The Impacts of i-Ready Personalized Instruction on Student Math Achievement Among Striving Learners

Ashley A. Grant, PhD
Michael A. Cook, PhD
Steven M. Ross, PhD

August 2023

JOHNS HOPKINS
SCHOOL of EDUCATION

Center for Research and Reform in Education

Contents

EXECUTIVE SUMMARY:
The Impacts of i-Ready Personalized Instruction on Student Math Achievement Among Striving Learners

In July 2021, The Center for Research and Reform in Education (CRRE) at Johns Hopkins University partnered with Curriculum Associates (CA) to conduct an efficacy study of the effects of i-Ready Personalized Instruction on student achievement in four school districts in California. The present report focuses on the effects of i-Ready Personalized Instruction on the student population of "striving learners" identified as scoring below 30% proficient on the Diagnostic pretest. Among this group of students, the study examines findings from quantitative analyses comparing math achievement gains, as measured by the Smarter Balanced Assessment (SBA). The analyses use multilevel modeling and propensity score matching to compare math achievement gains between students who experienced both i-Ready Personalized Instruction and Diagnostic testing and students who only participated in i-Ready Diagnostic testing. Additional analyses examined relationships between i-Ready usage metrics and SBA scores. Outcome data from the 2021-22 school year were used in all analyses.

Research questions for this evaluation were:

1. What is the impact of i-Ready Personalized Instruction on Smarter Balanced Assessment math achievement for students identified as striving learners?

2. How do program effects on math achievement vary for student subgroups and by school characteristics?
    a. By student demographic characteristics, including grade level, gender, race/ethnicity, disability status, and English language learner status?
    b. By school characteristics, including Title I eligibility, total enrollment size, and the percentage of students of minority race.

3. How do program effects on math achievement vary by implementation as measured by student usage?

The study sample consisted of 11,034 "striving learner" students in Grades 3-6 from four school districts in California. All four districts used i-Ready Diagnostic assessments, but schools in one district assigned some students to receive the Personalized Instruction product in addition to the Diagnostic product. Striving learner students in this district assigned to both the Personalized Instruction and Diagnostic products were compared to striving learner students in the other districts who were only assigned to the Diagnostic product.

Key findings of the current study include:

i-Ready Personalized Instruction was associated with greater math achievement gains for striving learners. Assignment to i-Ready Personalized Instruction was associated with greater gains on the SBA math exam for striving learners. This impact was strongest for students in Grades 4 and 5 and students identified as Black.

Greater i-Ready Personalized Instruction usage was associated with math achievement gains. Students who used i-Ready Personalized Instruction for longer amounts of time and number of lessons had significantly higher math achievement on the SBA in relation to comparison students.

Conclusions

The key results and conclusions of this evaluation are as follows:

- Assignment to i-Ready Personalized Instruction was associated with significantly higher SBA math scores for striving learners.
- Students assigned to i-Ready Personalized Instruction in Grades 4 and 5 experienced greater growth in math achievement than students in other grades who were assigned to i-Ready Personalized Instruction. This was potentially driven by higher usage rates by Grades 4 and 5 students compared to the other grades.
- Students identified as Black assigned to i-Ready Personalized Instruction experienced greater growth in math achievement than students of other races who were assigned to i-Ready Personalized Instruction. Students of White, Hispanic, and Other race backgrounds also benefited from treatment more than students of Asian racial identity.
- i-Ready Personalized Instruction usage metrics including total time and lesson count variables were positively associated with significantly higher math achievement, in relation to comparison students.

The Impacts of i-Ready Personalized Instruction on Student Math Achievement Among Striving Learners

In July 2021, The Center for Research and Reform in Education (CRRE) at Johns Hopkins University partnered with Curriculum Associates (CA) to conduct an efficacy study of the effects of i-Ready Personalized Instruction on student achievement in four districts in California. The present report examines findings from quantitative analyses comparing math achievement gains, as measured by the Smarter Balanced Assessment (SBA). The analyses use multilevel modeling and propensity score matching to compare math achievement gains between striving learner students assigned to both i-Ready Personalized Instruction and Diagnostic testing and striving learner students only assigned to i-Ready Diagnostic testing. Additional analyses examined relationships between i-Ready usage metrics and SBA scores. Outcome data from the 2021-22 school year were used in all analyses.

The i-Ready Diagnostic assessment is an adaptive assessment designed to provide teachers with actionable insight into student needs. The Diagnostic assessment offers a complete picture of student performance and growth, eliminating the need for multiple, redundant tests. The i-Ready Diagnostic assessment pinpoints student ability level, identifies specific skills students need to learn to accelerate their growth, and charts a personalized learning path for each student.

The i-Ready Personalized Instruction suite delivers online lessons for Grades K-8 students that provide tailored instruction that meets learners at their level, helps them problem solve, and keeps students motivated to continue their progress. Personalized Instruction uses data obtained from the i-Ready Diagnostic assessment to deliver personalized learning paths for each student, balancing rigor and reachability. Online lessons offer students explicit instruction when they need it, along with systematic practice and scaffolded feedback that helps to promote a growth mindset.

This study investigates the efficacy of i-Ready Personalized Instruction by comparing treatment students in one district—students assigned to i-Ready Personalized Instruction in schools that used i-Ready Personalized Instruction (for 5-100% of their students) throughout the school year—to comparison students in three other districts who were not assigned to i-Ready Personalized Instruction in schools that did not use i-Ready Personalized Instruction (or only had up to 5% of students using the program).

