Supporting Middle School Math Learning With a Technology-Based Intervention

Impact, Moderators, and Usage

Mingyu Feng
Chunwei Huang
Kelly Collins

Working Paper
WestEd

September 2024

Suggested citation: Feng, M., Huang, C-W., & Collins, K. (2024). *Supporting middle school math learning with a technology-based intervention: Impact, moderators and usage*. Working Paper. WestEd.

Contents

Introduction

The issue of educational achievement gaps between demographic groups remains a persistent concern at the state and national level (Bohrnstedt et al., 2015). In the wake of the COVID-19 pandemic, these gaps widened as disruptions for K–12 students' education illuminated and exacerbated educational inequities in the United States (Bacher-Hicks, Goodman & Mulhern, 2020; Engzell, Frey, & Verhagen, 2020; Maldonado & de Witte, 2020; Stein, 2020; U.S. Department of Education, 2021). Achievement gaps in mathematics were particularly impacted and widened significantly (Kuhfeld et al., 2020).

As a promising tool for improving mathematics education and closing the achievement gap, the use of educational technology has expanded dramatically in K–12 education in recent years, accelerated by the pandemic. Yet the promise of improved student outcomes can only come to fruition if the educational technology programs being used are designed and implemented effectively. Evidence-based practices can help ensure that technology is built to be effective and used in a way that will most likely benefit students. One such evidence-backed program is ASSISTments.

Developed by researchers at Worcester Polytechnic Institute, ASSISTments (Heffernan & Heffernan, 2014) is a free online formative assessment platform that applies innovative technology in education. The ASSISTments intervention aligns with theory- and empirically based instructional practices of formative assessment (Heritage & Popham, 2013) and skill development (Koedinger, Booth & Klahr, 2013).

Evidence to support the effectiveness of ASSISTments in improving students' math outcomes has been, and continues to, grow as studies on the subject are persistently conducted. One particularly relevant study on the effectiveness of ASSISTments was conducted in Maine from 2013 to 2016 (Roschelle et al., 2016).

The Maine study was a randomized control trial in which 87 grade 7 mathematics teachers and their students (totaling 2,850 students in the final study sample) participated. Teachers, who were randomized at the school level, remained in their schools' assigned condition for 2 years. Student outcomes were measured at the end of the second year, and an analysis was conducted using a hierarchical linear regression model. The results showed that students who used ASSISTments performed statistically significantly better than did comparison students ($p < 0.001$, effect size Hedges' $g = 0.22$) (Murphy et al., 2020).

From 2018 to 2020, WestEd led a randomized controlled trial in North Carolina to determine the replicability of the Maine study findings (Roschelle et al., 2016) in a heterogeneous population that more closely resembled the demographic diversity of the United States as a whole. The North Carolina replication study was designed to examine the efficacy of ASSISTments for advancing grade 7 students' math learning.

Just like the Maine study, the North Carolina replication study was a randomized control trial experiment in grade 7 mathematics classrooms. Sixty-three participating schools were randomly assigned to either the intervention or comparison conditions. ASSISTments was implemented by all grade 7 mathematics teachers in intervention schools over 2 consecutive school years starting in fall 2018. Students who were in grade 7 during the 2019–20 school year comprised the analytic sample for the replication study.

Based on the study participants from the replication study, WestEd designed and conducted a follow-up study to measure the possible long-term impact of ASSISTments in grade 8 in 2021, one year after the intervention was over.

The following five research questions (RQs) motivated this follow-up study:

- RQ1: What is the long-term impact of ASSISTments on student math outcomes measured at the end of grade 8, one year after the completion of the intervention?
- RQ2: Do the long-term impacts of ASSISTments, if they exist, differ between students with different levels of prior math achievement?
- RQ3: Do the long-term effects of ASSISTments, if they exist, vary for students of different socioeconomic status, race/ethnicity, or other policy-relevant characteristics?
- RQ4: Do the long-term effects of ASSISTments, if they exist, vary for students in schools with different compositions?
- RQ5: To what extent does the usage of ASSISTments features measured in grade 7 relate to student math outcomes measured at the end of grade 8?

In a previously published research brief (Feng, Huang, & Collins, 2023), WestEd reported initial results that briefly addressed RQ1, RQ3, and RQ4. The results showed that ASSISTments demonstrated a significant long-term impact on student learning and increased student scores on the state standardized assessment as compared with a control group that continued with existing math practices one year after the implementation was finished. According to the preregistered preanalysis plan, RQ1 and RQ2 are *confirmatory* research questions, and RQ3, RQ4, and RQ5 are *exploratory* research questions.

This technical report extends the scope of the previously published research brief and further examines the differential impacts of ASSISTments on subgroups of student populations and the relationship between the usage of ASSISTments and students' long-term math outcomes. Analytical details of the analysis are provided for the primary confirmatory analysis and the exploratory analysis.

Study Overview

Setting and Participants

Sixty-three schools from 41 districts in North Carolina took part in the replication study, including 48 Title I schools and schools from rural, town, suburban, and urban communities. The schools taught various grade spans (i.e., K–8, 6–8, and 8–12). In all, 102 grade 7 mathematics teachers were enrolled in the study.

Study Design

During the replication study, schools were randomly assigned to either the intervention or the business-as-usual comparison condition in spring 2018. Then, all grade 7 math teachers in intervention-group schools implemented ASSISTments over 2 consecutive school years. During the "warm-up" first year (2018–19) intervention, teachers learned about ASSISTments and used the tool in their grade 7 classrooms. They then continued using ASSISTments during the second year (2019–20) with a new cohort of grade 7 students. This new cohort of grade 7 students served as the analytic sample for the replication study and the follow-up study.

Study-related supports for intervention-group schools ended at the conclusion of the 2019–20 school year. At that time, grade 7 math teachers at comparison schools were offered ASSISTments training. Because ASSISTments is a free program, there were no restrictions on the use of ASSISTments after the conclusion of the replication study. Any teacher at an intervention- or comparison-group school could use the program with their students, but no efforts were made by the research team or the developer to support teachers' use of ASSISTments after the 2019–20 school year.

The follow-up study that aimed to test the long-term impact of the ASSISTments platform examined the grade 8 performance of students from the analytic sample. These students maintained their original conditions, and their scores on the 2021 North Carolina End-of-Grade (EOG) state standardized mathematics test for grade 8 were used as the measure of grade 8 performance.

Intervention and Comparison Groups

ASSISTments (Heffernan & Heffernan, 2014) was introduced to the intervention group as an online tool for math homework. Many teachers also used the platform to support classwork. Developed by researchers at Worcester Polytechnic Institute, the ASSISTments platform aligns with the theoretically and empirically based instructional practices of formative assessment (Heritage & Popham, 2013). The tool provided students with instant feedback on their math work while teachers received real-time data and reports on students' learning progress. ASSISTments follows four steps of formative assessments:

- **Step 1:** Teachers create math assignments aligned with grade-level standards.
- **Step 2:** While completing assignments independently, students receive immediate feedback, hints, and explanations in ASSISTments to support their understanding and problem-solving.
- **Step 3:** Teachers obtain real-time assignment reports that summarize individual and class performance and common wrong answers, and they gain insights on students' needs for support.
- **Step 4:** Teachers analyze the data with the class and discuss common errors, using the information from the reports to inform their subsequent instruction and meet students' needs.

The ASSISTments platform contains textbook-based problems from the school's existing textbooks and curricula, along with other prebuilt content available for teachers to assign to students through the ASSISTments platform. The ASSISTments development team also added practice problems created by the intervention-group teachers to the platform upon their request.

