

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

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:	1
The Impacts of i-Ready Personalized Instruction on Student Math Achievement.....	1
School-wide i-Ready Personalized Instruction was associated with greater math achievement gains and proficiency.....	2
Greater i-Ready Personalized Instruction usage was associated with math achievement gains.	2
Conclusions	2
The Impacts of i-Ready Personalized Instruction on Student Math Achievement.....	1
Method.....	2
Research Design.....	2
Participants	2
Measures	4
Smarter Balanced Assessment scores (post-test).....	4
Demographic variables.	5
i-Ready Diagnostic Scores (pre-test).....	5
i-Ready Personalized Instruction Assignment and Usage data.....	5
Analytical Approach	6
Results	9
i-Ready Personalized Instruction usage.....	9
Achievement descriptive statistics.....	11
Main achievement analyses	11
SBA Scores.	11
Subgroup analyses.....	13
Usage Analyses	15
Math usage.....	15
Weeks used (categories) and math achievement.....	16
CA usage guidelines.	17
CA usage guidelines and math achievement.....	18
Discussion	19
Math Achievement Gains.....	19

Usage Patterns	20
Conclusion	20
References	21
Appendix A: Attrition Table.....	22

EXECUTIVE SUMMARY:

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

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 a large, suburban school district in southern 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 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?
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 5,330 students in 22 schools from Grades 3-6 from one school district in southern California. All schools used i-Ready Diagnostic assessments, but some schools were considered “full instruction” and assigned all students to receive both the Diagnostic and Personalized Instruction products. Students in these schools were compared to “diagnostic-only” students in other “partial instruction” schools where only some students received the Instruction product in addition to the Diagnostic product. (These Diagnostic and Instruction students in partial instruction schools are not included in this study.)

Key findings of the current study include:

School-wide i-Ready Personalized Instruction was associated with greater math achievement gains and proficiency. School-wide implementation of i-Ready Personalized Instruction was associated with greater gains on the math SBA exam and an increased likelihood of scoring proficient on the math SBA exam. This impact was strongest for students in Grade 4 and in Title I eligible schools.

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. Additionally, math Personalized Instruction students who met usage guidelines scored higher than did comparison students.

Conclusions

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

- School-wide i-Ready Personalized Instruction usage was associated with significantly higher SBA math scores and was also significantly associated with greater odds of scoring proficient on the SBA math test.
- Students assigned to i-Ready Personalized Instruction in Grade 4 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 Grade 4 students compared to the other grades.
- Students assigned to i-Ready Personalized Instruction in Title I eligible schools experienced greater growth in math achievement than students in non-Title I eligible schools who were assigned to i-Ready Personalized Instruction.
- i-Ready Personalized Instruction usage metrics including total time and lesson count variables were significantly positively associated with math achievement.
- Meeting i-Ready Personalized Instruction usage guidelines was associated with significantly higher math achievement, in relation to comparison students.

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

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 a large, suburban school district in southern 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 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.

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 in one school district by comparing treatment students—those in schools that used i-Ready Personalized Instruction school-wide (for at least 70% of their students) throughout the school year—to comparison students who were not assigned to i-Ready Personalized Instruction in schools that used i-Ready Personalized Instruction for only some students (between 5-33%).

Research questions for this evaluation were:

1. What is the impact of i-Ready Personalized Instruction on Smarter Balanced Assessment math achievement?

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—those in schools that used i-Ready Personalized Instruction school-wide (for at least 70% of their students) throughout the school year—to comparison students who were not assigned to i-Ready Personalized Instruction in schools that used i-Ready Personalized Instruction for only some students (between 5-33%). All schools used the i-Ready Diagnostic assessment.

Specifically, SBA math scores from the 2021-22 school year were obtained for all students in Grades 3-6. 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 in school-wide i-Ready Personalized Instruction schools and comparison students only using the Diagnostic tool (in schools where only some students received the Instruction product). 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 6,559 students from one school district in California. This number includes all students in Grades 3-6 in this district. There were 5,799 students eligible for inclusion in this study, excluding 813 students in partial treatment schools who received both Instruction and Diagnostic products. From this study sample, some students ($n = 93$) were missing school demographic data, another group ($n = 4$) were excluded due to lack of common

