



# Beyond the Classroom: Impact of a High-Dosage Tutoring Program on Student Literacy Achievement

Ayman Shakeel

Abt Global

This study examines the impact of a high-dosage tutoring program, characterized by low tutor-to-student-ratio, on the literacy achievement of students in grades two through five in a mid-sized suburban school district in the southeastern United States. Using a student-level randomized controlled trial, 333 students were randomly assigned to either receive tutoring during the intervention period of the school day or continue with business as usual. While the average treatment effect on literacy achievement is not statistically significant across the full sample, analysis reveals a significant positive impact of 0.2 standard deviations for students with lower baseline achievement, those most in need of high-dosage tutoring, accompanied by a negative impact for the higher achieving students. These findings underscore the importance of tailoring high-dosage tutoring interventions to student needs and suggest that differentiated implementation may be important to maximizing benefits of such programs across diverse learner profiles.

VERSION: November 2025

Suggested citation: Shakeel, Ayman. (2025). Beyond the Classroom: Impact of a High-Dosage Tutoring Program on Student Literacy Achievement. (EdWorkingPaper: 25-1348). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/e6nm-k121>

# **Beyond the Classroom: Impact of a High-Dosage Tutoring Program on Student Literacy Achievement<sup>1</sup>**

Ayman Shakeel, Abt Global

This study examines the impact of a high-dosage tutoring program, characterized by low tutor-to-student-ratio, on the literacy achievement of students in grades two through five in a mid-sized suburban school district in the southeastern United States. Using a student-level randomized controlled trial, 333 students were randomly assigned to either receive tutoring during the intervention period of the school day or continue with business as usual. While the average treatment effect on literacy achievement is not statistically significant across the full sample, analysis reveals a significant positive impact of 0.2 standard deviations for students with lower baseline achievement, those most in need of high-dosage tutoring, accompanied by a negative impact for the higher achieving students. These findings underscore the importance of tailoring high-dosage tutoring interventions to student needs and suggest that differentiated implementation may be important to maximizing benefits of such programs across diverse learner profiles.

---

<sup>1</sup> This study was funded by Accelerate. I acknowledge the support by Matthew Steinberg, Allan Porowski, and Beth Boulay. I also thank participants at the Society of Research for Educational Effectiveness (2025) for their feedback. I am especially grateful to the tutoring provider and district partner for their collaboration.

# Introduction

Improving literacy among adolescents in the United States remains a critical educational priority, as national assessments continue to reveal downward trends in reading achievement. According to the National Assessment of Educational Progress (NAEP), commonly referred to as the Nation's Report Card, reading scores for both fourth and eighth graders declined in 2024, continuing a downward trajectory that began prior to the Covid-19 pandemic (National Center for Education Statistics, 2024). These declines are particularly pronounced among lower-achieving students, exacerbating existing educational disparities and raising concerns about long-term consequences for workforce readiness. As researchers and policymakers seek solutions, there is growing consensus that targeted and evidence-based interventions, such as high-dosage tutoring, are essential to correcting these trends and ensuring all students develop the skills necessary for academic and lifelong success.

High-dosage tutoring, often defined as tutoring delivered at least three times per week in one-on-one or small-group settings (Robinson et al., 2021; Washington DC Office of the State Superintendent of Education, 2021; Dobbie and Fryer, 2013), has long been a focal point for educators and researchers as a promising strategy to enhance student learning outcomes. Recent studies underscore the effectiveness of tutoring programs not only in boosting academic performance but also in reducing student absenteeism (Fryer & Howard-Noveck, 2020; Lee et al., 2024; Nickow, Oreopoulos, & Quan, 2024). A recent meta-analysis reports a substantial pooled effect size of 0.37 standard deviations, indicating a strong positive impact of high-dosage tutoring on student achievement (Nickow, Oreopoulos, & Quan, 2024).

Despite the robust evidence supporting efficacy of high-dosage tutoring, programs vary widely in their design and implementation, and there is no clear consensus on what constitutes an

ideal model. Program effectiveness depends on several factors, including tutoring frequency, tutor qualifications, group size, delivery setting (in-school vs. out-of-school), and curriculum. Moreover, research highlights persistent challenges in the implementation of these programs, underscoring the critical importance of monitoring and tracking implementation (Carbonari et al., 2024).

This study analyzes the impact of a high-dosage and low tutor-to-student-ratio (1:3) program on the literacy achievement of students in grades two to five in a midsized suburban school district in the southeastern United States. I conducted a randomized controlled trial (RCT) in which 144 students in the treatment group were randomly assigned to receive tutoring during the intervention period of the school day, while 189 students in the control group continued with business-as-usual and did not receive structured support during that time. Tutoring was introduced over a three-week rollout period in November 2024 and continued through the end of the academic year. This study answers the following research questions:

- 1) Implementation Question: Was the tutoring intervention implemented with fidelity?
- 2) Impact Question: What is the impact of tutoring on literacy benchmark assessments for second to fifth graders compared to business-as-usual?
- 3) Exploratory Question: Does the impact of tutoring on literacy benchmark assessments differ by subsamples?

The analysis reveals that on average, there is no statistically significant difference between literacy achievement of the treatment and control groups. However, meaningful heterogeneity in treatment effects emerges when disaggregating results by prior achievement. Specifically, students with prior achievement below the grade-level median experience a statistically significant improvement of 0.20 standard deviations ( $p < 0.05$ ), while students with prior achievement above

the grade-level median exhibit a negative impact of 0.15 standard deviations ( $p < 0.10$ ). These findings align with prior research, including Kraft (2015) who demonstrated that individualized after-school tutorials benefit lower-performing students most and Robinson et al. (2024) who found that virtual tutoring was particularly effective for students who initially scored below benchmark levels.

The tutoring provider primarily serves students identified as Tier 2 or Tier 3, that is those considered at risk for academic failure and in need of additional instructional support.<sup>2</sup> The findings in this paper suggest that interventions designed for a specific subgroup of students may yield unintended consequences when applied to students outside the intended target group. The differential effects by prior achievement also underscore the importance of tailoring tutoring interventions to student academic needs and starting points. Understanding and accounting for such variation is essential for designing programs that maximize learning gains across the achievement spectrum.