Along with providing access to the ASSISTments platform, the ASSISTments developers provided intervention-group teachers with professional development and continuous support. Teachers attended a 2-day, in-person professional development session just before the start of each school year. A local coach provided ongoing support to teachers during the school year. The coach was available via phone and email, visited each intervention-group teacher at three points during the year, and provided personalized guidance and recommendations on the use of ASSISTments with students. Intervention-group teachers were expected to follow a specified use model to assign 20 to 30 minutes of ASSISTments assignments at least twice per week and were guided to review data reports of student work on assignments regularly.

Comparison-group teachers were asked to maintain their regular curriculum and their regular practices regarding homework and participating in professional development. They were also asked to continue with their existing use of technology tools but were given no access to ASSISTments or ASSISTments professional development.

Teachers in both the intervention and comparison groups followed the North Carolina Standard Course of Study (NCSCOS). They all participated in their district-provided professional development and continued with their existing curricula as usual.

A note about the broader study context

The North Carolina replication study was in its second year of implementation during the 2019–20 school year when the COVID-19 pandemic caused school closures in March 2020. Subsequently, the North Carolina EOG state standardized test was canceled for spring 2020. ASSISTments usage was interrupted due to school closures and then resumed in approximately half of the intervention-group teachers' classrooms during remote learning. The grade 7 state EOG test scores were missing for all students in the study during the 2019–20 school year.

Methods

Data Collection and Measures

The WestEd research team obtained demographic, school and course enrollment, and state assessment performance data for all students in the 63 study schools from the statewide database through the North Carolina Education Research Data Center (NCERDC).

Outcome and Baseline Measure

The outcome analysis followed the preregistered analysis plan and focused on grade 8 students in the participating schools who enrolled in regular grade 8 courses during the 2020–21 school year. The spring 2021 North Carolina EOG math assessment for grade 8 (EOG-MA08) was used as the primary outcome measure for examining ASSISTments' long-term effects on students' math learning. The North Carolina EOG math assessments are aligned with the Common Core State Standards for Mathematics (CCSSM) and have been used to measure students' proficiency since 2010. These assessments are administered annually to students in grades 3 through 8. Grade 6 mathematics test (EOG-MA06) scores from spring 2019 served as the baseline measure (Table 1).

Table 1. Overview of Student Outcome and Baseline Measures for the Study

School year	Grade	(Intended) measurement
2018–19	6	EOG-MA06 as baseline
2019–20	7	EOG-MA07 as an immediate outcome (missing due to the COVID-19 pandemic)
2020–21	8	EOG-MA08 as primary, long-term outcome

In addition, to study the differential treatment impact (moderator analysis as described in the discussion of RQ2 in the Results and Findings section), the research team split students into two performance-level groups according to the median of their EOG-MA06 scale scores: low-performing (below median) and high-performing (median and above). Note that the median is the same cutoff point used by North Carolina to determine students' grade 6 proficiency level (North Carolina Department of Public Instruction [NCDPI], 2022). Based on the state's guidance, we also categorized students into two groups according to their achievement levels: not ready for college (Level 3 and below) and ready for college (Levels 4 and 5) (NCDPI, 2022).

Demographics

Individual demographic data included gender, race/ethnicity, economically disadvantaged status (EDS), and disability status. School-level background data included average EOG-MA06 scores, grade 7 enrollment size, Title 1 eligibility, percentage of students with EDS, and percentage of ethnic groups (Black, Hispanic, and White). These variables served primarily as covariates in the impact analyses. Individual demographic data and school-level background data were also used to study the differential treatment impact.

Course Enrollment

Course enrollment data included the math courses that students enrolled in each semester during grade 7 and grade 8. Although most students took the EOG-MA08, some students enrolled in North Carolina's high school level courses and took the End-of-Course (EOC) math assessments, the NC Math 1 (EOC-MA1) and the NC Math 3 (EOC-MA3). The EOC assessments are used to estimate student's mastery of the material in a particular content area as specified in the NCSCOS, which defines the appropriate content standards for each grade or proficiency level and each high school course. This smaller group of students was not included in the preregistered analysis plan, and their data were analyzed separately and are not included in this report.

Intervention Usage Indicators

The ASSISTments platform automatically captures student actions, inputs, and teacher usage in real time with a time stamp. The research team collected the computer records for grade 7 (2019–20) ASSISTments usage from the ASSISTments team and constructed a series of usage metrics at the student level and the class level. Student-level metrics captured exposure and dosage. Class-level metrics summarized teachers' actions, including assigning math work for classes and accessing various types of data reports generated by the platform. Table 2 shows the student-level and class-level ASSISTments usage indicators that were examined in the subsequent exploratory analysis to address RQ5.

Table 2. Student-Level and Class-Level Metrics

Level	Metrics	Measure of . . .
Student	[a] Total number of minutes students worked in ASSISTments [b] Total number of problems completed [c] Total number of assignments completed	Dosage and students' commitment to math work in the system
Student	[d] Average number of days per week a student worked in ASSISTments [e] Average number of minutes per week	Teachers' compliance with the expected use model
Class	[f] The rate at which teachers reviewed the corresponding student performance reports at least once	Teachers' compliance with the expected use model
Class	[g] Total number of weeks in which the teacher assigned at least one ASSISTments assignment to the class [h] Total number of assignments assigned to the class over the school year	Student's exposure to the ASSISTments intervention in general
Class	[i] Percentage of assignments that were from textbooks or open educational resources (OER) (not from Skill Builders)	Students' exposure to mastery-based problem sets

Note: For metric [f], the report reviewing rate is calculated by dividing the number of assignments for which a teacher opened the corresponding student performance report at least once by the total number of assignments the teacher assigned to a class.

To examine the association between student math achievement and the amount of intervention students received on their math work during independent problem-solving practices, the research team utilized student-level indicators of [a] the total number of minutes students worked in ASSISTments, [b] the total number of problems completed, and [c] the total number of assignments completed as a measure of dosage and students' commitment to math work in the system.

The class-level indicator of [f] the rate at which teachers reviewed the student performance reports were used in conjunction with two student-level indicators that were aggregated to the class-level—[d] the average number of days and [e] the average number of minutes per week students worked in ASSISTments—as proxy measures of teachers' compliance with the research team's expectation for the use of ASSISTments and its features to support instruction.

We considered the class-level indicators of [g] the total number of weeks in which assignments were made in ASSISTments and [h] the total number of assignments the teacher made for a class as measures of student's exposure to the ASSISTments intervention in general.

The [i] percentage of assignments that were regular assignments was included to measure the extent to which teachers assigned regular problem sets versus the system's Skill Builder problem library. This allowed the team to analyze whether higher exposure to mastery-based problem sets is associated with higher math achievement scores. During the original ASSISTments studies in Maine, there was a significant correlation between the number of Skill Builders assigned and student learning (Roschelle et al., 2016; Murphy et al., 2020). We were interested in seeing if the findings were replicated in the North Carolina study.

We compiled measures of teachers' use of ASSISTments (e.g., reviewing reports, making assignments) by aggregating data from the platform's user log files at the class level. Many teachers in the intervention schools taught multiple classes (about four classes per teacher on average). There was variation in how teachers between and within schools used ASSISTments and how they used the platform across different classes. Measures of students' use were calculated for assignments that teachers assigned for specific enrolled classes. Consequently, student-level metrics are nested within classes.

