achievement distribution*. (EdWorking Paper: 24-1031). Annenberg Institute at Brown University.
- Sprague, K., Zaller, C., Kite, A., & Hussar, K. (2012). *Springfield-Chicopee School Districts Striving Readers program final report Years 1–5: Evaluation of implementation and impact*. Providence, RI: The Education Alliance at Brown University.
- Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.
- Swanlund, A., Dahlke, K., Tucker, N., Kleidon, B., Kregor, J., Davidson-Gibbs, D., & Halberg, K. (2012). *Striving Readers: Impact study and project evaluation report: Wisconsin Department of Public Instruction (with Milwaukee Public Schools)*. Naperville, IL: American Institutes for Research.
- U.S. Department of Education (2016). *WWC intervention report: Read 180*. Institute of Education Sciences, What Works Clearinghouse.

U.S. Department of Education (2025). *NAEP report card. Reading*. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).
https://www.nationsreportcard.gov/reports/reading/2024/g4_8/?grade=4

White, R., Haslam, B. M., & Hewes, G. (2006). *Improving student literacy in the Phoenix Union High School District 2003–04 and 2004–05*. Washington, DC: Policy Studies Associates.