## High-Dosage Tutoring Background

High-dosage tutoring has emerged as one of the most effective educational interventions for addressing learning gaps, particularly in the wake of the COVID-19 pandemic. After the pandemic, several districts and schools around the US implemented tutoring given the robust body of empirical evidence demonstrating its capacity to significantly improve student achievement across grade-levels and subject areas. A meta-analysis by Nickow, Oreopoulos, and Quan (2024) summarized findings from almost 100 experimental studies and reported an average effect size of 0.37 standard deviations across reading and mathematics outcomes, with particularly strong

---

<sup>2</sup> These tiers are part of the Response to Intervention (RTI) framework. The RTI framework is a multi-tiered approach used in schools to identify and support students with learning and behavior needs. It's designed to provide early, systematic assistance to children who are struggling, before they fall too far behind.

impacts observed in early-grade reading and upper-grade mathematics interventions. These findings have been reinforced by more recent studies, some of which specifically focus on literacy outcomes.

Kortecamp and Peters (2024) conducted a randomized controlled trial to examine the effects of one-on-one high-dosage tutoring on the reading achievement of kindergarten and first-grade students. The study found that students who received tutoring two to three times per week significantly outperformed their peers on nationally normed literacy assessments. Similarly, Cortes et al. (2025) evaluated a scalable early literacy tutoring program embedded within kindergarten classrooms. The study included 818 students in a large southeastern U.S. district and found that those who received high-dosage tutoring were more than twice as likely to reach grade-level reading benchmarks by the end of the school year compared to their peers.

The effectiveness of high-dosage tutoring relies heavily on the program design features. Programs that deliver tutoring during the school day tend to outperform those offered after school or during the summer, often owing to better attendance and alignment with classroom instruction (Robinson, Kraft, and Loeb, 2025). One-on-one tutoring is also found to be the most impactful (Nietzel et al., 2022), however, small-group sizes, comprising three or four students, can also be effective when implemented with fidelity and help reduce cost. Research also shows that qualifications and training of tutors are an important determinant of the success of tutoring. While certified teachers and paraprofessionals tend to generate the greatest impact, college students and other non-traditional tutors can be effective when provided with structured curricula and ongoing coaching (White, Groom-Thomas, & Loeb, 2023).

Despite the promise of high-dosage tutoring programs, challenges remain in scaling these programs sustainably (Kraft, Edwards, & Canatta, 2024). Widespread adoption is often limited by

factors like the cost of one-on-one tutoring, need for a reliable pipeline of trained tutors, and logistical barriers such as scheduling and space constraints.

Importantly, high-dosage tutoring may not work equally well for students across the achievement spectrum. Evaluating the impact of high-dosage tutoring on different student subgroups is essential for understanding the overall effectiveness. Students enter tutoring programs with varying levels of prior achievement, language proficiency, socioeconomic background, and learning needs, which often influence how they respond to additional instructional support. Without disaggregating effects by subgroup, average treatment effects may obscure important variation, masking substantial gains for some students and unintended consequences for others.

## Program Details

This section outlines the sample selection process for the tutoring intervention and provides a detailed description of both the treatment and control conditions.

### Sample Selection

This study is set in a mid-sized suburban school district in the southeastern United States. Students in grades two through five from the two schools in the district were randomly assigned to receive the tutoring intervention or continue with business-as-usual during the intervention period of the school day. In October 2024, the school district provided a list of students eligible for random assignment. This list excluded students in Tier 3 (those who require the most intensive level of academic or behavioral support) or those required to receive other mandatory services during the intervention period, as stipulated by the state (such as special education services). It also excluded students participating in other interventions who did not have sufficient time during any intervention period to receive treatment. Eligible students were stratified by school, grade, and

intervention period, and then randomly assigned to either the treatment or control group within those strata.

According to state law where this study was conducted, students were required to participate in some tutoring for a full school year or attend summer camp learning if they did not score proficient on the state English Language Arts (ELA) test. The study sample includes some fourth- and fifth-grade students who were required to receive state-mandated tutoring. These students met for two sessions per week, each lasting one hour. Half of each session was spent in face-to-face tutoring, and the other half completing i-Ready lessons online. Within this subgroup, students assigned to the treatment group received both the high dosage tutoring intervention and after-school tutoring. Those assigned to the control group received only after-school tutoring.

It is important to note that the tutoring intervention studied here is primarily designed for Tier 2 or Tier 3 students, however, there was a delay in program implementation and many students in the target group were already participating in other interventions by the time tutoring began. As a result, the sample also included higher-achieving students who were not part of the original target population.

## Treatment Condition

This subsection provides contextual background on the structure and delivery of the tutoring program, including details on dosage, instructional content, and tutor preparation.

### Dosage

Tutoring was delivered in a high-dosage format and maintained a low tutor-to-student ratio of 1:3, consistent with best practices for high-impact tutoring (Nickow, Oreopoulos, & Quan, 2024). According to the tutoring provider, the program dosage, defined as the *total amount of tutoring a student would receive under ideal conditions* (Accelerate, 2025), was 24 weeks of tutoring for 3

sessions per week with 45 minutes per session. However, as there was a delay in establishing a partnership and contract between the tutoring provider and the district, scheduled dosage, defined as the *amount of tutoring an individual student would receive if they attended all scheduled tutoring sessions* (Accelerate, 2025) was approximately 18 weeks of tutoring for 3 sessions per week with 45 minutes per session. The district began implementing the program in early November 2024, with a phased rollout over three weeks due to tutor availability. Tutoring continued through mid-May 2025, extending nearly to the end of the academic year.

According to the tutoring provider, an attendance rate of at least 80% of the program dosage constitutes adequate dosage, defined as the *threshold above which providers believe a student would realize substantively meaningful improvements in learning* (Accelerate, 2025). As tutoring started later than the provider anticipated, none of the students received enough sessions to meet this threshold.

### *Tutoring Structure and Instructional Content:*

At the start of the program, each tutor was assigned an average of three students, with the expectation that these groupings would remain consistent throughout the academic year.<sup>3</sup> In instances where a tutor was absent and another tutor had an available slot (for example, due to a student absence), students were temporarily reassigned for that session. If no alternate tutor was available, the session was canceled for the student.

Tutoring sessions were held in the cafeteria for students in grades two, three, and five, while grade four students received tutoring in various classrooms. In both settings, students did not share the space with peers outside of the intervention group. The shared tutoring environment enabled tutors to observe one another's sessions, fostering a unique accountability structure. This

---

<sup>3</sup> Some tutors were assigned two students but none of them had more than three students.

arrangement allowed tutors to exchange instructional strategies, collaboratively address common challenges, and make real-time adjustments to maintain consistent lesson pacing.