Post-Study Usage Indicators

The research team attempted to pull grade 8 ASSISTments usage data for students who participated in the study during grade 7, regardless of the original experimental group. Since teachers and students accessed the ASSISTments platform via their Google Classroom or Canvas accounts, it was not possible to systematically collect school identification based on user account data. In addition, teachers and students were not required to enter their school information directly into their ASSISTments account profiles. Moreover, grade 8 teacher users, who were not study participants, did not consistently provide state, town, district, school, or grade information in their accounts or the ASSISTments classes they created. For these reasons, it was difficult to reliably and systematically identify the grade 8 teachers and students in the participating schools during the 2020–21 school year, which in turn prevented the team from reliably identifying and collecting post-intervention usage data from the grade 7 students who participated in the study the previous year and entered grade 8 in 2020–21.[1] Therefore, no usage data in grade 8 was included for the subsequent analysis.

Data Analysis Approach

An intent-to-treat (ITT) approach was used to conduct the impact analysis. A two-level hierarchical linear regression model (HLM) was used to account for the effect of clustering due to the nesting structure of data. In this study, students were nested within schools.

For RQ1, we modeled the EOG-MA08 for students in the intervention- and comparison-group schools, controlling for the students' EOG-MA06 test scores and other student- and school-level covariates as described in the previous section, "Data Collection and Measures." The HLM takes the following form:

$$\mathrm{Math}_{ij} = a0 + b_1\mathrm{PreMath}_{ij} + b_2\mathrm{Tx}_j + \Sigma b_j S_j + \Sigma b_i I_{ij} + \mu_j + e_{ij},$$

where subscripts *i* and *j* denote student and school, respectively; Math represents student achievement in math (scale score from the EOG-MA08 in spring 2021); PreMath represents the baseline measure of math performance (scale score from the EOG-MA06 in spring 2019); Tx is a dichotomous variable indicating student enrollment in a school that was assigned to either the intervention group (Tx = 1) or the comparison group (Tx = 0); and I and S are vectors of student-level covariates and school-level covariates, respectively, measured prior to exposure to the intervention. Lastly, μ_j and e_{ij} represent the random effect of school and student, respectively. In this model, the long-term impact of ASSISTments is represented by b_2.

For RQ2 and RQ3, a cross-level interaction term of Tx by the subgroup variable (a binary variable) was added to the model to examine whether the intervention has a differential impact. These subgroup variables included the following: performance level (low-performing versus high-performing) at the baseline; college-readiness status (not ready for college versus ready for college) at the baseline; gender (female versus male); Black versus non-Black; Hispanic versus non-Hispanic; White versus non-White; economically disadvantaged status; and disability status.

For RQ4, a school-level interaction term of Tx by the school-level characteristic (a binary variable) was added to the model to examine the intervention's differential impact. In addition

[1] The research team identified some teachers from selected North Carolina districts that included at least one participating school based on the teachers' registered district email addresses. The team examined the usage data in these teachers' classes during the 2020–21 school year. In the limited data sample, no reliable evidence that students used ASSISTments consistently in grade 8 was found.

to the baseline measure of math performance, the percentage of White students and the percentage of economically disadvantaged students were considered. The median was used to determine the low versus high percentage for each variable of interest.

For RQ5, one-level and two-level regression models were used to study the association between student usage data and outcome. The two-level model is similar to the equation shown above, controlling for the same set of covariates as those in the impact analysis. The one-level model was used to estimate the effect size of the association (i.e., partial eta squared) for interpretation (Richardson, 2011). Given that the majority of variance in the outcome was explained at the student level, it was appropriate to use the one-level model. For class-level usage data, a three-level model (student, class, school) was used. To do that, students who were enrolled in more than one class (678 students, 28.6%) were excluded. Students enrolled in very small classes (no more than 5 students) were also excluded.

The missing-indicator method (White & Thompson, 2005) was used to account for missing values on the covariates (not the outcome variables) in the impact analysis models. The missing-indicator method retains all observations with missing values on covariates in the analysis. Indicator variables were created for missing values on each variable (0 = observed, 1 = missing), and missing values on the covariates were coded to a constant. Both the recoded covariates and the missing value indicator variables were included in the regression model. In a randomized controlled trial, in which randomization helps ensure that the baseline covariates are balanced, the missing-indicator method appears to refine the precision of impact estimates and standard errors (White & Thompson, 2005).

Observations with missing values on outcome variables were excluded from the impact analyses. Deletion of observations with missing outcome variables has been shown to result in accurate impact estimates and standard errors when outcomes are missing at random, conditional on the covariates (Allison, 2001; von Hippel, 2007).

Results and Findings

Study Sample

Students from all 63 original schools (32 intervention-group schools and 31 comparison-group schools) were included in the study sample if both criteria were met: (1) they had grade 6 EOG scores from the 2018–19 school year when the schools were randomly assigned to conditions and (2) they were enrolled in one of the 63 participating schools as a grade 6 student during the 2018–19 school year or as a grade 7 student during the 2019–20 school year if the school did not serve grade 6. In total, the schools included 9,073 students (4,495 in intervention-group schools and 4,578 in comparison-group schools) as the baseline reference sample. Table 3 shows the school-level characteristics of the 63 schools in the study.

Table 3. School-Level Characteristics of All Participating Schools

Characteristic	Median	Mean	Min	Max
Grade 7 enrollment	183	207	27	430
Female	48.48%	48.63%	36.43%	59.55%
Black	12.68%	16.89%	0.00%	82.40%
Hispanic	13.08%	15.46%	2.46%	60.00%
White	64.00%	60.52%	1.09%	94.59%
EDS	58.80%	53.77%	5.00%	95.00%
ELL	3.57%	3.80%	0.00%	14.44%

Note: EDS refers to economically disadvantaged status; ELL refers to English language learners.

The final analytic sample for the EOG-MA08 analysis (EOG Sample) consisted of 5,991 students who took the EOG-MA08 test, including 2,961 students in the intervention group and 3,030 students in the comparison group. The student-level demographic characteristics of the EOG Sample are shown in Table 4.

Table 4. Student Demographics of the EOG Sample

Demographic	Intervention	Comparison	Both conditions
Female	48.63%	49.54%	49.09%
Black	19.15%	20.26%	19.71%

Demographic	Intervention	Comparison	Both conditions
Hispanic	22.15%	21.55%	21.85%
White	50.69%	51.35%	51.03%
EDS	63.02%	59.77%	61.38%
ELL	5.34%	6.34%	5.84%

A group of 1,962 students, 936 in the intervention group and 1,026 in the comparison group (EOC Sample), took the End-of-Course Math 1 (EOC-MA1) test and were analyzed separately. These students typically performed at the top 25 percent and enrolled in high school–level courses during grade 8. Analysis of this sample and outcome are not included in this report.

Sample Attrition

Given the study's clustered randomized trial design, the sample attrition was examined at both the cluster and individual levels. The EOG Sample included data from all participating schools. Therefore, the cluster-level attrition (for schools) is 0 percent.

We tried two approaches to estimate individual-level attrition. First, we estimated the individual attrition rates based on the availability of EOG-MA08 test results (Table 5). The sample attrition would be considered low considering the overall attrition (33.97%) and differential attrition (0.32%) per What Works Clearinghouse (WWC) (2002) guidance.