Research questions for this evaluation were:

1. What is the impact of i-Ready Personalized Instruction on Smarter Balanced Assessment math achievement for students identified as striving learners?

2. How do program effects on math achievement vary for student subgroups and by school characteristics?
    a. By student demographic characteristics, including grade level, gender, race/ethnicity, disability status, and English language learner status?
    b. By school characteristics, including Title I eligibility, total enrollment size, and the percentage of students of minority race.

3. How do program effects on math achievement vary by implementation as measured by student usage?

Method

Research Design

This study was a quasi-experimental design (QED) that analyzed end of year summative Smarter Balanced Assessment (SBA) data and i-Ready Diagnostic assessment and usage data from the 2021-22 school year. Specifically, student achievement gains on the SBA were compared from treatment students—students assigned to use i-Ready Personalized Instruction in schools in one district that used i-Ready Personalized Instruction school-wide or partially (for 5-100% of their students) throughout the school year—to comparison students in schools in three other districts that did not use i-Ready Personalized Instruction (or only had up to 5% of students using the program). All schools used the i-Ready Diagnostic assessment.

Specifically, as this study focused on striving learners, SBA math scores from the 2021-22 school year were obtained for all students in Grades 3-6 who scored below 30% proficient on the Diagnostic pretest. SBA math scores were also obtained for Grades 7-8, but since no students participated in the treatment district in these grades, analyses were not able to include students in these grades. We also obtained i-Ready Diagnostic scores from the fall of the 2021-22 school year, along with i-Ready usage data for students who used i-Ready Personalized Instruction. Fall 2021 i-Ready scores were used as the prior achievement variable. As i-Ready Personalized Instruction usage tends to be implemented by school, Hierarchical Linear Modeling (HLM) was used to compare student achievement between students assigned to use i-Ready Personalized Instruction and students in schools that only used the Diagnostic testing. We also conducted similar analyses to examine the relationships between usage variables and math achievement.

Participants

Student data were originally obtained from a total of 15,684 students from one treatment school district and three other comparison districts, all in California. This total number includes all striving learner students in Grades 3-6 in these districts. From this sample, 11,878 students were eligible for inclusion in this study, excluding 3,806

students in full or partial treatment schools who only received the Diagnostic product. From this study sample, some students (*n* = 124) were missing school demographic data, another group (*n* = 5) were excluded due to lack of common support during the matching process, and a third group (*n* = 714) did not have outcome data. This resulted in a final sample of 11,034 students from 103 schools, including 22 treatment (partial or full Personalized Instruction) and 81 comparison (Diagnostic-only) schools. Additional attrition calculations are provided in Appendix Table A1, specifically comparing the complete matched sample (11,749) to the final analytic sample.

Student demographics for participants in this study are displayed in Table 1. "Other Race" is defined as races other than White, Black, Hispanic, and Asian, and includes Multiracial, Filipino, Hawaii and Pacific Islander, American Indian, and Alaska Native (which were combined due to variety within and the small numbers of students in these groups). Additionally, Hispanic ethnicity was incorporated separately from race, in accordance with the student demographic information as presented by the districts. Finally, school demographics are also displayed in the latter part of Table 1. Additional information on student socioeconomic status and school urbanicity was not included due to the large number of students missing information on these factors.

Table 1

Student and school characteristics for analytic sample, by treatment group

	Treatment	Comparison
Student Characteristics		
% White	70.0% *	18.3%
% Hispanic	9.5%	57.1% *
% Black	3.0%	12.8% *
% Asian	10.9% *	3.9%
% Other Race	6.1%	7.9% *
% Hispanic (ethnicity)	63.0%	75.8% *
% Female	51.6%	51.4%
% English Language Learners (ELLs)	42.5% *	36.9%
% Students with Disabilities (SWD)	27.6% *	17.5%
School Characteristics		
% Title I	57.9%	88.1% *
Total enrollment	641	772 *
% Students of Minority Race	63.4%	93.9% *
n (students)	1,185	9,850
j (schools)	22	81

Note: * $p < .05$.

Looking at the substantive differences between the treatment and comparison groups, Table 1 shows that the treatment sample contained significantly higher percentages of White and Asian students and lower percentages of Hispanic (race) and Black students relative to the comparison group. Additionally, the treatment group

contained significantly more students who are English Language Learners (ELLs) and Students with Disabilities (SWD).

Comparing school demographics, the bottom part of Table 1 shows that comparison schools had significantly larger enrollments, contained more students of minority race and were more likely to be Title I eligible than treatment schools.

Measures

Data sources for the current study include student i-Ready Diagnostic scores, i-Ready Personalized Instruction usage data, student demographic data, and student SBA achievement data. Math scores were obtained from both i-Ready and SBA assessments. Student achievement data from the 2021-22 school year were analyzed to compare achievement gains between striving learner students assigned to use i-Ready Personalized Instruction and comparison striving learner students in Diagnostic only schools. In addition, i-Ready Personalized Instruction usage data were analyzed to examine relationships between i-Ready usage and SBA test scores.