Tutors used grade-level appropriate i-Ready workbooks aligned with classroom instruction and pacing. These materials corresponded with the curriculum, designated by the state Department of Education as a High-Quality Instructional Material for K–5 English Language Arts. Additionally, the program incorporated state-approved resources to provide explicit phonics practice during each session. Tutoring lessons were designed to mirror classroom content, offering targeted reinforcement of key skills.

At the start of the program, tutors administered a baseline cold reading assessment, followed by biweekly assessments at the conclusion of each i-Ready lesson. These assessments served as informal progress checks, helping tutors identify areas where students required additional support (for example, identifying point of view or main character).

Tutors were also responsible for tracking student attendance. The tutoring provider employed a Tutor Support Manager who audited the attendance records prior to entry in the official tracker.

### *Tutor Qualifications:*

Nineteen tutors were hired to support this program, all employed on a part-time basis with the option to work either 9 or 18 hours per week. The tutor cohort included college students (not enrolled in Educator Preparation Programs) and community members, though none served in a volunteer capacity. Minimum hiring criteria required a high school diploma, with preference given to candidates who had some college experience and prior work with school-age children. Tutors were required to pass the state education department foundational skill test prior to starting.

Compensation was determined based on educational background and prior experience with the program, with hourly wages ranging from \$20 to \$30.

### *Tutor Trainings:*

The tutoring provider ensured that the tutors were well-prepared to deliver high-dosage tutoring. Initial trainings included sessions on human resources and employee handbook, reporter protocols, Foundational Skills/i-Ready curriculum specific instruction, and Connected Literacy/Sounds First material developed by the state education department. Tutors also participated in a range of asynchronous virtual and in-person trainings held over school breaks. In addition, the Tutor Support Manager facilitated monthly cohort trainings, with topics tailored to observed instructional needs. Appendix Exhibit A1 lists some of the sessions. The Tutor Support Manager also regularly conducted both formal and informal observations of tutoring sessions, providing real-time feedback and individualized coaching to tutors. In addition, the manager offered professional development opportunities to the full tutor cohort.

### **Business-as-Usual Condition:**

Although neither the tutoring provider nor the district formally tracked the activities in the business-as-usual group, district staff reported that these students were expected to engage with computer-assisted instruction using the i-Ready online platform during the intervention period of the school day. Teachers facilitated this time and occasionally provided small-group support tailored to individual student needs. Notably, these students were not located in the same physical space as those participating in the tutoring intervention.

## **Data**

The school district and the tutoring provider provided the relevant data for this study. The student-level test scores and demographic data were provided by the school district and the

tutoring provider tracked and provided tutoring attendance data for the students who received tutoring.

The analysis primarily focuses on i-Ready reading diagnostic scores as the tutoring utilized i-Ready instructional materials and the test score data were available for all grades. The i-Ready Reading Diagnostic Assessment is a computer-adaptive test designed to evaluate K-12 students' reading abilities across key domains such as phonological awareness, phonics, vocabulary, comprehension of literature, and comprehension of informational text. Prior research has demonstrated a strong alignment between i-Ready scores and standardized state assessments. For example, in Ohio, the correlation between i-Ready and state test scores was reported as 0.80 for English Language Arts. (Curriculum Associates, 2020). During the intervention period, the district administered the i-Ready diagnostic assessment at the beginning and end of the academic year. The beginning-of-year i-Ready diagnostic was administered between August 12<sup>th</sup> and August 30<sup>th</sup>, 2024, and the end-of-year i-Ready diagnostic between April 21<sup>st</sup> and May 9<sup>th</sup>, 2025.

I also obtained access to end-of-year state ELA test scores and Common Assessments to Support Education (CASE) ELA benchmark assessment scores for students in grades three through five, which I use as supplemental outcome measures. CASE assessments are closely aligned with state-specific academic standards in core subjects like Math, ELA, Science, and Social Studies. They are designed to provide actionable insights into student learning and predict performance on end-of-year state assessments. The end-of-year CASE assessment was administered during the second half of March 2025, while the state ELA assessment was conducted between April 14<sup>th</sup> and April 18<sup>th</sup>, 2025. Additionally, demographic data were available, including information on students' gender, race/ethnicity, economic disadvantage status, and an indicator of whether a student was identified as requiring tutoring under the state law.

# Research Design and Empirical Model

This section describes the study design and the empirical model used to estimate the impact of the tutoring intervention.

## Research Design

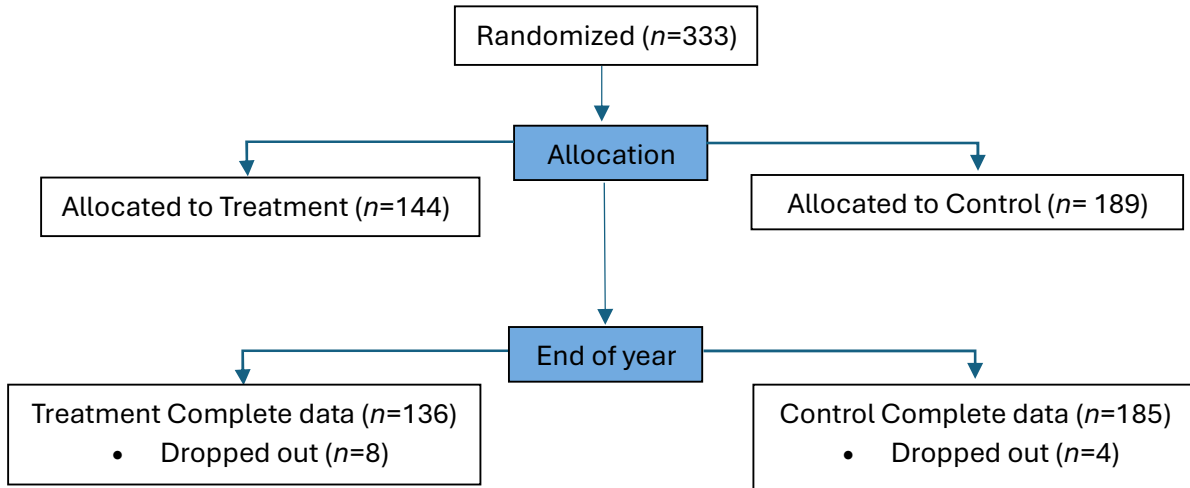
This study employs a student-level randomized controlled trial to evaluate the impact of the tutoring intervention on end-of-year literacy achievement. The CONSORT diagram presented in Figure 1 summarizes the random assignment process and resulting sample. The initial sample included 333 students, with 144 randomly assigned to the treatment group and 189 to the control group. The Minimal Detectable Effect Size (MDES) for the study was 0.21.<sup>4</sup> Randomization was conducted within blocks, hereafter referred to as *randomization blocks*, defined by school, grade, and intervention period. Prior to randomization, the tutoring provider specified the number of students they could accommodate in each intervention period. Based on the capacity of the tutoring provider, students were randomly assigned to the available tutoring slots within each randomization block, while the remaining students in those randomization blocks were assigned to the control group. Appendix Exhibit A3 presents the number of randomized students within each randomization block.

