Impact of Imagine Learning Illustrative Mathematics in California: A School-Level Quasi-Experimental Study

Lucia M. Reyes, Michael A. Cook, PhD,
Steven M. Ross, PhD

January 2026

Johns Hopkins University

Impact of Imagine Learning Illustrative Mathematics in California: A School-Level Quasi-Experimental Study

Lucia M. Reyes, PhD, Michael A. Cook, PhD, Steven M. Ross, PhD

Center for Research and Reform in Education
Johns Hopkins University School of Education
2800 N. Charles St
Baltimore, MD 21218
https://education.jhu.edu/crre

Preferred citation: Reyes, L. M., Cook, M. A., Ross, S. M. (2026). *Impact of Imagine Learning Illustrative Mathematics in California: A school-level quasi-experimental study.* Center for Research and Reform in Education, Johns Hopkins University.

Contents

EXECUTIVE SUMMARY

In this quasi-experimental study, we examined the impact of the Imagine Learning Illustrative Mathematics curriculum on mathematics achievement in 1,309 California schools. The study was designed according to the *What Works Clearinghouse Procedures and Standards Handbook v5.0* (WWC, 2022). Measures included school-level achievement (i.e., percent of students meeting or exceeding proficiency benchmarks) on the Smarter Balanced Assessment Consortium (SBAC) mathematics assessment and school-level demographics.

- The study used a school-level Intent-to-Treat (ITT) design. Schools in districts that contracted with Imagine Learning to use the Imagine Learning Illustrative Mathematics curriculum were matched to comparison schools in districts that received business-as-usual math instruction.
- Propensity score matching was used to achieve a well-balanced analytic sample of 3,128 schools, including all 1,309 treatment schools with available data and 1,819 comparison schools with similar characteristics.
- The analytic sample included demographically diverse elementary, middle, and high schools across the state of California.
- Data sources were publicly available datasets from the California Department of Education (CDE) including school-level demographics and SBAC achievement percentages from the Spring of 2024 and the Spring of 2025.
- Impact analyses using multiple linear regression with cluster robust standard errors (to account for schools clustered within districts) and controlling for the prior year's school-level achievement and demographics showed a directionally positive and statistically significant effect of Imagine Learning Illustrative Mathematics on school-level math achievement.
- On average, schools that received Imagine Learning Illustrative Mathematics had a 1.38 percentage point increase (ES = .07 SDs) in students that met or exceeded SBAC proficiency benchmarks in 2025 relative to similar comparison schools that received business-as-usual math instruction.

INTRODUCTION

Overview of Imagine Learning Illustrative Mathematics

Imagine Learning Illustrative Mathematics is a core curriculum that emphasizes problem-based learning, student discourse, conceptual understanding, and procedural fluency. It blends print materials with digital resources, manipulatives, and videos to provide all students, including multilingual learners and diverse learners, access to grade-level content. Lessons include five core structural components:

1. Lesson Warm-Ups,
2. Lesson Instructions and Activities,
3. Lesson Synthesis,
4. Lesson Cool-Downs, and
5. Centers (K-5 only).

Each unit also includes pre-assessments, problem checkpoints, and practice problems for students to engage in, and the curriculum includes end-of-unit and end-of-course assessments. Teachers shape lesson plans to meet their students' learning needs and use assessments to monitor student progress. Previous studies using student-level data in Missouri and Iowa showed favorable effects of the Imagine Learning Illustrative Mathematics curriculum, particularly for Black students and students with special education needs (Cook, Eisinger, & Ross, 2023; Cook & Ross, 2025).

Overview of the Evaluation

In the fall of 2025, Imagine Learning partnered with the Center for Research and Reform in Education (CRRE) at Johns Hopkins University to evaluate the impact of its Imagine Learning Illustrative Mathematics curriculum across the state of California. To assess program impact, Imagine Learning shared with CRRE unaltered publicly available school-level demographic and performance data retrieved from the California Department of Education (CDE) website, along with a list of districts that implemented Imagine Learning Illustrative Mathematics during the 2024-2025 school year. Data included Spring 2024 (pretest) and Spring 2025 (posttest) school-level Smarter Balanced Assessment Consortium (SBAC) mathematics proficiency percentages for all California schools. The following research question guided the study design: What are the impacts of using the Imagine Learning Illustrative Mathematics curriculum on school-level SBAC math achievement gains?