Smarter Balanced Assessment scores (post-test). The Smarter Balanced Assessment was developed by the Smarter Balanced Assessment Consortium (SBAC) in collaboration with numerous state education agencies to produce valid, reliable, and fair information about students' English Language Arts and math achievement levels relative to the Common Core State Standards (SBAC, 2018). It is a large, computer-adaptive assessment employed for state and federal level accountability and has been shown to be valid and reliable for individual students, at the school level, and for subgroups of students. SBA math scores were obtained from the spring of the 2021-22 school year for all Grades 3-6 students. Spring 2022 math scores were used as the outcome variables in our analyses.

SBA is a vertically scaled assessment used to capture current student achievement and growth over time. Thus, scores across grades can be compared (i.e., a score of 2100 in Grade 4 is equivalent to a score of 2100 in Grade 5). Overall, SBA scale scores fall on a continuum, ranging from approximately 2000–3000. Table 2 shows the average and range of SBA scores by grade level in the sample, which illustrate the increase in achievement scores as grade levels increase. Average SBA scores and proficiencies are low in this sample due to its focus on striving learners.

Table 2

SBA math achievement scores and proficiency, by grade level

Grade level	Average Score	Range of Scores	% Proficient	*n*
3	2352.2	2190 – 2591	9.5%	2,786
4	2393.6	2205 – 2598	6.9%	2,953
5	2415.3	2220 – 2635	4.9%	2,923
6	2420.4	2235 – 2650	3.6%	2,373

SBA additionally separates scores into four achievement levels (1, Not Met; 2, Nearly Met; 3, Met; 4, Exceeded) based on thresholds of proficiency at each grade level. In this analysis we specifically examine whether students were more likely to be proficient (rated Level 3 or 4). The proportion of students who scored proficient at each grade level is also presented above in Table 2.

Demographic variables. The analyses also included a series of demographic variables about students including grade level, race/ethnicity, gender, special education, and English Language Learner variables. Student demographic data was provided by both the i-Ready system and the district. Comparisons between these two data sources revealed minimal discrepancies; district data tended to be more complete and thus was primarily used.

Additionally, school-level demographic variables were included to capture school size (total enrollment), school Title I eligibility, and the proportion of students in the school from minoritized racial/ethnic backgrounds. School demographic data were collected from the National Center for Education Statistics (NCES; https://nces.ed.gov).

i-Ready Diagnostic Scores (pre-test). Overall, i-Ready Diagnostic assessment scores were obtained for Grades 3-6 in the fall of the 2021-22 school year. The fall scores were used as a prior achievement adjustment variable in our main achievement analyses; they were included both in propensity score matching and in the final analytic models as a covariate. i-Ready Diagnostic assessment scores range from 0-800 and are vertically scaled and nationally normed across grades, meaning that scores can be directly compared to each other, regardless of a student's current grade level. In our analyses, i-Ready Diagnostic scores tended to range between 300-600.

i-Ready Personalized Instruction Assignment and Usage data. i-Ready data were obtained for all students who were tested by i-Ready (using the Diagnostic assessments) in the 2021-22 school year. Data regarding the assignment (at the individual and school level) and usage of the Personalized Instruction product was provided by Curriculum Associates. The usage data consists of time spent on i-Ready Personalized Instruction lessons and instruction only and thus, does not include time spent on Diagnostic assessments. Usage metrics provided by Curriculum Associates

included: total lessons completed, unique lessons completed, passed lessons, lesson passing rate, total minutes of usage, weeks with at least one completed lesson, and average minutes per week.

Regarding discrepancies between assignment and usage, in this study we did not observe any cross-over students from treatment to comparison, and thus all treatment students had non-zero values on all usage metrics. However, there were 25 cross-over students from comparison to treatment: comparison (Diagnostic-only students) who had non-zero values on usage metrics. These cross-over students all had low usage values (in the bottom 30% of all users). Additionally, 126 (10.6%) treatment students completed only one lesson. Both cross-over and low usage students were retained in their original assigned condition. Supplemental analyses without the cross-over or low usage students produced the same conclusions presented below in our main analysis.

Analytical Approach

Data for students in Grades 3-6 were analyzed by descriptively examining patterns of SBA and i-Ready Diagnostic scores and usage, as well as by comparing achievement patterns between "striving learner" students who were assigned to use i-Ready Personalized Instruction (Treatment students) and students assigned to only receive i-Ready Diagnostic assessments (Comparison students). Hierarchical Linear Modeling (HLM) was used to compare differences in achievement, as measured by the SBA, between treatment and comparison students due to the substantial intraclass correlation value of the outcome (ICC = .071) indicating that a large amount of the variation in the final test score was due to factors at the school level. Schools were chosen as the clustering variable for Level 2, as i-Ready Personalized Instruction assignment varies by school (rather than classroom). The final HLM model used to estimate the impact of treatment on math achievement was:

Level-1 (Student) Equation:

$$SBAMathScore_{ij} = \beta_{0j} + \beta_{1j}Pretest_{ij} + \beta_{kj} * \sum Student\ Covariates_{ij} + r_{ij}\ [weight_i]$$