After randomization, some students exited the district over the course of the year. As a result, the analytic sample includes 321 students with complete i-Ready data: 136 in the treatment group and 185 in the control group (Figure 1). During the intervention period, 21 students in the final analytic sample who were initially assigned to receive tutoring experienced scheduling changes and could no longer be tutored. Most of these students were either pulled into other programs, such as math interventions or gifted education services.

---

<sup>4</sup> The details for the MDES calculation are in Appendix Exhibit A2.

Figure 1: Consolidated Standards of Reporting Trials (CONSORT) diagram



## Empirical Model

I standardized the test scores by grade to have a mean of zero and standard deviation of one so that the scores are comparable across different grades. The data were pooled across grades, and an intent-to-treat (ITT) model specified in Equation (1) was estimated to identify the impact of the tutoring intervention on student literacy outcomes.:

$$Y_i = \beta_0 + \beta_1 Treat_i + \beta_2 X_i + \beta_3 Base_i + \gamma_{pgs} + \epsilon_i \quad (1)$$

Where:

- $Y_i$  is the outcome (i-Ready/state ELA test /CASE) for student  $i$ .
- $Treat_i$  is an indicator that equals one if the student is assigned to the tutoring intervention, and 0 otherwise.
- $X_i$  is a vector of student covariates (economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring)
- $Base_i$  is the baseline achievement (Fall i-Ready)
- $\gamma_{pgs}$  are school-grade-intervention-period (Randomization Block) fixed effects
- $\epsilon_i$  is the error term

In the model,  $\beta_1$  captured the effect of Tutor 901 on the literacy outcome. It represents the average difference in achievement scores between the treatment and control groups, conditional on baseline i-Ready achievement, student demographics, and randomization block fixed effects. The key identifying assumption of the model is that, conditional on random assignment, students in the treatment and control groups are equivalent in both observable and unobservable characteristics. Thus, any differences in outcomes can be attributed to the treatment. While it is not possible to test for balance on all potential confounders, I present mean differences in baseline covariates between the treatment and control groups in the Results section. To assess robustness, I also estimate alternative specifications that exclude baseline achievement and student demographic controls.

## Results

This section presents the impact of the tutoring intervention on literacy achievement. I start by presenting a baseline covariate balance test and then outline the implementation and impact and analyses.

### Covariate Balance

Table 1 presents the covariate balance analysis. The first three columns of the table present baseline analysis for the randomized sample and the last three columns present baseline analysis for the sample of students with follow-up i-Ready data. Because randomization was conducted within school, grade, and intervention period, all covariates are residualized by the randomization blocks. As expected, the first three columns of Table 1 show that the two groups are well balanced on baseline characteristics and none of the differences are statistically different from zero.

Table 1. Descriptive Statistics and Baseline Equivalence

Variable	Randomized Sample			Final Analytic Sample		
	(1) Control	(2) Treatment	(3) Model Adjusted Difference	(4) Control	(5) Treatment	(6) Model Adjusted Difference
Baseline i-Ready Achievement	0.435 (0.657)	0.438 (0.594)	0.021 (0.066)	0.449 (0.656)	0.456 (0.595)	0.024 (0.068)
Male	0.397 (0.491)	0.479 (0.501)	0.056 (0.054)	0.400 (0.491)	0.485 (0.502)	0.060 (0.055)
Black or African American	0.434 (0.497)	0.451 (0.499)	0.004 (0.055)	0.432 (0.497)	0.441 (0.498)	-0.005 (0.056)
White	0.460 (0.500)	0.431 (0.497)	-0.010 (0.055)	0.459 (0.500)	0.434 (0.497)	-0.006 (0.056)
Asian	0.037 (0.189)	0.014 (0.117)	-0.015 (0.018)	0.038 (0.191)	0.015 (0.121)	-0.014 (0.018)
Economically Disadvantaged	0.296 (0.458)	0.382 (0.488)	0.072 (0.052)	0.297 (0.458)	0.382 (0.488)	0.075 (0.053)
Require State Mandated Tutoring	0.101 (0.302)	0.111 (0.315)	0.005 (0.032)	0.097 (0.297)	0.103 (0.305)	0.001 (0.032)
Observations	189	144	333	185	136	321

Note: The “Randomized Sample” comprises all students who are randomly assigned to the treatment of control groups, while the “Final Analytic Sample” comprises students with non-missing fall *and* spring scores. Columns (1), (2), (4), and (5) report the means and standard deviations (in parentheses) for each variable and columns (3) and (6) present the residualized difference between the control and treatment groups for each variable along with the standard error (in parentheses) for the difference. The residuals are obtained by regressing the variables on the randomization block fixed effect. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Because the analytic sample excludes some students from the original randomized sample, it is important to assess baseline balance between treatment and control group students who contribute to the final analysis. If missing data is nonrandom, it may bias the estimated effects, limiting the ability to attribute observed differences to the tutoring intervention. However, if the missing data is random and not systematically related to treatment assignment, I would expect baseline characteristics to remain balanced between the treatment and control groups in the final analytic sample. The last three columns of Table 1 present mean comparisons of baseline covariates between treatment and control students included in the final analysis ( $F=0.85$ ,  $P = 0.57$ ).

These results suggest that there are no statistically significant differences across groups, supporting the assumption of baseline equivalence and missingness at random.<sup>5</sup>

## Implementation Analysis

### *Research Question 1: Was the tutoring intervention implemented with fidelity?*

Table 2 presents implementation statistics for treatment group students in the final analytic sample. The average tutoring group size was 2.9 students, which aligns with initial expectations. Due to the staggered rollout of tutoring over a three-week period, the total duration varied between 14 and 19 weeks, corresponding to approximately 43 to 56 sessions. On average, students were offered 49 sessions and attended 43, resulting in an average attendance rate of 89%, calculated as the percentage of offered sessions attended.