Table 5. Number of Students in the Reference Sample Versus the Number of Students With EOG-MA08 Scores

Group	# of students in the baseline reference sample	# of students with MA08 scores	Individual-level data attrition
Tx	4,495	2,961	34.13%
Cx	4,578	3,030	33.81%

Group	# of students in the baseline reference sample	# of students with MA08 scores	Individual-level data attrition
Total	9,073	5,991	33.97%
Differential attrition	n/a	n/a	0.32%

Considering that a significant portion of grade 8 students enrolled in high school–level courses and took the NC EOC tests instead of the EOG test, we tried an alternative approach to estimate individual attrition rates, using different reference samples. In this approach, we estimated the reference sample for each condition based on the grade 8 course enrollment data provided by the NCERDC and calculated individual attrition rates accordingly. We assumed that a grade 8 student was expected to take the state's EOG-MA08 assessment only if they enrolled in a grade-level math course (such as math grade 8, math grade 8 advanced/accelerated). Overall, 69.59 percent of grade 8 students in the intervention-group schools and 69.86 percent of students in the comparison-group schools enrolled in a grade-level math course. The difference between the two groups was not statistically significant.

Table 6 lists the number of students who were expected to take the EOG-MA08 test versus those whose scores were available in the data file. Based on this information, we calculated the individual overall and differential attrition rates to be 5.30% and 0.09%, respectively. Based on the WWC guidance, the individual-level attrition would also be considered low.

Table 6. Number of Students Expected to Take the EOG-MA08 Test Versus the Number of Students Whose Scores Were Available

Group	# of students expected to take MA08	# of students whose scores were available	Individual-level data attrition rates
Tx	3,128	2,961	5.34%
Cx	3,198	3,030	5.25%
Total	6,326	5,991	5.30%
Differential attrition	n/a	n/a	0.09%

Baseline Equivalence

We tested the baseline equivalence between intervention and comparison groups using EOG-MA06 test scores (Table 7). For those who took the EOG-MA08 in 2021, the baseline difference was small and not statistically significant at the 0.05 level (Hedges' g = 0.08), indicating that both groups performed equally at the end of grade 6 in 2019 prior to participating in the study.

Table 7. Baseline Equivalence Based on EOG-MA06 Test Scores in 2019

Tx (mean)	Tx (SD)	Tx (*n*)	Cx (mean)	Cx (SD)	Cx (*n*)	Diff	SE	*p*	Hedges' *g*
545.70	8.09	2,961	545.03	7.98	3,030	0.66	0.452	0.143	0.08

RQ1: ASSISTments had a significant positive impact on student learning in the long term.

The ITT analysis showed that students in the intervention-group schools assigned to use ASSISTments in grade 7 scored 0.80 points higher than did the comparison-group students on the EOG-MA08 test (Table 8, Figure 1). The estimated mean score on the EOG-MA08 for the intervention group is 535.13 (standard deviation = 8.494) versus 534.33 (standard deviation = 7.715) for the comparison group. This treatment impact is statistically significant at the 0.05 level (p = 0.011) and corresponds to an effect size of 0.10 (Hedges' g). The result suggests that implementing ASSISTments had a significant positive impact on student learning in math in the long term because the effect was sustained after 1 year of program implementation, which ended at the end of grade 7.

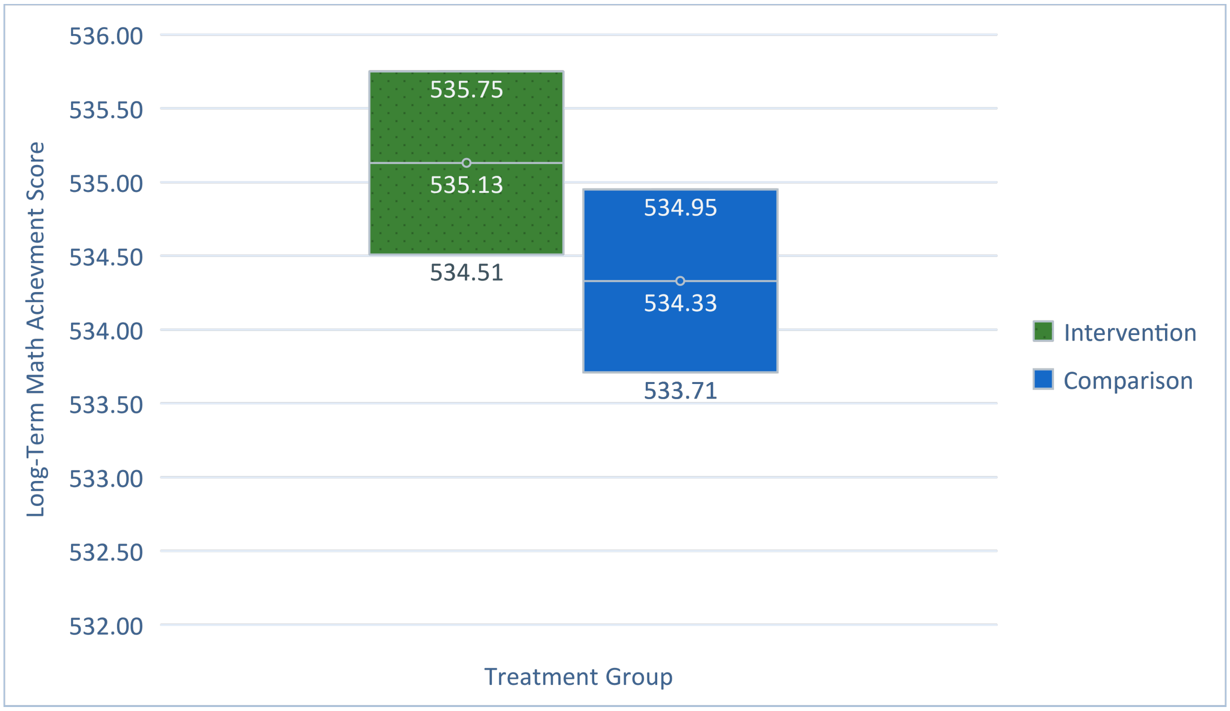
One might wonder what the 0.80-point score difference or the 0.10 effect size means. The research team used several approaches to contextualize the difference and explain the effect size, building on the guidance of Lipsey et al., (2012). We referred to North Carolina's statewide EOG scale score ranges for different achievement levels; score differences between ethnic groups; historic testing trends; and results from similar studies and interventions. We concluded that the effect size of the current study for the long-term impact is practically meaningful. The details of the approaches we used to interpret the effect can be found in the previously released research brief (Feng, Huang, & Collins, 2023, pp. 12–16).

Furthermore, we examined the group difference in terms of the percentage of students being proficient. The results indicate that 21.45 percent of students in the intervention group scored at the proficient level or above compared with 15.84 percent of students in the comparison group (the difference of 5.61% is significant at the 0.05 level, p = 0.018). This suggests that the estimated mean score difference of 0.80 points may not actually be small because it is reflected in the performance-level change.

Table 8. Impact Estimates Based on EOG-MA08 Test Scores in 2021

Tx (adjusted mean)	Cx (adjusted mean)	Diff	SE	*p*	Hedges' *g*	Tx (*n*)	Tx (SD)	Cx (*n*)	Cx (SD)
535.13	534.33	**0.80**	0.315	**0.011**	0.10	2,961	8.494	3,030	7.715

Figure 1. Estimated Impact of ASSISTments on EOG-MA08 Test Scores



RQ2: There was no differential impact by experimental condition between low-performing and high-performing students.

Roschelle et al. (2016) examined the differential treatment effects of ASSISTments by performing a median split of the study sample and found a significant difference in ASSISTments' treatment effects in lower-performing versus higher-performing students. The interaction was statistically significant: $t(2,770) = 2.432$, and $p = 0.015$. We followed the same approach and grouped students in the EOG Sample into two groups, low-performing versus high-performing, based on the median of EOG-MA06 test scores (546) for students who took the EOG-MA08. Though the difference between the intervention and comparison groups in each subgroup is positive (0.26 and 1.36, respectively), the contrast of these two differences

(–1.10) is not statistically significant at the 0.05 level (but is on the border, $p = 0.05$) and has an effect size of –0.14. This indicates that there was no differential treatment impact by performance level within the EOG Sample (Table 9).