Level-2 (School) Equations:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Treatment_{0j} + \gamma_{0n} * \sum School\ Covariates_{0j} + u_{0j}$$
$$\beta_{1j} = \gamma_{10}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{kj} = \gamma_{k0}$$

where $SBAMathScore_{ij}$ is the SBA math score in spring 2022 for student *i* in school *j*, γ_{00} is the covariate-adjusted grand mean test score for the comparison group; γ_{10} is the regression coefficient for the pretest; *Pretest* is the student's i-Ready Diagnostic math

score in fall 2021; γ_{k0} are the vector of regression coefficients for the *k* student covariates; *Student Covariates* are the vector of student covariates (grade level, gender, race/ethnicity, ELL status, and special education status); r_{ij} is the student-level residual; $[weight_i]$ is the propensity score weight (from the matching process further described below) for student *i*; γ_{01} is the average treatment effect; *Treatment* is the binary treatment indicator for school *j*; γ_{0n} are the regression coefficients for the *n* school covariates; *School Covariates* is the vector of school covariates (Title I eligibility, enrollment, and percentage of students of minority race); and u_{0j} is the random school effect for school *j*. All continuous covariates (pretest, school enrollment, and school percentage of students of minority race) were grand-mean centered to facilitate interpretation of the intercept. Effect sizes were calculated using the unadjusted pooled standard deviation of the outcome.

Proficiency Outcome. We additionally investigated the impact of i-Ready Personalized Instruction on the likelihood of students scoring proficient. This model used the same input variables but used a logistic regression to predict the binary outcome that indicated whether each student scored proficient (or not) on the Spring 2022 SBA assessment. Results are presented in odds ratio form and are also translated into probabilities: from log odds to odds (by exponentiation) and then probabilities (p = odds/1+odds) to illustrate how many out of 100 students would be predicted to score proficient. Estimated probabilities for the comparison group are estimated in the same way by converting the intercept coefficient into a probability. The treatment group probabilities accordingly combine the intercept added to the treatment coefficient.

Subgroup Variation. We also analyzed how the impact of i-Ready Personalized Instruction varied by student subgroups. To test if the impact of treatment was different for different types of students, we included an interaction term between the student subgroup (such as students of Hispanic ethnicity) and treatment. This allowed us to estimate the relationship between treatment and achievement growth specific to a student subgroup (such as students of Hispanic ethnicity versus those not of Hispanic ethnicity). We tested all subgroups included in the model (used as covariates) and present results on the significant differences we observed: student grade level and student race.

Usage Analyses. For usage analyses, HLM models tested the unique effect of each usage metric by adding the continuous usage measures to the model with the binary treatment variable (i-Ready Personalized Instruction vs. Diagnostic testing only). This allowed us to estimate the effect of individual units of instruction, such as one hour of usage or one completed lesson, on SBA scores in relation to students who did not use i-Ready Personalized Instruction. Usage variables were not mean centered.

Propensity Score Weighting to Achieve Baseline Equivalence. Baseline equivalence was met for fall 2021 i-Ready math scores. Baseline equivalence is defined as being met if the standardized mean difference between treatment and comparison

groups is less than 0.25 SD (WWC, 2020). Unadjusted means for fall 2021 math i-Ready scores are presented in Table 3. However, multiple large, standardized differences exceeding 0.25 SD between the treatment and comparison groups on the covariates indicated that there were still substantial differences between the groups. The apparent balance on prior achievement but not on other covariates is not surprising in this study because of the restricting to striving learners (effectively selecting students of similar performance at baseline).

Table 3

Baseline equivalence on covariates, unadjusted

	Treatment		Comparison		Stan. Mean
Outcome	Mean	SD	Mean	SD	Diff.
Fall 2021 i-Ready score	420.82	27.48	416.95	27.48	.141
White (race)	.700	.459	.183	.387	1.218
Hispanic (race)	.095	.294	.571	.495	−1.169
Black (race)	.030	.169	.128	.334	−.371
Asian (race)	.109	.312	.038	.192	.272
Other (race)	.061	.23*9*	.079	.269	−.070
Hispanic (ethnicity)	.630	.483	.758	.428	−.283
Female	.516	.500	.514	.500	.004
ELL	.425	.495	.369	.483	.115
SPED	.276	.447	.175	.380	.244
Grade 3	.213	.409	.257	.437	−.105
Grade 4	.242	.429	.271	.444	−.065
Grade 5	.284	.451	.263	.440	.047
Grade 6	.262	.440	.209	.407	.123
n	1,185		9,850		

Note: SD=standard deviation.

To adjust for the large standardized mean differences between treatment and comparison students on baseline achievement, propensity score matching (PSM) was used in all analyses for the purpose of creating comparison groups that were as similar as possible to groups of treatment students. First, prior to the receipt of any outcome data, treatment students were matched to similar comparison students (using a radius matching approach with a caliper of .05 standard deviations of the propensity score). All comparison cases within the caliper range were matched to that treatment case. Next, during the analysis, treatment students were each given a weight of one, and comparison students were each weighted based on the total number of treatment cases to which they were matched.

The result of these PSM and weighting procedures was that comparison students who were more similar to treatment students (in terms of prior achievement and demographic covariates) were weighted more heavily in the analyses, and comparison students who were less similar to treatment students were weighted less. After these weights were applied to comparison students, baseline equivalence was maintained for fall 2021 math i-Ready Diagnostic scores and achieved for most covariates of concern, with standardized mean differences of magnitude of less than 0.24. Table 4 shows the adjusted baseline differences on the prior achievement measure.