Table 2. Implementation Statistics for the Treatment Group Analytic Sample

	Mean
Group Size	2.87
Sessions Offered	49.0
Sessions Attended	43.8
Fraction of Offered Sessions Attended	0.89
Observations	120

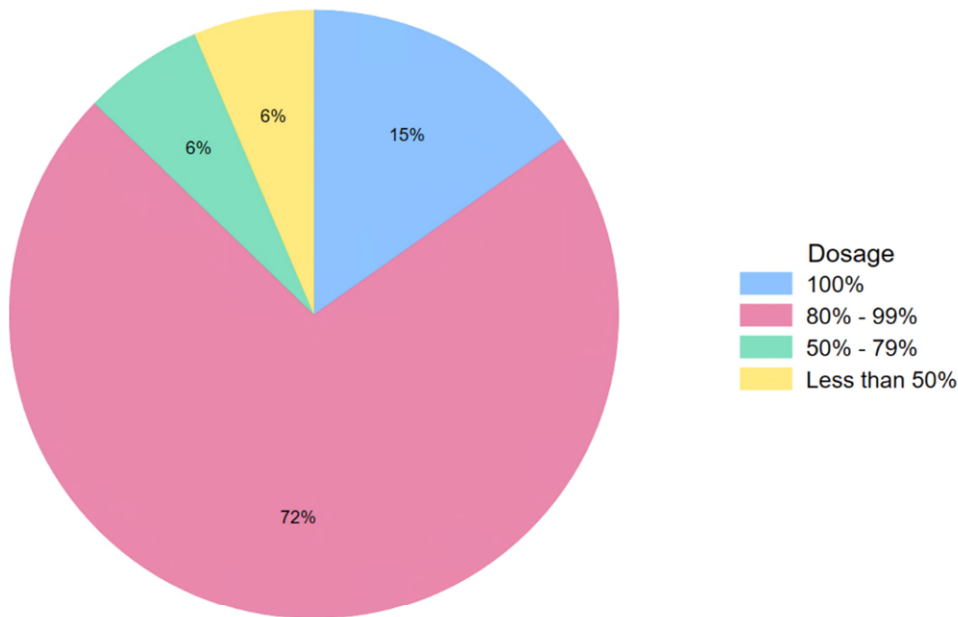
Figure 2 displays the detailed distribution of student attendance during the intervention period, measured as the percentage of offered tutoring sessions that students attended. Fifteen percent of students attended all sessions, while 72 percent attended between 80% and 90% of the offered sessions. Six percent attended between 50% and 79%, and another six percent attended

---

<sup>5</sup> Appendix Table A4 presents the balance statistics for the state ELA and CASE score analytic samples. Overall, covariates appear balanced between the treatment and control groups, except for some imbalance in the proportion of male students. The overall F-statistics are 1.35 ( $p = 0.21$ ) for the state ELA sample and 1.22 ( $p = 0.28$ ) for the CASE sample.

less than 50%. Overall, approximately 88% of students had an attendance of at least 80%. As mentioned earlier, tutoring started later than the provider had initially anticipated, so none of the students met the threshold for adequate dosage.

Figure 2: Percentage of Offered Sessions Attended by Students



Notes: Dosage is calculated as the percentage of sessions attended divided by the percentage of sessions offered

## Impact Analysis:

Research Question 2: What is the impact of tutoring on literacy benchmark assessments for second to fifth graders compared to business-as-usual?

Table 3 presents end-of-year results for the i-Ready impact analysis. Column 1 reports the estimated effect of the tutoring intervention on i-Ready test scores, controlling only for randomization block fixed effects. Column 2 adds a control for baseline achievement, and Column 3 includes the full set of covariates as specified in Equation (1). Across all specifications, the estimates reveal that there is no statistically significant difference between the i-Ready scores of the treatment and control groups, with the point estimates being close to zero.

Table 3. Effect of the Tutoring Intervention on i-Ready Achievement

Outcome	(1) Standardized Score	(2) Standardized Score	(3) Standardized Score
i-Ready	0.005 (0.073)	-0.006 (0.062)	0.0002 (0.060)
Randomization Block FE	Yes	Yes	Yes
Baseline Achievement	No	Yes	Yes
Demographic Characteristics	No	No	Yes
Observations	321	321	321

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

To assess the consistency of the estimated impacts, I also examine the effect of the tutoring intervention on state standardized ELA test scores and CASE assessment scores for students in grades three through five. As shown in Appendix Table A5, these estimates are also close to zero and statistically insignificant, suggesting that the intervention did not produce measurable effects on these additional outcomes.

## Exploratory Impact Analysis:

Research Question 3: Does the impact of tutoring on literacy benchmark assessments differ by subsamples?

Table 4 explores the heterogeneity of treatment effects across a wide variety of subsamples, including grades, gender, race, economic disadvantage status, and prior achievement.<sup>6</sup> Splitting the sample by grades, gender, and race yields results that are either insignificant or inconsistent across test scores (refer to Appendix Table A7 to see estimates for

<sup>6</sup> Appendix Table A6 tests for the robustness of these estimates to the exclusion of baseline achievement and demographic controls.

state ELA and CASE scores). This suggests that the impact of tutoring does not vary meaningfully across these dimensions. There is some evidence that the impact of tutoring is negative for economically disadvantaged students.

The most robust partition of the data is by student prior achievement. I stratify students into two groups, high-achieving and low-achieving, and estimate the model separately for each subgroup. High-achieving students are defined as those with baseline i-Ready achievement above the grade-level sample median, while low-achieving students fall below the grade-level sample median. The average baseline achievement of the low-achieving student is 0.17 standard deviations (49<sup>th</sup> percentile) and the average for high-achieving students is 0.80 standard deviations (80<sup>th</sup> percentile).<sup>7</sup>

Table 4 reveals that the tutoring intervention has a strong positive impact on low-achieving students relative to high-achieving students. On average, low achieving students in the treatment group score 0.20 standard deviations ( $p < 0.05$ ) higher than the low achieving students in the control group. In contrast, tutoring appears to have had a marginally significant negative effect on high-achieving students, with a reduction of 0.15 standard deviations. The difference in impact between high- and low-achieving students is 0.35 standard deviations and is statistically significant at the 1% level. Appendix Table A6 shows that these findings are robust to the exclusion of baseline achievement and demographic controls. Furthermore, Appendix Table A7 confirms that the results are consistent across both state ELA and CASE scores.

---

<sup>7</sup> Defining high- and low-achieving students using nationally normed i-Ready percentile cutoffs, such as the 50th percentile which distinguishes whether a student is performing above or below grade-level, may be more intuitive from a policy standpoint. However, the sample we studied includes relatively few students below the 50th percentile. Appendix Table A8 categorizes students based on whether they score above or below this threshold, and the findings align with the results presented in Table 4.