METHOD

Research Design

This study used a quasi-experimental, Intent-to-Treat (ITT) design (QED) following the *What Works Clearinghouse Procedures and Standards Handbook v5.0* (WWC, 2022). Prior to the start of the 2024-2025 school year, 38 California school districts contracted with Imagine Learning to use the Imagine Learning Illustrative Mathematics core curriculum, thus being assigned to the treatment group. The outcome measure for this study was the school-level percentage of students that met or exceeded the state's SBAC mathematics achievement benchmark in the Spring of 2025.

Matching

Since assignment to the treatment group was not random (i.e., school districts chose to contract with Imagine Learning to implement the Illustrative Mathematics curriculum), Propensity Score Matching (PSM) was used to create a matched comparison group with similar baseline characteristics relative to the treatment group. The propensity score represents the probability of treatment assignment conditional on observed baseline characteristics and is widely used as a balancing score (Austin, 2011). Propensity scores for all schools were calculated by using school-level characteristics to predict the probability of assignment to the treatment condition using a logistic function. School-level characteristics used during matching included pretests (i.e., the percent of students that met or exceeded Spring 2024 SBAC achievement benchmarks) and the percentage of students that were identified as Hispanic, White, socioeconomically disadvantaged, and homeless. Matches were selected from the comparison pool using a 2:1 nearest neighbor approach, meaning that up to two comparison schools could be matched to each treatment school. A caliper (measure of how different each treatment student is allowed to be from their matched comparison counterparts) of .2 was used to ensure a narrow threshold of similarity between propensity scores of matched schools. All matching procedures were conducted in Stata (v19).

Attrition and Missing Data

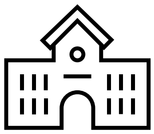
All schools with available pre-test data were included in the matching process. Seven hundred and sixty-seven (8.53%) schools that did not meet inclusion criteria for matching due to missing pre-test data were excluded prior to matching. After matching, thirty-three (1.04%) of the schools in the matched sample did not have post-test data and were thus considered attrited. The proportion of matched schools that were excluded from the analytic sample due to attrition was similar in the treatment (1.06%) and matched comparison (1.03%) groups (see Figure A1 in the Appendix). Descriptive characteristics of the analytic sample relative to the full school population are shown in Table A1.

Missing demographic data in the matched analytic sample was less than 10% for each variable, and we used replacement with a constant (the grand mean) to impute missing observations. An indicator of observations that were imputed was included in the final regression models.

Participants

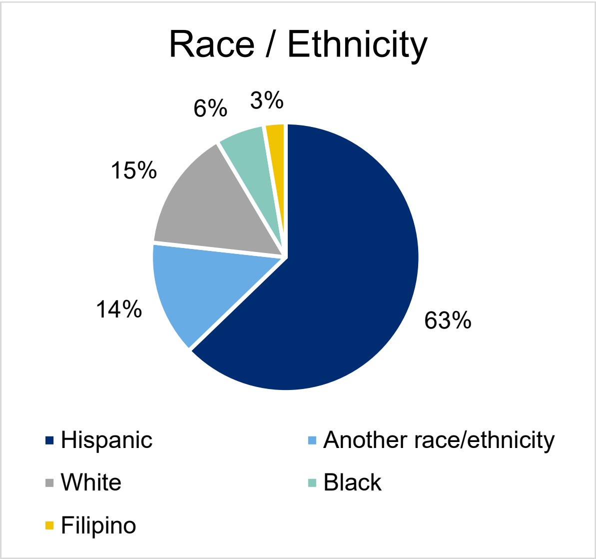
Details about participating schools included in the analytic sample are presented below.

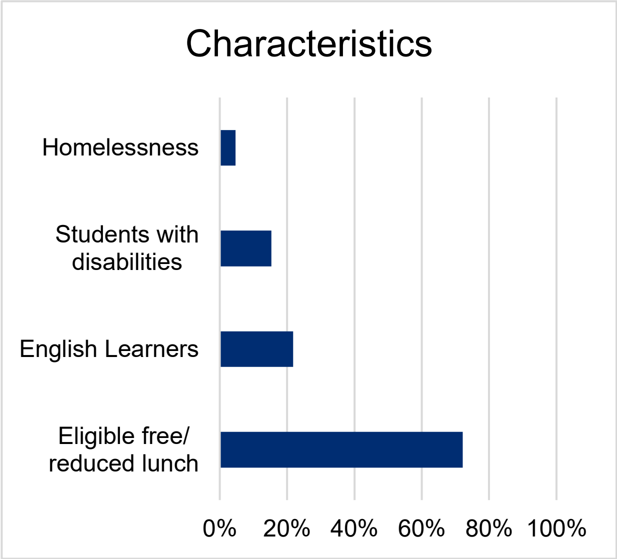
533 school districts
(33 T; 500 C)