Table 4

Baseline equivalence on covariates, adjusted

	Treatment		Comparison		T vs. C Difference	Pooled Unadjusted	Stan. Mean
Outcome	Mean	SD	Mean	SD	(adjusted)	SD	Diff.
Fall 2021 i-Ready score	420.82	27.48	418.39	27.67	2.428	27.505	.088
White (race)	.700	.459	.300	.458	.400	.426	.939
Hispanic (race)	.095	.294	.448	.498	–.353	.500	–.706
Black (race)	.030	.169	.038	.191	–.008	.322	–.025
Asian (race)	.109	.312	.121	.326	–.012	.209	–.058
Other (race)	.061	.23*9*	.087	.282	–.026	.266	–.099
Hispanic (ethnicity)	.630	.483	.612	.487	.017	.436	.039
Female	.516	.500	.523	.500	–.008	.500	–.015
ELL	.425	.495	.414	.493	.012	.484	.024
SPED	.276	.447	.226	.418	.050	.389	.129
Grade 3	.213	.409	.249	.432	–.036	.434	–.083
Grade 4	.242	.429	.270	.444	–.027	.443	–.062
Grade 5	.284	.451	.252	.434	.032	.441	.072
Grade 6	.262	.440	.230	.421	.032	.411	.077
n	1,185		9,850				

Notes: 1. SD=standard deviation; T = treatment; C = comparison. 2. All estimates include propensity-score weights.

Results

i-Ready Personalized Instruction usage. We first descriptively examine patterns of i-Ready usage by grade level. "i-Ready usage" data refers only to usage of i-Ready Personalized Instruction, not including any time spent on Diagnostic assessments. Table 5 shows descriptive statistics relating to i-Ready math usage metrics for all treatment students in the analytic sample for the analyses that follow. As previously discussed, some comparison students (25) had non-zero usage; this group of students are not included as treatment students in analyses and thus are represented in the usage metric averages below.

Table 5

i-Ready Personalized Instruction usage means and standard deviations for treatment students in math, by grade level

	Grade 3	Grade 4	Grade 5	Grade 6
Total lessons	20.10	19.17	14.35	13.26
	(19.91)	(21.43)	(18.12)	(14.23)
Unique lessons	18.83	17.20	12.59	11.34
	(18.30)	(18.67)	(15.69)	(11.79)
Passed lessons	18.28	16.27	11.64	10.21
	(17.80)	(17.55)	(14.83)	(10.66)
Lesson passage rate	91.8%	86.2%	81.5%	77.9%
	(13.0)	(16.9)	(22.5)	(22.9)
Minutes of Usage	410.79	443.81	340.72	385.36
(total)	(408.50)	(536.56)	(468.38)	(444.71)
Weeks of Usage	12.26	11.82	9.64	10.18
	(7.12)	(7.46)	(7.04)	(7.02)
Minutes Per Week	28.89	32.45	29.29	33.29
(average)	(16.66)	(19.37)	(17.56)	(20.17)
n	252	287	336	310

Note: Standard deviations are presented in parentheses below the mean.

On average, students in the treatment group completed 16 lessons, using i-Ready for a total of 392 minutes, over a total of 11 weeks for an average of 31 minutes per week. Usage metrics were generally highest for Grades 3 and 4. Although Grades 5 and 6 students completed fewer overall lessons and over fewer weeks than other grades, their usage was more concentrated, as indicated by their high average minutes per week usage. Another trend across grade levels is that students had lower passage rates in higher grades.

Distributionally, total usage figures (like total lessons, total minutes) were mostly positively skewed (which is evidenced in the large standard deviation numbers in Table

5, which are nearly as large as the mean for some measures). This means that a large number of students had infrequent usage, with one-quarter of students completing fewer than 4 lessons or 94 minutes of activity (and averaging less than 19 minutes per week). Metrics for the total number of weeks used and the average minutes per week had more evenly distributed values (and fewer students clustered close to 0) but were still positively skewed. The only exception to this right skew was the lesson passage rate which had a negative skew with most students having high passage rates close to 100%.

Achievement descriptive statistics. In Table 6, we present, by treatment group, fall 2021 i-Ready and spring 2022 SBA math scores, as SBA scores were the main outcome variable in our analyses.

Table 6

Average unadjusted i-Ready and SBA math scores, 2021-22, by treatment group

	Treatment	Comparison
Fall i-Ready score	420.82	416.95
Spring SBA score	2413.51	2392.39
% proficient	9.3%	6.0%
N	1,185	9,850

Note: Means are unadjusted.

As noted previously, prior to any adjustments, treatment students scored higher on the fall i-Ready assessment than did comparison students. These differences on test scores, however, are minimal, even prior to adjustment as all students are striving learners (30% proficient or below). Pre-test scores were skewed negatively to the left for both groups (due to the ceiling effect from the sample restriction). Unadjusted spring SBA scores and percent proficient also tended to be slightly higher, on average, for treatment students. The main analyses below use propensity score matching and covariates to adjust for these potential differences.

Main achievement analyses

In this section, we present the results of analyses examining the effect of i-Ready Personalized Instruction on math achievement among the subgroup of striving learners. We first present results for all Grades 3-6 on SBA scores and then SBA proficiency.

SBA scores. Results of analyses examining the impact of treatment on SBA math scores are found in Table 7. We report unstandardized regression coefficients, standard errors, and effect sizes in this table.