Table 9. Differential Treatment Impact by Prior Performance Level on the EOG-MA08

Level	Inter-vention (adjusted mean)	Tx (*n*)	Compar-ison (adjusted mean)	Cx (*n*)	Mean diff	Contrast (inter-action)	SE	*p*	Hedges' *g*
Low-performing	534.57	1,426	534.31	1,517	0.26	—	—	—	—
High-performing	535.67	1,535	534.31	1,513	1.36	–1.10	0.563	0.050	–0.14

The results on the differential treatment impact by prior performance level should be interpreted with the caveat that the low-performing and high-performing subgroups were determined *within* the EOG Sample as opposed to the full grade 8 sample in participating schools. As noted above, the EOG Sample included only regular students who took the EOG-MA08 test and did not include the top-performing students in grade 8 in the participating schools who enrolled in high school–level courses and thus took EOC tests. We examined the grade 6 performance of students in the EOG Sample versus those in the EOC sample. The data showed that the EOG-MA06 mean was 545.36 (standard deviation = 8.03) for students in the EOG sample and 558.01 (standard deviation = 6.53) for those in the EOC sample, and the difference (12.65) was statistically significant ($p < 0.001$).

Similarly, there was no differential treatment impact between not-college-ready and college-ready students. The contrast was –0.87 with Hedges' *g* of –0.11 ($p = 0.211$).

RQ3: ASSISTments helped close achievement gaps, especially for Hispanic students.

We also conducted a series of analyses to study the differential impacts by subgroup based on student characteristics (Table 10). Using EOG-MA08 test scores as the outcome measure, the analyses indicated that

- the intervention benefited Students of Color[2] significantly more than it benefited White students (p = 0.003, Hedges' g = –0.14),
- implementing ASSISTments benefited Hispanic students more than it did non-Hispanic students (p = 0.014, Hedges' g = 0.13), and
- no differential treatment impacts were found among other subgroups.

[2] The group of Students of Color included all students who were not identified as "White" in the state database.

Table 10. Differential Impact Estimates by Student-Level Characteristics

Category	Tx (adjusted mean)	Tx (*n*)	Cx (adjusted mean)	Cx (*n*)	Mean diff	Contrast (inter-action)	SE	*p*	Hedges' *g*	Tx (*n*)	Tx (SD)	Cx (*n*)	Cx (SD)
White	534.53	1,501	534.26	1,556	0.27	—	—	—	—	—	—	—	—
Students of Color	535.76	1,460	534.39	1,474	1.37	–1.10	0.372	0.003	–0.14	2,961	8.494	3,030	7.715
Hispanic	535.63	656	534.02	653	1.61	—	—	—	—	—	—	—	—
Non-Hispanic	534.99	2,305	534.40	2,377	0.59	1.02	0.412	0.014	0.13	2,961	8.494	3,030	7.715
Black	534.27	567	533.37	614	0.90	—	—	—	—	—	—	—	—
Non-Black	535.34	2,394	534.56	2,416	0.78	0.12	0.380	0.732	0.01	2,961	8.494	3,030	7.715
Female	534.99	1,440	534.31	1,501	0.68	—	—	—	—	—	—	—	—
Male	535.25	1,521	534.35	1,529	0.90	–0.22	0.313	0.486	–0.03	2,961	8.494	3,030	7.715
EDS	534.62	1,866	533.83	1,811	0.79	—	—	—	—	—	—	—	—
Non-EDS	535.93	1,095	535.12	1,219	0.81	–0.02	0.351	0.948	0.00	2,961	8.494	3,030	7.715
DS	534.47	428	534.05	509	0.42	—	—	—	—	—	—	—	—
Non-DS	535.23	2,505	534.38	2,492	0.85	–0.43	0.527	0.412	–0.05	2,933	8.511	3,001	7.714

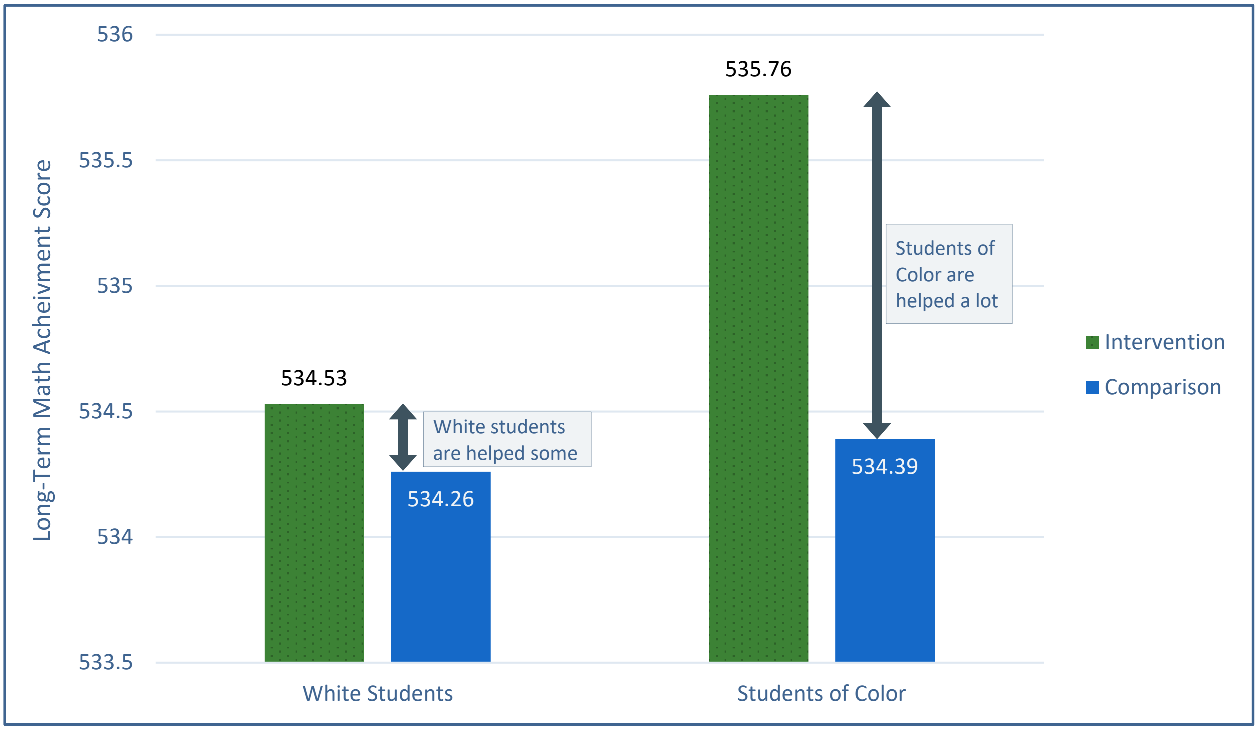
Note: EDS refers to economically disadvantaged status; DS refers to disability status.

The finding that the intervention improved math learning for Hispanic students more than it did for other groups of students is not typically seen in other studies. At the baseline in the intervention and comparison groups, Students of Color scored 2 points lower than White students scored on the EOG-MA06 (Table 11). At the end of grade 8, both Students of Color and White students in the intervention group scored higher than their counterparts did in the comparison group. However, Students of Color in the intervention group scored significantly higher than Students of Color in the comparison group did, whereas there was no significant difference between the intervention and comparison groups for White students (Figure 2a). The results suggest that traditionally underrepresented groups (in particular, Hispanic students, Figure 2b) benefited more from the ASSISTments program than did White students. Therefore, the intervention helped narrow the achievement gap at least among Hispanic students.