3,128 Grade K-12 schools
(1,309 T; 1,819 C)

Demographic snapshot of participating schools[1]





Most students in participating schools within the analytic sample were identified as Hispanic (63%), followed by students that identified as White (15%) and another race or ethnicity (14%). Almost three-quarters of the analytic sample received free or reduced lunch, indicating vast economic disadvantage. Table 1 displays baseline characteristics of the matched sample by study condition.

Baseline difference effect sizes were used to compare means and proportions of baseline characteristics across the treatment and comparison groups, using the procedures outlined by the What Works Clearinghouse (WWC, 2022). For continuous variables, the Hedges' *g* effect size was computed by dividing the unstandardized mean difference by the pooled within-group standard deviation and multiplying the result by a small-sample correction factor (Hedges, 1981). For dichotomous variables, effect size

[1] School-level averages. Publicly available data acquired from the California Department of Education.

differences were calculated using Cox's Index, which is designed to produce a comparable effect size measure to Hedges' *g* (WWC, 2022). Cox's Index applies a logarithmic to the odds ratio for the intervention group, subtracts the same for the comparison group and then divides by 1.65. As shown in Table 1, the matched groups were well balanced in terms of all characteristics.

Table 1

Demographics of Schools in the Matched Analytic Sample (N = 3,128)

	Treatment (n = 1,309)	Matched comparison (n = 1,819)	Effect size[1]
2024 math achievement (pretest) (M)	33.77	34.07	*0.01*
% White	13.20	15.82	*-0.21*
% Hispanic	64.34	61.67	*0.11*
% Disability status	15.93	14.8	*0.09*
% Homeless	4.45	4.73	*-0.06*
% Socioeconomically disadvantaged	75.29	71.28	*0.21*
% Eligible free/reduced lunch	74.86	70.57	*0.22*
% English language learner	21.04	22.67	*-0.10*
% Female	48.16	48.33	*0.00*
Total students (*M*)	604.01	615.79	*-0.03*
Grade K-5 students (*M*)	260.91	288.61	*-0.12*
Grade 6-8 students (*M*)	128.51	131.71	*0.01*
Grade 9-12 students (*M*)	166.59	169.86	*0.01*

[1]Hedges' g or Cox's Index. *M* = mean.

Measures

To address the research questions, the CRRE study team processed and analyzed administrative data housed at CDE from the Spring of 2024 and the Spring of 2025. Measures included school-level SBAC mathematics percent achievement benchmarks, demographic characteristics, and treatment assignment, as described below:

School-level Smarter Balanced Assessment Consortium (SBAC) math achievement benchmarks were obtained from the California Department of Education's (CDE) publicly available administrative datasets. Local education agencies and charter schools in California systematically submit Census Day TK/K-12 enrollment data in the California Longitudinal Pupil Achievement Data System (CALPADS) each fall. These data are certified by the district superintendent, charter school administrators or their designees prior to submission (CDE, 2025). The SBAC utilizes computer-based tests and performance tasks to measure English language arts/literacy and mathematics competencies and is administered in Grades 3-8 and 11. The state of California uses four levels (Standard Not Met, Nearly Met, Met, Exceeded) to indicate readiness based on results. The outcome measure for the current study was the school-level percentage of students who tested at or above the state benchmark on the math component of the SBAC assessment. This percentage is calculated as the sum of students who tested with a score at Standard Exceeded and students who tested with a score at Standard Met, divided by the sum of students who tested with a score at any performance level, and multiplied by 100. The What Works Clearinghouse considers administrative data, such as CDE CALPADS records, to be valid and reliable (WWC, 2022).

School-level demographic characteristics. CDE administrative datasets submitted by local school districts were also used to obtain demographic indicators, including race/ethnicity, free/reduced school lunch status, English language learner status, disability status, homelessness, district size, and grade bands.