Table 7

Analyses of i-Ready Personalized Instruction on SBA math scores

Outcome	Estimate	Standard Error	*p* value	Effect size
Treatment	13.178**	4.630	.004	.175
Constant	2422.422	6.197		

Notes: 1. *N* = 11,034; *j* (schools) = 103. 2. Adjusted estimates with PSM weights and covariates. 3. ** *p* < .01.

There was a statistically significant effect of treatment on SBA math scores in relation to the comparison condition. The impact estimate in Table 7 can be interpreted as the average difference between treatment and comparison students. Thus, the regression estimate indicates that after adjusting for prior achievement and demographics, treatment students scored an average of 13 points higher on the SBA math test than did comparison students.

Interpreting the effect size, after adjusting for prior achievement and demographics, treatment students scored an average of 0.18 standard deviations higher on the SBA math assessment than did comparison students. In relation to effect sizes from randomized control trials evaluating the impact of interventions on math scores in large samples, this effect size is in the 70-90th percentile of study impacts, indicating a substantial impact and efficacious intervention (Kraft, 2020). Interpreted as percentile growth, the average comparison student would be predicted to score 6.9 percentile points higher (moving from the 50 to 56.9 percentile rank) if they had received the intervention.

SBA Proficiency. We also examined the impact of treatment on students' likelihood of achieving SBA proficiency (a yes/no outcome). In this study, however, there was no significant impact on this outcome, which is likely due to the limited variation in proficiency levels in the sample of striving learners. The results of this analysis are presented in Appendix Table B1.

Subgroup analyses

We also conducted a series of supplementary analyses in which we examined the impact of treatment across different student subgroups including grade level, student race/ethnicity, gender, ELL status, SPED status, and school Title I eligibility. Specifically, these models examine whether the impact of treatment was different for different types of students; for example, did Black students in treatment gain more in math achievement than Black students not in treatment or than White students in treatment? The models used in these analyses are identical to those used in the main analysis, with the addition of an interaction term between the subgroup and the treatment condition to estimate separate treatment impacts for each subgroup. We did not find significant

differences in treatment effect when examining student gender, ELL status, or SPED status. However, we did find differences by grade level and student race.

Grade level. We first examined how the impact of treatment varied across grade levels. The results of this analysis are shown in Table 8.

Table 8

Impacts of i-Ready Personalized Instruction on SBA math achievement, by grade level

Grade	Estimate	Standard Error	*p* value	Effect size	*n*
3	10.670	6.356	.093	0.142	2,786
4	23.160***	5.582	.000	0.308	2,953
5	13.362*	6.307	.034	0.177	2,923
6	5.834	6.565	.374	0.077	2,373

Notes: 1. *N* = 11,034; *j* (schools) = 103. 2. Adjusted estimates with PSW weights and covariates. 3. * *p* < .05, *** *p* < .001.

Across the grade levels, results were more positive for Grades 3 through 5 although the effect of treatment only reached statistical significance in Grades 4 and 5 when separated out by grade (*p* < .001 and .05, respectively). In Grade 4, students in the treatment condition averaged 23 points higher on the SBA math assessment than their Grade 4 peers in the comparison condition, which equated to an effect size of 0.3 standard deviations.

Student Race. We also examined how students of different racial backgrounds were differentially impacted by treatment. The results of this analysis are shown in Table 9.

Table 9

Impacts of i-Ready Personalized Instruction on SBA math achievement, by student race

Race	Estimate	Standard Error	*p* value	Effect size	*n*
White	14.291**	4.456	.001	0.190	2,653
Black	23.176*	9.216	.012	0.308	1,292
Asian	0.635	6.580	.923	0.008	507
Hispanic	15.600*	7.622	.041	0.207	5,737
Other	14.371	8.688	.098	0.191	845

Notes: 1. *N* = 11,034; *j* (schools) = 103. 2. Adjusted estimates with PSW weights and covariates. 3. * *p* < .05, ** *p* < .01.

Looking at the impact of treatment between these racial groups, there are different trends across the groups and a substantial difference in the treatment impact by race. Treatment was estimated to have a positive impact on math achievement for

all races of students. This impact was highest for Black students; treatment was significantly positively associated with a 23-point greater gain in math achievement ($p < .05$) than comparison students who were also Black, representing an effect size of 0.3 standard deviations. White, Hispanic and Other race students had the next highest gains between 14 and 16 points, compared to their peers of the same race in the comparison group (representing an effect size of around 0.2 standard deviations). Finally, students of Asian race assigned to treatment performed similarly to their peers in the comparison group, with an effect of practically 0.

Referring back to Table 1, comparing the demographics of treatment and comparison students, only a small percentage (3%) of students in treatment schools were of Black race compared to a larger share (12.8%) in comparison schools. This greater benefit for Black students is encouraging within the treatment group schools. Comparing treatment and comparison groups, however, there could be other differences between the Black student populations and their school experiences that explains this large differential impact for this group.

Usage Analyses

Next, we present a series of analyses examining the associations between i-Ready Personalized Instruction usage metrics and math achievement. These analyses are identical to the previous achievement analyses, with the addition of an i-Ready usage variable in each model. A separate analysis (and model) was run for each i-Ready usage measure. i-Ready usage metrics used in these analyses include the number of completed lessons, number of unique lessons, and number passed lessons, along with total minutes and weeks of usage. We follow this up with similar analyses examining the relationships between instructional usage categories (based on weeks used) and achievement. Finally, we examine the relationship between usage of i-Ready in accordance with program guidance and achievement.