support during the matching process, and a third group ($n = 319$) did not have outcome data. This resulted in a final sample of 5,330 students from 22 schools, including 7 treatment (school-wide Personalized Instruction) and 15 comparison (partial Personalized Instruction) schools. Additional attrition calculations are provided in Appendix Table A1, specifically comparing the complete matched sample (5,649) 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 district. 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	68.3% *	63.1%
% Hispanic	7.5% *	4.5%
% Black	2.7%	2.2%
% Asian	16.1%	24.6% *
% Other Race	4.9%	4.0%
% Hispanic (ethnicity)	54.4% *	28.1%
% Female	49.5%	47.8%
% English Language Learners (ELLs)	31.1% *	10.4%
% Students with Disabilities (SWD)	13.3% *	10.4%
School Characteristics		
% Title I	72.6% *	21.3%
Total enrollment	741 *	637
% Students of Minority Race	68.4% *	53.2%
n (students)	1,952	3,378
j (schools)	7	15

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 Hispanic (in both race and ethnicity) students and lower percentages of Asian 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, Table 1 shows that treatment schools had significantly larger enrollments and a greater percentage of students of minority race and were three times more likely to be Title I eligible than comparison 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 the Smarter Balanced Assessment. Student achievement data from the 2021-22 school year were analyzed to compare achievement gains between students in school-wide i-Ready Personalized Instruction schools and comparison students in partial i-Ready Personalized Instruction schools who only took the Diagnostic assessment. 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 student current 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.

Table 2

SBA math achievement scaled scores and proficiency, by grade level

Grade level	Average Score	Range of Scores	% Proficient	<i>n</i>
3	2459.3	2190 – 2600	63.7%	1,263
4	2512.7	2205 – 2700	67.0%	1,257
5	2527.3	2220 – 2740	54.4%	1,389
6	2559.1	2235 – 2780	55.5%	1,421

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 comparison to treatment, and thus all Diagnostic-only (comparison students) had 0's on all usage metrics. However, in this study, there were 282 (14.4%) cross-over students from treatment to comparison (i.e., treatment students with no usage metrics) and 146 (8.7%) treatment students who completed only one lesson. Both cross-over and low usage students were retained in their original

condition (treatment). 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 scores and usage, as well as by comparing achievement patterns between students in schools using i-Ready Personalized Instruction school-wide (Treatment students) and students who only received i-Ready Diagnostic assessments and who were in schools only partially using i-Ready Personalized Instruction (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 = .151) 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:

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

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 ($p = \text{odds}/1+\text{odds}$) 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 school Title I eligibility.

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 Matching and Weighting to Achieve Baseline Equivalence. Initially, baseline equivalence was not 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 and other covariates are presented in Table 3.

Table 3*Baseline equivalence on covariates, unadjusted*

Outcome	Treatment		Comparison		Stan. Mean Diff.
	Mean	SD	Mean	SD	
Fall 2021 i-Ready score	449.41	36.26	465.91	33.73	-.471
White (race)	.683	.465	.632	.482	.109
Hispanic (race)	.075	.263	.045	.207	.126
Black (race)	.027	.161	.023	.148	.027
Asian (race)	.161	.368	.246	.431	-.213
Other (race)	.049	.215	.040	.195	-.044
Hispanic (ethnicity)	.544	.498	.282	.450	.552
Female	.495	.500	.478	.500	.035
ELL	.311	.463	.104	.305	.529
SPED	.133	.339	.104	.305	.090
Grade 3	.221	.415	.246	.431	-.060
Grade 4	.220	.415	.245	.430	-.058
Grade 5	.283	.450	.248	.432	.079
Grade 6	.276	.447	.262	.439	.034
<i>n</i>	1,952		3,378		

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 achieved for fall 2021 math i-Ready Diagnostic scores and all covariates of concern, with a standardized mean difference of magnitude of less than 0.24, as shown in Table 4.