District-level Imagine Learning Illustrative Mathematics assignment. Imagine Learning provided CRRE with a list of districts that contracted to use the Imagine Learning Illustrative Mathematics core curriculum. Usage metrics were not available for the current study and not necessary, given the Intent-to-Treat design.

Analytical Approach

Data for schools were analyzed descriptively by examining the mean, standard deviation, and range of SBAC achievement benchmarks, as well as change from pretest (2024) to outcome (2025). Change was calculated by subtracting pretest scores from outcome scores and comparing differences with t-tests. Multiple Linear Regression (MLR) was used to determine the impact of Imagine Learning Illustrative Mathematics on SBAC achievement benchmarks, controlling for demographic variables including race/ethnicity, free/reduced school lunch status, English language learner status, disability status, homelessness, and home district grade bands. Covariates were grand mean centered in order to obtain an interpretable intercept. Since assignment of the Imagine Learning Illustrative Mathematics treatment took place at the district level, we used clustered robust standard errors to account for the potential non-independence of schools nested within districts. Additionally, we conducted a supplementary hierarchical linear model (HLM) to account for potential random effects of school districts (see Appendix B). All quantitative analyses were completed in Stata (version 19).

To facilitate the interpretation of the treatment effect, the unstandardized regression coefficient for the dichotomous predictor was transformed into a standardized effect size. Following conventional procedures (see Bornstein et al., 2019), the effect size was calculated by dividing the unstandardized regression coefficient by the pooled within-group standard deviation of the dependent variable. This transformation allowed for a scale-free measure of the treatment's magnitude relative to the variability of the outcome.

RESULTS

The results section of the report begins with descriptive achievement results, followed by impact analyses comparing achievement outcomes controlling for baseline characteristics.

Descriptive Achievement Results

Table 2 describes gains in school-level math achievement. Gains were calculated by subtracting pretest (Spring 2024 math achievement) from the outcome (Spring 2025 math achievement), with dependent t-tests used to determine statistical significance of Spring 2024 to Spring 2025 gains.

Table 2
Descriptive Analysis of School-Level Math Achievement

	Treatment n = 1,309		Comparison n = 1,819		Sig.	Effect size[1]
	Mean	SD	Mean	SD		
Spring 2024 (pretest)	33.77	19.78	34.07	20.55	0.678	
Spring 2025 (posttest)	36.49	19.26	35.32	20.40	0 .106	
Pre-to-post gains	2.72	5.72	1.25	4.85	< .001	.32

[1] Bias corrected standardized mean difference (i.e., Hedges *g*) of unadjusted change in % met or exceeded from 2024 to 2025 between treatment and control groups.

As shown in Table 2, schools in the treatment and comparison groups had similar proportions (~34%) of students that met or exceeded state benchmarks for math achievement during the Spring of 2024. While schools in both groups showed improvements in achievement in the Spring of 2025, gains were greater for the treatment group. On average, schools in the treatment group gained 2.72% points from pre- to post-test, compared to 1.25% in the comparison group. This difference in gains was statistically significant, suggesting that schools using Imagine Learning Illustrative Mathematics outperformed schools using business as usual.

Imagine Learning Illustrative Mathematics' Impact Results

The impact analyses compared Spring 2025 math achievement between groups while accounting for baseline differences, including demographic characteristics (i.e., race/ethnicity, free/reduced lunch, disability status, homelessness, English language learner status, school size), and pretest scores. Table 3 presents impact results for the multiple regression model.

Table 3

Impact Analysis of Imagine Learning Illustrative Mathematics on 2025 School-Level Math Achievement (N = 3,128)

Variable	Estimate	Standard error	*p* value	95% CI	Effect size
Imagine Learning Illustrative Mathematics	1.38**	0.48	.004	0.43 – 2.33	0.07
Constant	35.23***	0.13	<.001	34.98 – 35.49	

Note. ** *p < .01, *** p < .001.*

The regression estimate can be interpreted as the average difference in 2025 math achievement between the treatment and comparison group, controlling for relevant baseline characteristics and the previous year's math performance. On average, the treatment group scored 1.38 percentage points higher than the comparison group on school-level Spring 2025 math achievement. This difference was directionally positive and statistically significant, indicating a favorable impact of Imagine Learning Illustrative Mathematics on school-level math achievement. Since covariates were centered on the grand-mean, the constant (i.e., intercept) can be interpreted as the school-level average percentage of students that met or exceeded the state's math achievement benchmark in the comparison group (~35%).