Math usage. We present the results of analyses examining the effects of i-Ready Personalized Instruction usage variables on math achievement. Table 10 shows the unstandardized coefficients of all usage variables, which estimate the impact of one unit of usage on achievement, compared to no usage (for those in the comparison or treatment group).

Table 10

Associations between i-Ready Personalized Instruction usage and math achievement

Usage Measure	Estimate	Standard Error
# of lessons	0.238**	0.072
# of unique lessons	0.287***	0.079
# of passed lessons	0.305***	0.075
Minutes of Usage (total)	0.008**	0.003
Weeks of Usage	0.073	0.245

Note: N = 11,034; ** $p < .01$, *** $p < .001$.

All the usage variables were significantly positively associated with SBA math scores with the exception of weeks of usage. Further, all but one of these variables were significant at the .01 alpha significance level. Coefficients in Table 10 can be interpreted as the expected change in SBA math score for every unit of a usage variable. For example, looking at the second to last row, every extra minute of usage was associated with a 0.008-point increase in SBA math score. Thus, usage for the average student, who completed 392 minutes of instruction, would be associated with a 3.1-point SBA math score increase. Similarly, the average user who completed 16 lessons would be expected to gain 3.8-points on their SBA math score. Furthermore, for the treatment student who followed the guidelines of 18 weeks of usage, this amount of usage would be associated with a 1.3-point SBA math score increase. However, it should be noted that while our model estimates this linear, similar impact across all usage values, there may, in reality, be different returns at different values of usage. For example, going from 10-20 minutes may increase scores more than going from 160 to 170 minutes.

Weeks Used (Categories). We next examined the relationships between i-Ready Personalized Instruction groupings, as measured by weeks of usage, and achievement, as measured by SBA scores. Usage categories were created because, in previous analyses relating curriculum usage and achievement, these relationships tended to be curvilinear, with the lowest and highest users tending to have the highest levels of achievement. Thus, these relationships are potentially underestimated in simple (linear) correlational and regression analyses. The construction of usage categories allows for a more partitioned investigation of the relationship between instructional usage and achievement for different levels of usage. We constructed usage categories based on the number of weeks of Personalized Instruction usage for students with non-missing SBA scores, with the low category capturing students with 0-9 weeks of usage, average capturing 10-17 weeks, and high 18 or more weeks. These categories were chosen in relation to the recommended usage of 18 weeks: the low category containing students meeting less than half the threshold, high category meeting the recommended threshold and then average falling in between. A positive association would indicate that a usage category was associated with higher SBA

scores, in relation to comparison students, who had no i-Ready Personalized Instruction usage.

Analyses were performed for math usage categories and achievement. The models used in these analyses are identical to those used in previous variables, with the replacement of the treatment variable with three dummy variables representing the categories of usage. Usage category ranges and sizes are shown in Table 11.

Table 11

Usage weeks categories ranges and sample sizes

Usage Category	Range	*n*	% of Treatment
1 – Low	0 – 9 weeks	579	48.9%
2 – Average	10 – 17 weeks	353	29.6%
3 – High	18 or more weeks	255	21.5%

Note: 9,848 students in comparison group with 0 weeks.

Looking at Table 11, there are smaller groups of students in each ascending category, i.e., fewer students use at the higher levels, including above the recommended usage level of 18 weeks. Usage in the high category ranged between 18 to 32 weeks.

Weeks used (categories) and math achievement. We now present the results of analyses examining the association between i-Ready Personalized Instruction usage categories (based on weeks used) and SBA math scores. Table 12 shows the unstandardized regression coefficients for each usage category, in relation to students with no usage (i.e., comparison students).

Table 12

Associations between i-Ready Personalized Instruction usage categories and math achievement

Usage Category	Estimate	Standard Error
1 – Low	13.689**	5.150
2 – Average	12.341**	4.373
3 – High	12.779*	5.541

Note: $N = 11{,}035$; * $p < .05$, ** $p < .01$.

A significant positive relationship between usage and math achievement was observed for users in all categories of i-Ready Personalized Instruction usage. The regression coefficients in Table 12 can be interpreted as the average increase in SBA math score for a student in a particular i-Ready Personalized Instruction usage

category, in relation to comparison students who did not use i-Ready Personalized Instruction. Thus, high usage (of at least 18 weeks) was associated with a 12.8-point increase in SBA math score, in relation to comparison students. A statistically significant positive relationship was found between all three usage categories (low, average, and high) and math achievement, indicating that students who used i-Ready Personalized Instruction at any level outscored, on average, comparison students. This relationship was of a similar magnitude (12.3–13.7) across all three usage categories.

CA usage guidelines. Curriculum Associates provides recommended i-Ready Personalized Instruction usage guidelines to educators. Specifically, Curriculum Associates recommends individual students aim for a consistent 30-45 minutes of i-Ready Personalized Instruction usage per subject per week over a minimum of 18 weeks. To identify students who met Curriculum Associates' recommended guidelines, we operationalized this guidance as follows:

- An average of at least 30 minutes per week of i-Ready Personalized Instruction use
- At least 18 weeks of i-Ready Personalized Instruction use

Students that met both of these guidelines were classified as having met usage guidelines, while those that did not meet both criteria were classified as not having met usage guidelines. Table 13 displays the percentages of i-Ready Personalized Instruction students who met i-Ready Personalized Instruction usage guidelines in each grade.