Table 4*Baseline equivalence on covariates, adjusted*

Outcome	Treatment		Comparison		T vs. C Difference (adjusted)	Pooled Unadjusted SD	Stan. Mean Diff.
	Mean	SD	Mean	SD			
Fall 2021 i-Ready score	449.41	36.26	447.03	36.24	2.387	35.577	.067
White (race)	.683	.465	.650	.477	.034	.477	.071
Hispanic (race)	.075	.263	.114	.318	-.039	.230	-.170
Black (race)	.027	.161	.025	.157	.001	.153	.009
Asian (race)	.161	.368	.145	.353	.015	.411	.038
Other (race)	.049	.215	.057	.231	-.008	.203	-.039
Hispanic (ethnicity)	.544	.498	.577	.494	-.033	.485	-.069
Female	.495	.500	.502	.500	-.006	.500	-.013
ELL	.311	.463	.336	.473	-.025	.384	-.066
SPED	.133	.339	.132	.338	.001	.318	.003
Grade 3	.221	.415	.238	.426	-.018	.425	-.041
Grade 4	.220	.415	.257	.437	-.037	.425	-.087
Grade 5	.283	.450	.254	.435	.029	.439	.066
Grade 6	.276	.447	.250	.433	.026	.442	.058
<i>n</i>	1,952		3,378				

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 treatment students (242, or 14.4% of all treatment students) had zero usage, but this group of students are still included as treatment students in analyses and 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	13.64 (14.96)	22.94 (20.80)	10.74 (15.54)	8.11 (13.78)
Unique lessons	12.99 (12.75)	21.06 (18.70)	9.54 (13.49)	6.92 (11.42)
Passed lessons	12.75 (14.07)	20.30 (18.09)	8.94 (12.94)	6.27 (10.47)
Lesson passage rate	85.0% (15.2)	84.5% (22.3)	69.0% (36.5)	56.9% (40.4)
Minutes of Usage (total)	287.99 (317.15)	544.88 (474.70)	264.30 (386.82)	266.41 (499.23)
Weeks of Usage	10.59 (6.59)	14.15 (7.69)	7.82 (6.91)	6.51 (7.08)
Minutes Per Week (average)	22.20 (14.44)	34.84 (19.65)	23.74 (18.27)	26.76 (29.63)
<i>N</i>	431	430	552	539

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 388 minutes, over a total of 11 weeks for an average of 31 minutes per week. Usage metrics were generally highest for Grade 4 and in similar ranges among the other grades. Although Grade 6 students completed fewer overall lessons over fewer weeks than other grades, their usage was more concentrated, as indicated by their high average minutes per week usage. Additionally, lesson passage rates decreased as grade levels increased.

Distributionally, total usage figures (like total lessons, total minutes) were mostly positively skewed (which is evidenced in the large standard deviation values in Table 5, which are nearly as large as the mean for some measures). This means that a large number of treatment students had infrequent usage, with one-quarter of students completing fewer than 2 lessons or 56 minutes of activity (and averaging less than 18 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

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

	Treatment	Comparison
Fall i-Ready score	449.36	466.03
Spring SBA score	2489.26	2531.76
% proficient	47.0%	67.3%
<i>n</i>	1,952	3,378

Note: Means are unadjusted.

As noted previously, prior to any adjustments, comparison students scored higher on the fall i-Ready assessment than did treatment students. This difference may be related to characteristics of schools that decided to implement school-wide i-Ready Personalized Instruction with some of their students as opposed to using the Diagnostic-only program. Additionally, in partial instruction schools, i-Ready Personalized Instruction is purchased as a supplement for students who are underperforming and need extra support. Unadjusted spring SBA scores (and percent proficient) also tended to be higher, on average, for comparison students. 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 for treatment students (in schools with school-wide i-Ready Personalized Instruction) on math achievement, in relation to comparison (only assigned to Diagnostic testing, in schools with only some i-Ready Personalized Instruction). We first present results 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	10.869*	4.359	.013	0.109
Constant	2530.224	4.818		

Notes: 1. $N = 5,330$; j (schools) = 22. 2. Adjusted estimates with PSM weights and covariates. 3. * $p < .05$

There was a statistically significant positive effect of treatment on SBA math scores in relation to the comparison condition. The treatment (impact) estimate in Table 7 can be interpreted as the average difference between treatment and comparison students. For example, the regression estimate for treatment indicates that after adjusting for prior achievement and demographics, treatment students scored an average of 10.9 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 0.11 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 60-80th 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 4.4 percentile points higher (moving from the 50 to 54.4 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). Results from this analysis are presented in Table 8. Results are presented as odds ratios, which can be interpreted as the odds, or likelihood, of being proficient.