Table 11. Baseline Difference of Average Score on the EOG-MA06: White Students and Students of Color

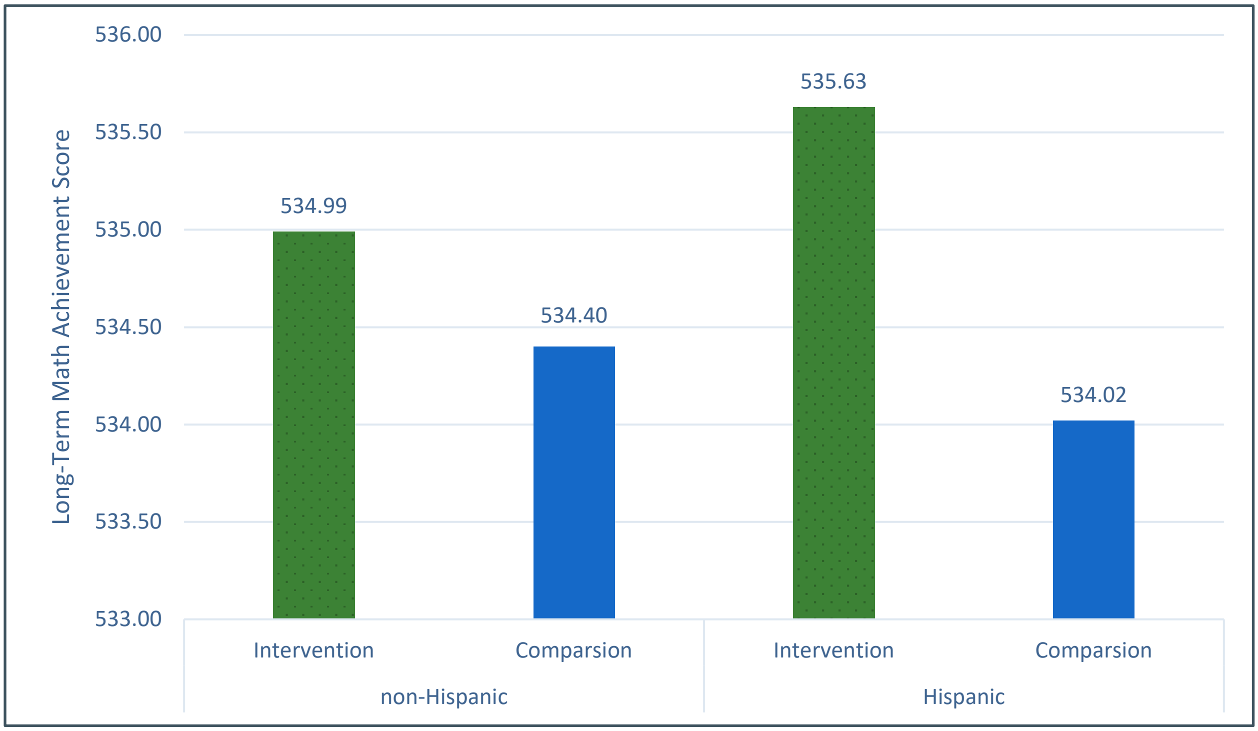
Student group	Intervention (mean score)	Intervention (*n*)	Comparison (mean score)	Comparison (*n*)
White students	546.56	1,501	546.27	1,556
Students of Color	544.72	1,465	543.79	1,474
Difference	1.84	n/a	2.48	n/a

Figure 2a. Average EOG-MA08 Score by Experimental Condition and Ethnicity (White Students Versus Students of Color)



Note: Both White students and Students of Color in the intervention group scored higher than their counterparts did in the comparison group. However, Students of Color in the intervention group scored significantly higher than Students of Color in the comparison group did, whereas there was no significant difference between the intervention and comparison groups for White students.

Figure 2b. Average EOG-MA08 Score by Experimental Condition and Ethnicity (non-Hispanic Versus Hispanic)



RQ4: ASSISTments had more benefits for students attending schools with a higher percentage of economically disadvantaged students and a lower percentages of White students.

We also examined the differential impacts of school-level backgrounds to see whether the effects of ASSISTments vary based on school composition of student populations. The results suggested that students in intervention-group schools with a higher percentage of students from economically disadvantaged communities benefitted more than did those in intervention-group schools with a lower percentage of students from economically disadvantaged communities (p = 0.045, Hedges' g = 0.22). Furthermore, students in the intervention-group schools with a lower percentage of White students benefitted more than did those in the intervention-group schools with a higher percentage of White students (p = 0.056, Hedges' g = 0.19). Studies have shown that school districts that serve more Black and Latino students and students from economically disadvantaged communities on average receive less funding than their whiter and wealthier counterparts (Morgan, 2022). The results demonstrated the potential positive impacts of the ASSISTments program in providing needed support to students and teachers in schools where access to educational resources might be limited.

These findings are similar to those for RQ3, which suggest that ASSISTments benefited Hispanic students more than those of other ethnic groups. Schools with fewer White students or more students from economically disadvantaged communities tended to have more Hispanic students (the correlation between the percentage of White students and the percentage of Hispanic students was –0.41; the correlation between the percentage of students from economically disadvantaged communities and the percentage of Hispanic students was 0.53).

RQ 5: Certain intervention features were associated with student learning outcomes.

Teacher and Student Use of ASSISTments

Teacher and student usage of ASSISTments was tracked during the 2019–20 school year. Table 12 shows the descriptive statistics of all ASSISTments usage variables.

On average, intervention teachers made 51 assignments in ASSISTments for each class in 17 weeks, approximiately 3 assignments per week. We also analyzed the proportion of ASSISTments reports that a teacher opened by clicking on a link either in an email sent to the teacher or in the ASSISTments platform. We used the opening of reports as an indicator that a teacher was using ASSISTments to review student math work performance, which is a precondition to using data to adjust instruction. The data showed that teachers on average reviewed performance reports generated by the system for 47 percent of the assignments they made.

On average, students in the intervention group spent 337 minutes (standard deviation = 290 minutes) in ASSISTments over 28 days and completed 233 problems. Students used ASSISTments for an average of 19 minutes over 1.62 days per week, slightly less than the study leaders' expectations of 20 to 30 minutes total over 2 days per week.

We found that on average 76 percent of the assignments made by teachers for a class were from their textbooks or prebuilt OER resources within ASSISTments and the rest were Skill Builder problems. Though Skill Builders were on average utilized for a lower percentage of assignments (24%), more than half of the intervention-group teachers assigned Skill Builders and the use of Skill Builders varied among teachers (standard deviation = 0.18).