Table 4. Effect of the Tutoring Intervention on i-Ready Achievement, By Subsamples

	(1)	(2)	(3)
<b>Impact by Gender</b>	Males	Females	Difference
	-0.001	-0.103	0.103
	(0.093)	(0.082)	(0.083)
<i>Observations</i>	140	181	321
<b>Impact by Race</b>	Black	Non-Black	Difference
	0.129	-0.102	0.231*
	(0.089)	(0.787)	(0.118)
<i>Observations</i>	142	179	321
<b>Impact by Economic Disadvantage Status</b>	Economically Disadvantaged	Not Economically Disadvantaged	Difference
	-0.181**	0.080	-0.261**
	(0.083)	(0.086)	(0.118)
<i>Observations</i>	214	107	321
<b>Impact by Grades</b>	Grades 2-3	Grades 4-5	Difference
	-0.077	0.098	0.175
	(0.074)	(0.096)	(0.122)
<i>Observations</i>	142	179	321
<b>Impact by Prior Achievement</b>	Achievement Below Median	Achievement Above Median	Difference
	0.198**	-0.153*	0.351***
	(0.086)	(0.083)	(0.119)
<i>Observations</i>	157	164	321
Randomization Block FE	Yes	Yes	Yes
Baseline Achievement	Yes	Yes	Yes
Demographic Characteristics	Yes	Yes	Yes

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Appendix Table A9 further partitions the sample into quartiles based on prior achievement. The results reveal a consistent pattern: students in the lowest quartile benefit the most from tutoring, while those in the highest quartile experience the greatest negative impact. However, caution is warranted in interpreting these results, as the sample sizes become increasingly smaller.

## Discussion

This paper evaluates the impact of a high-dosage tutoring program, on the literacy achievement of students in grades two to five. The analysis shows that students in the treatment group did not have significantly higher literacy achievement than the control group. One likely explanation for the lack of overall impact is that students did not receive the adequate tutoring dosage intended by the provider, due to a delay in program implementation. However, the overall effect masks heterogeneity of effects by prior achievement. I find a strong and statistically significant impact of 0.2 standard deviations for students with baseline achievement below grade-level median accompanied by a marginally significant negative impact of 0.15 standard deviations for students with baseline achievement above grade-level median. There is also some suggestive evidence that the impact of tutoring is higher for students that are not economically disadvantaged. This may indicate that tutoring alone is insufficient to support learning for students experiencing poverty, though this remains a hypothesis.

The concentration of positive effects among lower-achieving students may be attributable to the design of the tutoring model itself. The tutoring program employed a universal approach, delivering instruction based on a fixed curriculum that was tailored to address foundational skill gaps, characteristics more aligned with the needs of lower-achieving students. This contrasts with tutoring models that personalize instruction to target individual learning needs. The tutoring provider typically serves students identified as Tier 2 and Tier 3. However, because the intervention

began later in the academic year, many students in these tiers were already engaged in other support programs and were therefore not eligible to participate. While the study included several students who scored on the higher end of the baseline i-Ready assessment distribution, the tutoring provider reported that they did not completely modify instructional materials or strategies for these students, aside from grouping students based on initial proficiency levels. The tutoring provider also explicitly noted that serving students with higher proficiency levels presented a challenge, and they were not entirely surprised by the differential findings based on prior achievement.

The results in this paper underscore the importance of examining treatment effects across subgroups, as average effects can obscure meaningful heterogeneity. From a policy perspective, these findings highlight the need to consider differentiated tutoring strategies that are responsive to the diverse learning profiles of students. Targeted interventions may be essential to ensure that tutoring programs yield beneficial outcomes across the achievement spectrum. The findings also suggest that tutoring programs should be strategically targeted rather than universally applied. Given scarcity of resources, policies should prioritize tutoring resources for those most in need as mandatory or blanket tutoring requirements could be counterproductive.

## References

- Accelerate. (2025, August 28). *Defining tutoring dosage for program implementation and applied research*. Retrieved from <https://accelerate.us/research/defining-tutoring-dosage-for-program-implementation-and-applied-research/>
- Dobbie, W., & Fryer Jr, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4), 28-60.
- Fryer, R.G. (2017). The production of human capital in developed countries: Evidence from 196 randomized field experiments. In *Handbook of Economic Field Experiments*, vol. 2, 95-322. North-Holland. [https://scholar.harvard.edu/files/fryer/files/handbook\\_fryer\\_03.25.2016.pdf](https://scholar.harvard.edu/files/fryer/files/handbook_fryer_03.25.2016.pdf)
- Carbonari, M. V., Davison, M., DeArmond, M., Dewey, D., Dizon-Ross, E., Goldhaber, D., Hashim, A. K., Kane, T. J., McEachin, A., Muroga, A., Morton, E., Patterson, T., & Staiger, D. O. *The Impact and Implementation of Academic Interventions During COVID: Evidence from the Road to Recovery Project*. CALDER Working Paper No. 275-0624-2.
- Cortes, K. E., Kortecamp, K., Loeb, S., & Robinson, C. D. (2025). A scalable approach to high-impact tutoring for young readers. *Learning and Instruction*, 95, 102021.
- Curriculum Associates. (2020). *i-Ready diagnostic: Linking study with Ohio's state tests (OST)*. <https://www.curriculumassociates.com/-/media/>
- Fryer, R. G., & Howard-Noveck, M. (2020). High-Dosage Tutoring and Reading Achievement: Evidence from New York City. *Journal of Labor Economics*, 38(2), 421-452.
- Karen Kortecamp & Michelle Lynn Peters. (2024). The Impact of a High- Dosage Tutoring Program on Reading Achievement of Beginning Readers: A Multi-Level Analysis, *Journal of Education for Students Placed at Risk (JESPAR)*, 29:3, 291-309, DOI: 10.1080/10824669.2023.2179056

Kraft, M. A., Sanderson Edwards, D., & Cannata, M. (2024). *The scaling dynamics and causal effects of a district-operated tutoring program* (EdWorkingPaper No. 24-1030). Annenberg Institute at Brown University. <https://doi.org/10.26300/zcw7-4547>

Kraft, M. A., Schueler, B. E., & Falken, G. (2024). *What Impacts Should We Expect from Tutoring at Scale? Exploring Meta-Analytic Generalizability*. (EdWorkingPaper: 24-1031). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/zygj-m525>

Lee, M. G., Loeb, S., & Robinson, C. D. (2024). *Effects of High-Impact Tutoring on Student Attendance: Evidence from the OSSE HIT Initiative in the District of Columbia*. (EdWorkingPaper: 24 -1107). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/wghb-4864>