Table 8*Analyses of i-Ready Personalized Instruction on SBA math proficiency*

Outcome	Odds Ratio	Standard Error	p value
Treatment	1.522***	0.172	<.001
Constant	48.408	10.421	

Notes: 1. $N = 5,330$; j (schools) = 22. 2. Adjusted estimates with PSM weights and covariates. 3. *** $p < .001$

Mirroring the direction of the impacts on SBA scores, Table 8 shows that treatment had a significant positive impact on the proportion of students who were categorized as proficient on the SBA math assessment, based on their SBA score ($p <$

.001). Specifically, the treatment impact in Table 8 shows that treatment students had 1.5 times greater odds than comparison students of scoring proficient on the SBA math assessment in spring 2022.

The positive statistically significant impact for students on both outcomes, overall score and likelihood of achieving proficiency, and the large magnitude of the estimated improvement—almost 11 points—indicates that i-Ready Personalized Instruction consistently improved student achievement in math.

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 examined 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 race/ethnicity, gender, ELL status, or SPED status. However, we did find differences by grade level and school Title I eligibility.

Grade level. We first examined how the impact of treatment varied across grade levels. The results of this analysis are shown in Table 9.

Table 9

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

Grade	Estimate	Standard Error	p value	Effect size	n
3	6.009	8.653	.487	0.060	1,263
4	16.777**	5.352	.002	0.167	1,257
5	12.935	6.655	.052	0.129	1,389
6	7.689	6.017	.201	0.077	1,421

Notes: 1. $N = 5,330$; j (schools) = 22. 2. Adjusted estimates with PSM weights and covariates. 3. ** $p < .01$

Across the grade levels, results were most positive for Grades 4 and 5, although the effect of treatment only reached statistical significance in Grade 4 ($p < .01$). In Grade 4, students in the treatment condition averaged nearly 16.8 points higher on the SBA math assessment than their Grade 4 peers in the comparison condition. This improved result for Grade 4 students could reflect the greater usage by treatment

students in this grade (see Table 5) and supports the hypothesis that greater usage in the other grades might lead to greater achievement gains in those grades as well.

School Title I Eligibility Status. We also examined how students in Title I eligible schools were differentially impacted by treatment, compared to students not in Title I eligible schools. The results of this analysis are shown in Table 10 that presents the treatment impact for students in non-Title I eligible schools followed by the treatment impact for students in Title I eligible schools.

Table 10

Impacts of i-Ready Personalized Instruction on SBA math achievement, by Title I eligibility status

School Title I status	Estimate	Standard Error	p value	Effect size	n
Non-Title I eligible schools	-2.485	4.402	.572	-0.025	3,194
Title I eligible schools	17.145***	3.158	<.001	0.171	2,136

Notes: 1. $N = 5,330$; j (schools) = 22. 2. Adjusted estimates with PSM weights and covariates. 3. *** $p < .001$

Looking at the impact of treatment for each of these groups, there is a substantial difference in the treatment impact. Results were more positive for treatment students in Title I eligible schools than treatment students in non-Title I schools: i-Ready Personalized Instruction was significantly associated with a 17-point greater gain in math achievement ($p < .001$) over comparison students. In contrast, treatment students in non-Title I schools performed no differently than their comparison peers.

Referring back to Table 1, comparing the demographics of treatment and comparison students, a much larger percentage of treatment students were in Title I schools. Specifically, only 536 (17%) of comparison students were in Title I eligible schools, and likewise, only 720 (34%) of students were in non-Title I eligible schools. Originally, i-Ready Personalized Instruction was intended for Title I eligible schools only and so this is potentially an alternate way to examine the treatment effect. These results are thus encouraging for continued focused use of i-Ready Personalized Instruction in these schools. However, the results could also reflect some underlying differences about non-Title I schools included in the study (such as schools struggling with leadership, etc.).

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 11 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 11

Associations between i-Ready Personalized Instruction usage and math achievement

Usage Measure	Estimate	Standard Error
# of lessons	0.356***	0.048
# of unique lessons	0.413***	0.061
# of passed lessons	0.454***	0.064
Minutes of Usage (total)	0.012***	0.002
Weeks of Usage	0.506**	0.164

Note: $N = 5,330$; ** $p < .01$, *** $p < .001$.

All the usage variables were significantly positively associated with SBA math scores. Further, all but one of these variables were significant at the .001 alpha significance level. Coefficients in Table 11 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 0.012-point increase in SBA math score. Thus, usage for the average student, who completed 388 minutes of instruction, would be associated with a 4.7-point SBA math score increase. Similarly, the average user who completed 16 lessons would be expected to gain 5.7-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 9.1-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 on the basis of 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 12.