Table 12. Descriptive Statistics of Usage Variables

Variable	Mean (SD)	Range	Percentile 20	Percentile 40	Percentile 50	Percentile 60	Percentile 80
STUDENT LEVEL **[a] Total minutes**	337.04 (290.34)	0–1835.63	109.08	195.21	254.07	317.65	528.51
STUDENT LEVEL **[b] Total problems completed**	233.41 (218.95)	0–1251	66	135	172	222	341
STUDENT LEVEL **[c] Total assignments completed**	25.46 (20.75)	0–104	7	16	21	27	40
STUDENT LEVEL **[d] Average days per week**	1.62 (0.47)	0–3.55	1.25	1.46	1.55	1.67	1.95
STUDENT LEVEL **[e] Average minutes per week**	19.22 (11.30)	0–164.08	10.83	15.10	17.14	19.17	25.66
CLASS LEVEL **[f] Reports reviewing rate**	0.47 (0.24)	0–1	0.26	0.38	0.49	0.54	0.71
CLASS LEVEL **[g] Total number of weeks with assignments**	17.44 (7.30)	1–32	10	16	18	19	23
CLASS LEVEL **[h] Total assignments assigned**	50.58 (28.34)	1–144	28	38	45	49	76
CLASS LEVEL **[i] Proportion of assignments from non–Skill Builders**	0.76 (0.18)	0.24–1	0.64	0.73	0.78	0.82	0.96

Note: The sample size used to compute the student-level usage data is 2,374; the sample size used to compute the class-level usage data is 1,696. As explained in the “Data Analysis Approach” section, some students were excluded from the class-level usage data analysis.

ASSISTments usage during the replication study in North Carolina was relatively lower than the reported usage during the original ASSISTments efficacy study in Maine. In the Maine study, on average the students worked on 967 mathematics problems and spent 14 hours in ASSISTments across a full school year (Roschelle et al., 2016) and on average the teachers opened 73 percent of the performance reports (Murphy et al., 2020).

There is reason to suspect that pandemic-related school closures in spring 2020 may have contributed to the usage differences between the original study and the replication study. Tables 13 and 14 shows class-level usage and student-level usage metrics before and after schools ended in-person learning (the cutoff date was March 15, 2020). There were approximately 26 school weeks of in-person learning and 10 weeks of remote learning.

As shown in Table 13, the number of active intervention-group teachers dropped from 64 to 36. Forty-four percent of the teachers stopped using ASSISTments during remote learning for various reasons (e.g., school/district policy regarding which supplemental programs could be used, students lacked home access to technology). However, the teachers who continued using ASSISTments during the pandemic continued making assignments in ASSISTments and reviewed student performance reports more frequently than they did before the pandmic. Report reviewing rate increased from 46 percent to 62 percent. This suggests that during remote learning, teachers relied more on the ASSISTments reports to understand students' progress and challenges. This was confirmed by teachers' responses in the post-intervention survey.

Table 13. Descriptive Statistics of Usage Variables Before and After Remote Learning (Teacher Usage)

Teacher-level usage variable	All of 2019–20 (*n*)	All of 2019–20 (mean)	All of 2019–20 (SD)	Pre-COVID/ in person (*n*)	Pre-COVID/ in person (mean)	Pre-COVID/ in person (SD)	During COVID/ remote (*n*)	During COVID/ remote (Mean)	During COVID/ remote (SD)
Total number of assignments assigned	64	126.08	109.60	64	104.62	90.47	36	36.83	36.34
Report-reviewing rate	64	0.49	0.25	64	0.46	0.41	36	0.62	0.33

On the other hand, student usage and commitment decreased overall after schools transitioned to remote learning. As shown in Table 14, the number of days students worked in ASSISTments per week dropped from 1.54 to 0.80 and the amount of time they spent dropped from 17.92 to 11.53 minutes. Students were less committed to ASSISTments work. They started only 45 percent of the assignments their teachers made during remote learning compared with 63 percent before the pandemic, and they completed only 38 percent of the assignments they were assigned.

Table 14. Descriptive Statistics of Usage Variables Before and After Remote Learning (Student Usage)

Student-level usage variable	All of 2019–20 (mean)	All of 2019–20 (SD)	Pre-COVID/ in person (mean)	Pre-COVID/ in person (SD)	During COVID/ remote (mean)	During COVID/ remote (SD)
Total number of problems completed by students	206.66	224.39	179.84	108.35	28.24	48.85
Proportion of assignments started by students out of the total assignments they were assigned	0.63	0.26	0.63	0.27	0.45	0.38
Average proportion of assignments completed by students out of the total assignments they were assigned	0.50	0.29	0.49	0.29	0.38	0.38
Average number of days per week students worked in ASSISTments	1.60	0.49	1.54	0.60	0.80	0.88
Average total minutes per week students worked in ASSISTments	19.09	12.40	17.92	12.18	11.53	17.79

Relationship Between ASSISTments Use and Student Achievement

To address RQ5, we examined how student's EOG math achievement varied with ASSISTments use within the intervention-group schools. Table 15 summarizes the association between the usage variables—one set at the student level and one set at the class level—and student learning outcomes measured by the EOG-MA08. Given a large sample size at the student level, we used $p = 0.01$ instead of 0.05 to determine whether the association between student-level usage and math outcome was statistically significant. In addition, when applicable (one-level model for student-level usage data), we used the partial eta squared to determine the effect size of the association.

Except for one usage variable (the average number of days per week students worked in ASSISTments for this class), generally the analysis showed a positive relationship between student-usage variables and the EOG-MA08. At the class level, only one variable was found to be positively associated with student math achievement: the percentage of assignments that were non–Skill Builder assignments. Note that these analyses are exploratory and cannot be used to make definitive claims about causal relationships between different types and levels of use of ASSISTments and student math achievement.

While the association for the student-level usage data tended to be statistically significant, the proportion of variance on the EOG-MA08 explained by any usage variable is small compared with other variables, especially prior math achievement. However, the [a] total number of minutes students worked on ASSISTments had the highest proportion of 1.3 percent (partial eta squared = 0.013), followed by the [e] average number of minutes per week students worked on assignments with a proportion of 1.0 percent (partial eta squared = 0.010). These results suggest that in contrast with other usage variables (such as the number of problems completed or the number of days worked in ASSISTments), the amount of time that students are exposed to the intervention or work on assignments might play a more pivotal role than other factors in learning math.

Table 15. The Association Between the Usage Data and Student Math Scores on the EOG-MA08

Usage variable	One-level model (coefficient)	One-level model (*p*)	One-level model (partial eta squared)	Two- or three-level model (coefficient)	Two- or three-level model (*p*)
STUDENT LEVEL **[a] Total number of minutes students worked in ASSISTments for this class**	0.003	**<0.001***	0.013	0.003	**0.002***
STUDENT LEVEL **[b] Total number of problems completed**	0.002	**<0.001***	0.007	0.005	**0.013***
STUDENT LEVEL **[c] Total number of assignments completed**	0.022	**0.001***	0.005	0.053	**<0.001***
STUDENT LEVEL **[d] Average number of days per week students worked in ASSISTments**	0.130	0.624	0.000	1.155	**0.002***
STUDENT LEVEL **[e] Average number of minutes per week students worked on assignments**	0.055	**<0.001***	0.010	0.067	**0.001***
CLASS LEVEL **[f] Rate at which the teacher reviewed reports for assignments**	n/a	n/a	n/a	1.650	0.065

Usage variable	One-level model (coefficient)	One-level model (*p*)	One-level model (partial eta squared)	Two- or three-level model (coefficient)	Two- or three-level model (*p*)
CLASS LEVEL **[g] Total number of weeks with one or more assignments assigned**	n/a	n/a	n/a	0.066	**0.034***
CLASS LEVEL **[h] Total number of assignments assigned**	n/a	n/a	n/a	0.017	0.060
CLASS LEVEL **[i] Percentage of assignments that were non–Skill Builder assignments**	n/a	n/a	n/a	3.270	**0.003***

Note: The two-level model was used for the student-level usage data; the three-level model was used for the class-level usage data. The one-level model did not apply to the class-level usage data. The *p*-values in **bold that end with an asterisk** indicate that they are smaller than 0.01 for student-level variables and smaller than 0.05 for class-level.