National Center for Education Statistics. (2024). *Summary of results: NAEP 2024 reading assessment for grades 4 and 8*. U.S. Department of Education. Retrieved from [https://www.nationsreportcard.gov/reports/mathematics/2024/g4\\_8/supporting-files/summary-of-results.pdf](https://www.nationsreportcard.gov/reports/mathematics/2024/g4_8/supporting-files/summary-of-results.pdf)

Neitzel, A. J., Lake, C., Pellegrini, M., & Slavin, R. E. (2022). A Synthesis of Quantitative Research on Programs for Struggling Readers in Elementary Schools. *Reading Research Quarterly*, 57(1), 149–179

Nickow, A., Oreopoulos, P., & Quan, V. (2024). The Promise of Tutoring for PreK–12 Learning: A Systematic Review and Meta-Analysis of the Experimental Evidence. *American Educational Research Journal*, 61(1), 74-107. <https://doi.org/10.3102/00028312231208687>

Robinson, C. D., Kraft, M. A., Loeb, S., & Schueler, B. E. (2021). Accelerating Student Learning with High-Dosage Tutoring. EdResearch for Recovery Design Principles Series. *EdResearch for Recovery Project*. <https://files.eric.ed.gov/fulltext/ED613847.pdf>

Robinson, C. D., Pollard, C., Novicoff, S., White, S., & Loeb, S. (2024). The Effects of Virtual Tutoring on Young Readers: Results from a Randomized Controlled Trial. *Educational Evaluation and Policy Analysis*. doi: 10.3102/0162373724128884

Washington DC Office of the State Superintendent of Education (2021). *High-dosage tutoring: A proven strategy to accelerate student learning. Guide for local education agencies*. Washington, DC: Author.  
[https://osse.dc.gov/sites/default/files/dc/sites/osse/page\\_content/attachments/HighDosageTutoringGuidance.pdf](https://osse.dc.gov/sites/default/files/dc/sites/osse/page_content/attachments/HighDosageTutoringGuidance.pdf)

# Appendix

## Exhibit A1: Topics for Monthly Tutor Training

Training Topics
<ul style="list-style-type: none"><li>• Aligning to instructional goals</li><li>• Utilizing wait time</li><li>• Tracking student Learning goals</li><li>• Cultivating a culture of learning and relationships</li><li>• Student management</li><li>• Deliberate practice of i-ready reading lesson plans</li><li>• Practicing wait-time in depth</li><li>• Giving precise directions</li><li>• Checks for understanding</li><li>• Lessening tutor talk time</li><li>• Increasing student independence</li></ul>

Exhibit A2: Minimal Detectable Effect Size (MDES) Calculation

Assumptions		Comments
Alpha Level ( $\alpha$ )	0.05	Probability of a Type I error
Two-tailed or One-tailed Test?	2	
Power ( $1-\beta$ )	0.80	Statistical power (1-probability of a Type II error)
P	0.43	Proportion of the sample randomized to treatment: 144/333
R <sup>2</sup>	0.55	Percent of variance in outcome explained by covariates. R <sup>2</sup> is assumed to be 0.55 as the covariates include blocks and pre-test, where the pre-test (i-Ready) is different from the post-test for the ELA TCAP outcome.
k*	10	based on blocks and covariates (school-grade-intervention-period blocks)
n (Total Sample Size)	333	total students
M (Multiplier)	2.81	Computed from T <sub>1</sub> and T <sub>2</sub>
T <sub>1</sub> (Precision)	1.97	Determined from alpha level, given two-tailed or one-tailed test
T <sub>2</sub> (Power)	0.84	Determined from given power level
MDES	<b>0.209</b>	Minimum Detectable Effect Size

Exhibit A3: Summary of the Number of Students in the Study by School, Grade, and Intervention Period

School	Grade	Intervention Period	Number of Students in the Treatment Group	Number of Students in the Control Group
School 1	2	1	30	39
School 1	3	1	30	48
School 2	4	1	18	51
School 2	4	2	30	10
School 2	5	1	18	11
School 2	5	2	18	30

Table A4. Descriptive Statistics and Baseline Equivalence

Variable	Randomized Sample			State ELA Analytic Sample			CASE Analytic Sample		
	(1) Control	(2) Treatment	(3) Model Adjusted Difference	(4) Control	(5) Treatment	(6) Model Adjusted Difference	(4) Control	(5) Treatment	(6) Model Adjusted Difference
Baseline i-Ready Achievement	0.435 (0.657)	0.438 (0.594)	0.021 (0.066)	0.354 (0.639)	0.375 (0.584)	0.046 (0.077)	0.356 (0.644)	0.360 (0.579)	0.042 (0.078)
Male	0.397 (0.491)	0.479 (0.501)	0.056 (0.054)	0.361 (0.482)	0.523 (0.502)	0.13** (0.06)	0.350 (0.479)	0.530 (0.502)	0.16** (0.06)
Black or African American	0.434 (0.497)	0.451 (0.499)	0.004 (0.055)	0.408 (0.493)	0.449 (0.500)	0.02 (0.06)	0.400 (0.492)	0.450 (0.500)	0.03 (0.06)
White	0.460 (0.500)	0.431 (0.497)	-0.010 (0.055)	0.463 (0.500)	0.393 (0.491)	-0.04 (0.06)	0.464 (0.501)	0.400 (0.492)	-0.04 (0.06)
Asian	0.037 (0.189)	0.014 (0.117)	-0.015 (0.018)	0.048 (0.214)	0.019 (0.136)	-0.02 (0.02)	0.050 (0.219)	0.020 (0.141)	-0.02 (0.02)
Economically Disadvantaged	0.296 (0.458)	0.382 (0.488)	0.072 (0.052)	0.286 (0.453)	0.411 (0.494)	0.11* (0.06)	0.271 (0.446)	0.410 (0.494)	0.13** (0.06)
Require State Mandated Tutoring	0.101 (0.302)	0.111 (0.315)	0.005 (0.032)	0.122 (0.329)	0.131 (0.339)	0.00 (0.04)	0.129 (0.336)	0.110 (0.314)	-0.02 (0.04)
Observations	189	144	333	147	107	254	140	100	240

Note: The "Randomized Sample" comprises all students who are randomly assigned to the treatment of control groups, while the "State ELA Analytic Sample" and "CASE Analytic sample" comprises students with non-missing fall i-Ready score *and* spring ELA or CASE scores, respectively. Columns (1), (2), (4), (5), (7), and (8) report the means and standard deviations (in parentheses) for each variable and columns (3), (6), and (9) present the residualized difference between the control and treatment groups for each variable along with the standard error (in parentheses) for the difference. The residuals are obtained by regressing the variables on the randomization block fixed effect. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Table A5. Effect of the Tutoring Intervention on State Standardized English Language Arts and CASE Assessment for Students in Grades 3-5