Table 13

Percentages of treatment students meeting i-Ready Instruction usage guidelines, by grade

Grade	Met i-Ready usage guidelines	*n*
3	20.6%	252
4	20.6%	287
5	12.2%	336
6	11.9%	310

Over all grades, 15.9% of students met both usage guidelines (43% meeting the first guideline of 30 minutes per week and 22% meeting the second guideline of 18 weeks). As shown in Table 13, about one-fifth of treatment students met usage guidelines in Grades 3 and 4, but percentages of students reaching usage guidelines were much lower in the upper grades.

CA usage guidelines and math achievement. We also conducted analyses that examined the relationship between meeting or not meeting i-Ready usage guidelines on achievement, in relation to comparison students. The models used in

these analyses are similar to those used in previous analyses, except a dichotomous variable indicating i-Ready Personalized Instruction students who did or did not meet usage guidelines is added in addition to the treatment variable. Comparison students were assigned a "0" for both dichotomous variables. This allowed us to uniquely estimate the relationship between SBA math achievement and Personalized Instruction usage that did or did not meet CA's guidelines. Results of these analyses are shown in Table 14.

Table 14

Association between meeting i-Ready usage guidelines and SBA math achievement

	Estimate	Standard Error	*n*
Did not meet usage guidelines	12.338*	4.861	996
Met usage guidelines	4.811	5.138	189

Notes: 1. These estimates are in comparison to the 9,850 comparison students. 2. * $p < .05$.

Achievement gains differed between treatment students who did and did not meet i-Ready Personalized Instruction usage guidelines in math, although both associations were positive. Students who did not meet i-Ready usage guidelines averaged 12.3-point greater gains on the SBA math exam than comparison students. The magnitude of the increase was lower for treatment students who met usage guidelines, and this association (estimating a 4.8-point gain over comparison students) was not statistically significant.

Discussion

The purpose of this evaluation was to examine the impact of i-Ready Personalized Instruction on math achievement, as measured by SBA scores. We compared "striving learner" students who were assigned to use i-Ready Personalized Instruction (Treatment students) and "striving learner" students assigned to only receive i-Ready Diagnostic assessments (Comparison students). We also examined associations between various i-Ready usage metrics and achievement.

In interpreting the findings of this evaluation, some limitations should be noted. First, while we adjusted for as many demographic variables as possible, some student information, namely economic disadvantage, were not available from all school districts involved in this evaluation, meaning we were unable to adjust for these variables or conduct relevant subgroup analyses. Similarly, we had access only to spring SBA scores and i-Ready score and usage data from the 2021-22 school year. This limited our analyses to only one year and to strictly quantitative measures, which precluded drawing any substantive conclusions regarding the fidelity of implementation within classrooms by teachers and students, outside of the quantitative usage data supplied to us by Curriculum Associates. Analyses of usage data suggested varied usage by

individual students and that these usage amounts potentially contributed to the observed impacts of the program.

Math Achievement Gains

We found that striving learner treatment students had statistically significantly higher SBA math scores than comparison students assigned to use only the Diagnostic testing. Specifically, students in the treatment group scored, on average, 13 points higher on the SBA math test than their peers in the comparison group.

Additionally, we observed that the relationship between treatment and math achievement varied by student subgroups. By grade level, treatment students in Grade 4 had the highest gains (23 points) in SBA math achievement, followed by students in Grade 5 (13.4 points). By student race, Black treatment students benefited the most from treatment, with gains of 32 points over their Black peers in the comparison group; however, Black students made up only a small proportion of the treatment group (3%) and may have been different in other ways from their Black peers in the comparison group.

Usage Patterns

Descriptive analysis of usage by grade revealed that usage metrics were generally highest in third and fourth grade but comparable across all grades. Students in Grade 6 tended to complete fewer lessons but participated in i-Ready Personalized Instruction in a more concentrated fashion over fewer weeks.

Multilevel, weighted regression analyses with usage variables showed that five of the six usage metrics we considered were significantly positively related to student math achievement. Specifically, average usage (392 minutes of instruction) was associated with a 3.1-point greater math achievement on the SBA math assessment.

Conclusion

Overall, this analysis presents promising evidence of i-Ready Personalized Instruction on student math achievement. This relationship between i-Ready Personalized Instruction and math achievement was strongest for students in Grades 4 and 5 and students identified as Black. Usage was also related to achievement and future studies should seek to further examine the reasons behind this variation in usage.

References

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241-253.

SBAC (2017). Smarter Balanced Assessment Consortium: 2017–18 Summative Technical Report. *Smarter Balanced Assessment Consortium.*

Appendix A: Attrition Table

Table A1

Attrition between assignment and analysis

	Pre-test group (n, at assignment)	Post-test group (n, final analytic sample)	Attrited students (n)	Attrition %
Treatment	1356	1185	171	12.61%
Comparison	10393	9850	543	5.22%
Total	11749	11035	714	6.08%
			differential attrition	7.39%

Appendix B: Supplementary Results Table

Table B1

Analyses of i-Ready Personalized Instruction on SBA math proficiency

Outcome	Odds Ratio	Standard Error	*p* value
Treatment	1.585	0.312	.019
Constant	6.675	1.219	

Notes: 1. *N* = 11,034; *j* (schools) = 103. 2. Adjusted estimates with PSM weights and covariates.