Table 12

Usage weeks categories ranges and sample sizes

Usage Category	Range	<i>n</i>	% of Treatment
1 – Low	0 – 9 weeks	1,067	54.7%
2 – Average	10 – 17 weeks	544	27.9%
3 – High	18 or more weeks	341	17.5%

Note. 3,378 students in comparison group with 0 weeks.

Looking at Table 12, 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 13 shows the unstandardized regression coefficients for each usage category, in relation to students with no usage (i.e., comparison students).

Table 13

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

Usage Category	Estimate	Standard Error
1 – Low	8.819	4.501
2 – Average	10.517	6.165
3 – High	18.730***	4.279

Note: $N = 5,330$; *** $p < .001$.

A significant positive relationship between usage and math achievement was observed for users in the highest category of i-Ready Personalized Instruction usage—students who used i-Ready for at least 18 weeks. The regression coefficients in Table 13 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 an 18.7-point increase in SBA math score, in relation to comparison students. A statistically significant positive relationship between usage category and math achievement was found only for this highest category of users. However, all of the remaining regression coefficients were positive, indicating that, even when an association did not reach statistical significance, students who used i-Ready Personalized Instruction outscored, on average, comparison students.

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 14 displays the percentages of i-Ready Personalized Instruction students who met i-Ready Personalized Instruction usage guidelines in each grade.

Table 14

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

Grade	Met i-Ready usage guidelines	<i>n</i>
3	9.7%	431
4	27.2%	430
5	8.7%	552
6	6.5%	539

Over all grades, 12.4% of students met both usage guidelines (36% meeting the first guideline of 30 minutes per week and 17% meeting the second guideline of 18 weeks). As shown in Table 4, about one-quarter of treatment students met usage guidelines in Grade 4, but percentages of students reaching usage guidelines were much lower in other 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 15.

Table 15

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

	Estimate	Standard Error	<i>n</i>
Did not meet usage guidelines	9.294*	4.700	1,710
Met usage guidelines	14.114**	4.149	242

Notes: 1. These estimates are in comparison to the 3,378 comparison students. 2. * $p < .05$, ** $p < .01$.

Achievement gains differed between treatment students who did and did not meet i-Ready Personalized Instruction usage guidelines in math, although both associations were significant and positive. Students who met i-Ready usage guidelines averaged 14.1-point greater gains on the SBA math exam than comparison students. Although the magnitude of the increase was lower for treatment students who did not meet usage guidelines, this association (estimating a 9.3-point gain over comparison students) was still 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 students in schools using i-Ready Personalized Instruction school-wide (Treatment students) and students who only received i-Ready Diagnostic assessments and who were in schools only partially using i-Ready Personalized Instruction (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 the school district 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 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 these usage amounts potentially contributed to the observed impacts of the program.

Math Achievement Gains

We found that treatment students had statistically significantly higher SBA math scores than comparison students. Specifically, students in the treatment group scored, on average, 10.9 points higher on the SBA math test than their peers in the comparison group. Supplementary analysis of students achieving proficiency in math on their SBA score showed that treatment also had a statistically significant impact on this outcome. Specifically, students in the treatment group had 1.5 times higher odds of scoring proficient (versus not scoring at least proficient) 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 (16.8 points) in SBA math achievement. At the school level, treatment students in Title I eligible schools also benefited more from treatment, compared to treatment students in non-Title I eligible schools and also compared to comparison students in Title I eligible schools.

Usage Patterns

Descriptive analysis of usage by grade revealed that usage metrics were generally highest in 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 all five usage metrics we considered were significantly positively related to student math achievement. Specifically, average usage (388 minutes of instruction) was associated with a 4.7-point greater math achievement on the SBA math assessment. When examining associations between categories of usage and achievement, a statistically significant positive association was observed between the highest category of usage (18 or more weeks) and SBA math scores. Related to this, students who met usage guidelines (of at least 18 weeks and 30 minutes per week usage) benefited more from their participation in the program compared to comparison students and users who did not meet usage guidelines.

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 Title I eligible schools, Grade 4, and among students who met the recommended usage guidelines. 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 (<i>n</i> , at assignment)	Post-test group (<i>n</i> , final analytic sample)	Attrited students (<i>n</i>)	Attrition %
Treatment	2219	1952	267	12.03%
Comparison	3430	3378	52	1.52%
Total	5649	5330	319	5.65%
			Differential attrition	10.52%