	State ELA Test Score			CASE Test Score		
	(1)	(2)	(3)	(1)	(2)	(3)
Treatment	0.051 (0.132)	0.001 (0.118)	0.039 (0.118)	0.026 (0.134)	-0.016 (0.119)	0.005 (0.119)
Randomization Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Achievement	No	Yes	Yes	No	Yes	Yes
Demographic Characteristics	No	No	Yes	No	No	Yes
Observations	254	254	254	249	249	249

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Table A6. Effect of the Tutoring Intervention on i-Ready Achievement, by Subsamples

	(1) Standardized Score	(2) Standardized Score	(3) Standardized Score
<b><i>Impact by Gender</i></b>			
Treatment	-0.123 (0.092)	-0.115 (0.080)	-0.103 (0.083)
Treatment*Male	0.188 (0.152)	0.154 (0.122)	0.102 (0.124)
<b><i>Impact by Race</i></b>			
Treatment	-0.057 (0.098)	-0.118 (0.079)	-0.102 (0.079)
Treatment*Black	0.172 (0.150)	0.276** (0.127)	0.231* (0.118)
<b><i>Impact by Economic Disadvantage Status</i></b>			
Treatment	0.067 (0.096)	0.061 (0.088)	0.080 (0.087)
Treatment*Economically Disadvantaged	-0.198 (0.145)	-0.257** (0.124)	-0.261** (0.119)
<b><i>Impact by Grades</i></b>			
Treatment	-0.119 (0.094)	-0.123 (0.077)	-0.077 (0.074)
Treatment*Grades 4 or 5	0.238 (0.144)	0.217* (0.121)	0.175 (0.122)
<b><i>Impact by Prior Achievement</i></b>			
Treatment	-0.255*** (0.085)	-0.199** (0.082)	-0.153* (0.083)
Treatment*Achievement Below Median	0.437*** (0.131)	0.405*** (0.122)	0.351*** (0.119)
Randomization Block FE	Yes	Yes	Yes

## Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Baseline Achievement	No	Yes	Yes
Demographic Characteristics	No	No	Yes
Observations	296	296	296

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. Estimates are obtained from a regression of the literacy achievement on an interaction of treatment and an indicator for subgroup of interest. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Table A7. Effect of the Tutoring Intervention on State ELA and CASE Scores, by Subsamples

	State ELA Test Score			CASE Test Score		
	(1)	(2)	(3)	(1)	(2)	(3)
<b>Impact by Gender</b>						
Treatment	0.060 (0.179)	0.062 (0.170)	0.081 (0.172)	-0.038 (0.172)	-0.031 (0.157)	0.092 (0.160)
Treatment*Male	-0.105 (0.291)	-0.235 (0.269)	-0.308 (0.269)	0.023 (0.303)	-0.182 (0.265)	-0.283 (0.284)
<b>Impact by Race</b>						
Treatment	0.105 (0.169)	0.042 (0.148)	0.090 (0.144)	-0.018 (0.180)	-0.065 (0.165)	-0.049 (0.167)
Treatment*Black	-0.085 (0.277)	-0.059 (0.249)	-0.109 (0.240)	0.134 (0.287)	0.136 (0.249)	0.116 (0.246)
<b>Impact by Economic Disadvantage Status</b>						
Treatment	0.195 (0.164)	0.147 (0.155)	0.193 (0.147)	0.172 (0.166)	0.119 (0.160)	0.144 (0.164)
Treatment*Economically Disadvantaged	-0.378 (0.271)	-0.475* (0.248)	-0.566** (0.256)	-0.258 (0.295)	-0.350 (0.255)	-0.373 (0.262)
<b>Impact by Grades</b>						
Treatment	0.193 (0.228)	0.112 (0.223)	0.119 (0.211)	0.159 (0.234)	0.080 (0.212)	0.140 (0.219)
Treatment*Grades 4 or 5	-0.209 (0.280)	-0.166 (0.265)	-0.124 (0.259)	-0.196 (0.285)	-0.150 (0.257)	-0.216 (0.266)
<b>Impact by Prior Achievement</b>						
Treatment	-0.248 (0.160)	-0.178 (0.148)	-0.026 (0.148)	-0.382** (0.172)	-0.351** (0.167)	-0.284 (0.176)
Treatment*Achievement Below Median	0.456* (0.252)	0.413* (0.228)	0.263 (0.231)	0.672*** (0.256)	0.656*** (0.242)	0.594** (0.243)
Randomization Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Achievement	No	Yes	Yes	No	Yes	Yes
Demographic Characteristics	No	No	Yes	No	No	Yes
Observations	254	254	254	249	249	249

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. Estimates are obtained from a regression of the literacy achievement on an interaction of treatment and an indicator for subgroup of interest. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

## Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Table A8. Effect of the Tutoring Intervention on i-Ready, by Prior Achievement Percentile

	(1) Below 50 <sup>th</sup> Percentile	(2) Above 50 <sup>th</sup> Percentile	(3) Difference
Treatment	0.285** (0.132)	-0.105 (0.066)	0.390*** (0.068)
Randomization Block FE	Yes	Yes	Yes
Baseline Achievement	Yes	Yes	Yes
Demographic Characteristics	Yes	Yes	Yes
Observations	79	242	321

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

## Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement

Table A9. Effect of the Tutoring Intervention on i-Ready, by Prior Achievement Quartile

	(1) Quartile 1	(2) Quartile 2	(3) Quartile 3	(4) Quartile 4
Tutoring	0.282** (0.118)	0.036 (0.117)	-0.012 (0.124)	-0.321** (0.138)
Randomization Block FE	Yes	Yes	Yes	Yes
Baseline Achievement	Yes	Yes	Yes	Yes
Demographic Characteristics	Yes	Yes	Yes	Yes
<i>Average Baseline Achievement</i>				
In Standard Deviations	-0.284	0.269	0.677	1.23
In Percentile	39	60	75	88
Observations	84	82	79	76

Notes: Baseline Achievement is standardized fall i-Ready score. Student demographic controls include gender, race/ethnicity, economic disadvantage status, and an indicator for whether a student requires state mandatory tutoring. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

# Impact of Tutor 901, a High-Dosage Tutoring Program, on Student Literacy Achievement