Conclusion and Discussion

WestEd conducted a rigorous, large-scale experiment of a schoolwide intervention at the intersection of three factors: educational technology, independent math practices, and formative assessment. In line with the prior efficacy trial of the same intervention (Roschelle et al., 2016), the results showed that ASSISTments and similar technology-based programs with accompanying professional learning and integrated practices may support teachers in incorporating formative assessment practices effectively in classrooms and thus better prepare students for immediate and future math learning. This study's findings contribute to an increased understanding of the potential of technology-based solutions for advancing mathematics learning and recovering from the learning loss caused by the COVID-19 pandemic (U.S. Department of Education, n.d., 2021).

The study’s findings have significant implications for improving mathematics learning in middle school for all students. The research was conducted in schools with diverse student populations, closely reflecting the makeup of student populations across the nation. The impact of the intervention was particularly pronounced for Hispanic students, a historically underrepresented racial and ethnic group in STEM education and careers (Fry, Kennedy, & Funk, 2021) and those attending traditionally underresourced schools. This highlights the potential of technology-based programs to reduce achievement gaps. These findings underscore the importance of scaling up effective technology-based programs to promote equitable access to high-quality math education for all students and to prepare them for future mathematics learning.

The findings also suggest that educators and schools may want to continue focusing efforts on using technology to improve math learning via formative assessment practices. Research has shown that regular independent practice and teachers’ use of resulting data to inform instruction are crucial for student learning (Irons & Elkington, 2022). However, traditional paper-based independent practice has limitations: When students complete their work, they may make errors and practice incorrectly, and teachers have limited time to analyze students’ progress, adjust their subsequent instruction, and determine students’ learning needs (Mendicino, Razzaq, & Heffernan, 2009). Thus, technology-based interventions like ASSISTments provide opportunities to address these challenges by providing immediate feedback directly to students and making independent practice a more productive learning opportunity for all students, particularly those with limited access to supportive adults and peers outside of school. Teachers reported that ASSISTments’ automatic scoring and reporting, which makes students’ math work performance immediately visible, enabled them to focus their energies away from grading and toward adapting their instruction.

A separate cost study was conducted to examine the cost-effectiveness of the ASSISTments program during the study (see Feng, Weiser, & Collins, 2024). The results showed that the estimated per student costs ranges from $40 to $70 to achieve the long-term effect size of 0.10, which suggests that ASSISTments is a relatively cost-effective solution compared with many other intervention options.

References

Allison, P. D. (2001). *Missing data*. Sage.

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2020). *Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time* [Working paper 27555]. National Bureau of Economic Research. https://www.nber.org/papers/w27555

Bohrnstedt G., Kitmitto S., Ogut B., Sherman D., & Chan D. (2015). *School composition and the Black–White achievement gap* [NCES 2015-018]. U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress. https://nces.ed.gov/nationsreportcard/subject/studies/pdf/school_composition_and_the_bw_achievement_gap_2015.pdf

Engzell, P., Frey, A., & Verhagen, M. D. (2020). Learning loss due to school closure during the COVID-19 pandemic. *SocArXiv Papers*. https://doi.org/10.31235/osf.io/ve4z7

Feng, M., Huang, C-W., & Collins, K. (2023). *Technology-based support shows promising long-term impact on math learning: Initial results from a randomized controlled trial in middle schools.* WestEd. https://www.wested.org/resources/technology-based-support-shows-promising-long-term-impact-on-math-learning-2/

Feng, M., Weiser, G., & Collins, K. (2024). *Cost and effectiveness of ASSISTments online math support: Analysis from a randomized controlled study in middle school.* WestEd. https://www.wested.org/resources/cost-and-cost-effectiveness-of-assistments-online-math-support/

Fry, R., Kennedy, B., & Funk, C. (2021). *STEM jobs see uneven progress in increasing gender, racial and ethnic diversity*. Pew Research Center. https://www.pewresearch.org/science/wp-content/uploads/sites/16/2021/03/PS_2021.04.01_diversity-in-STEM_REPORT.pdf

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, *24*(4), 470–497. https://doi.org/10.1007/s40593-014-0024-x

Heritage, M., & Popham, W. J. (2013). *Formative assessment in practice: A process of inquiry and action*. Harvard Education Press.

Irons, A., & Elkington, S. (2022). *Enhancing learning through formative assessment and feedback*. Routledge.

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342*(6161), 935–937. https://www.science.org/toc/science/342/6161

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549–565. https://doi.org/10.3102/0013189X20965918

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSER 2013-3000). Institute of Education Sciences, National Center for Special Education Research. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

Maldonado, J. E., & de Witte, K. (2020). *The effect of school closures on standardized student test outcomes*, Discussion Paper Series DPS20.17, Katholieke Universiteit Leuven.

Mendicino, M., Razzaq, L., & Heffernan, N. T. (2009). A comparison of traditional homework to computer-supported homework. *Journal of Research on Technology in Education*, *41*(3), 331–359. https://doi.org/10.1080/15391523.2009.10782534

Morgan, I. (2022). *Equal is not good enough: An analysis of school funding equity access across the U.S. and within each state.* The Education Trust. https://edtrust.org/wp-content/uploads/2014/09/Equal-Is-Not-Good-Enough-December-2022.pdf

Murphy, R., Roschelle, J., Feng, M., & Mason, C. A. (2020). Investigating efficacy, moderators and mediators for an online mathematics homework intervention. *Journal of Research on Educational Effectiveness*, *13*(2), 235–270. https://doi.org/10.1080/19345747.2019.1710885

North Carolina Department of Public Instruction. (2022, July 22). *EOG mathematics achievement level ranges and descriptors.* https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-mathematics-achievement-level-ranges-and-descriptors

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online mathematics homework increases student achievement. *AERA Open*, *2*(4). https://doi.org/10.1177/2332858416673968

Stein, P. (2020, October 30). *Washington Post: In D.C., achievement gap widens, early literacy progress declines during pandemic, data show*. The DC Line.

https://thedcline.org/2020/10/30/washington-post-in-d-c-achievement-gap-widens-early-literacy-progress-declines-during-pandemic-data-show/

U.S. Department of Education. (n.d.). *Supporting schools' use of evidence to guide EdTech adoption*. https://tech.ed.gov/evidence/

U.S. Department of Education. (2021). *Education in a pandemic: The disparate impacts of COVID-19 on America's students*. https://www2.ed.gov/about/offices/list/ocr/docs/20210608-impacts-of-covid19.pdf

von Hippel, P. T. (2007). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, *37*(1), 83–117. https://doi.org/10.1111/j.1467-9531.2007.00180.x

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Handbooks

White, I. R., & Thompson, S. G. (2005). Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine, 24*(7), 993–1007. https://doi.org/10.1002/sim.1981

This study was conducted independent of Worcester Polytechnic Institute and The ASSISTments Foundation, who are the founder and distributor of ASSISTments. A local coach was hired by Worcester Polytechnic Institute to provide the professional learning and support to the teachers in the intervention group that was part of the intervention package. All data collection and analysis activities were carried out independently by the study team.

Acknowledgements: We extend thanks to Dr. Yvonne Kao for her feedback on the research and the ASSISTments team for their support during the study. We gratefully acknowledge the support of district leaders and the participation of the teachers and students within the districts where this study was conducted.

This material is based on work supported by Arnold Ventures and the Institute of Education Sciences of the U.S. Department of Education under Grant R305A170641. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Arnold Ventures or the Institute of Education Sciences.

WestEd.org