In the same model, the school-level percent of students identified as homeless (p = .019) and as English Language Learners (p < .001) were negatively associated with 2025 math achievement. As expected, Spring 2024 math achievement was positively and significantly associated with Spring 2025 achievement ranks (p < .001). All other covariates included in the model (i.e., race/ethnicity, disability status, free/reduced school lunch, school size, and indicators of imputed observations) were not statistically significant.

Finally, a supplemental analysis was conducted using hierarchical linear modeling (HLM) to estimate potential influences of the clustering of schools within districts. While the regression coefficient (0.58) for the impact of the treatment was directionally positive, showing a trend towards higher achievement in the treatment group, the effect was not statistically significant (p = .076) in the HLM. However, the results of this model should be interpreted with caution given the small number of treated clusters (n = 33) relative to the comparison (n = 500). In an HLM, this imbalance might lead to underestimated or unstable standard errors for the treatment effect, limit power to detect small effects, and heighten sensitivity to influential clusters. Detailed results of the HLM are provided in Appendix B.

DISCUSSION

The purpose of the present study was to evaluate the impact of the Imagine Learning

Illustrative Mathematics core curriculum on school-level math achievement. A quasi-experimental design was conducted to compare school-level SBAC math achievement percentages for schools in districts that contracted to use Imagine Learning Illustrative Mathematics relative to similar schools who continued business-as-usual instruction. Propensity Score Matching was performed to achieve a well-balanced analytic sample for the comparison.

Imagine Learning Illustrative Mathematics' Impact on Math Achievement

Results showed a significant positive effect of the Imagine Learning Illustrative Mathematics curriculum on Spring 2025 school-level math achievement (i.e., the percentage of students that met or exceeded the state's benchmark for the SBAC math assessment). The main model's effect size of .07 is considered small to medium in magnitude. Notably, in applied program evaluations, effects in the .05-.10 range represent educationally meaningful improvements in student outcomes (Kraft, 2020).

Methodological Considerations and Future Directions

The study was conducted according to WWC standards (2022), which conventionally form the basis determining the Every Student Succeeds Act (ESSA) tiers of evidence. Analyses were conducted by a third-party evaluator (i.e., CRRE) and relied on publicly available administrative data from the California Department of Education. All California districts assigned to use Imagine Learning Illustrative Mathematics in the 2024-25 school year that had SBAC score data were included in the study, consistent with an ITT design. Comparison schools were chosen based on prior math proficiency and school-level demographic variables. Both overall and differential attrition were low (<2%). Baseline differences on all variables were less than .25 standard deviations between the treatment and matched comparison samples. These design components helped to minimize potential bias in impact estimates and fulfill the criteria associated with ESSA Tier 2 ("Moderate") evidence.

Future studies that can assess Imagine Learning Illustrative Mathematics implementation can contribute to a more precise understanding of the conditions under which the intervention is most effective. Nonetheless, the current ITT design constitutes a robust test of real-world effectiveness with broad generalizability. Analyses using student-level data to explore subgroup variation in achievement across study conditions were beyond the scope of this study but would further illuminate patterns of response to the intervention. In conclusion, the current California statewide study provides rigorous evidence supporting the benefits of Imagine Learning Illustrative Mathematics relative to business-as-usual math instruction in raising student achievement.

REFERENCES

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Borenstein, M. & Hedges, L. V. (2019). Effect sizes for meta-analysis. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.*), The handbook of research synthesis and meta-analysis* (3rd ed., pp. 207–244). Russell Sage Foundation.

California Department of Education (CDE; November, 2025). *File Structure: Census Day Enrollment Data*. Retrieved from https://www.cde.ca.gov/ds/ad/fsenrcensus.asp#GenDes

Cook, M.A., Eisinger, J., Ross, S.M. (2025). *Impact Evaluation of Imagine Learning Illustrative Mathematics in Fort Zumwalt School District*. Center for Research and Reform in Education, The Johns Hopkins University.

Cook, M.A., Ross, S.M. (2025). *Efficacy Study of Imagine Learning Illustrative Mathematics in Iowa City Community School District*. Center for Research and Reform in Education, The Johns Hopkins University.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics, 6*(2), 107–128. https://eric.ed.gov/?id=EJ248027

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241-253. https://doi.org/10.3102/0013189X20912798

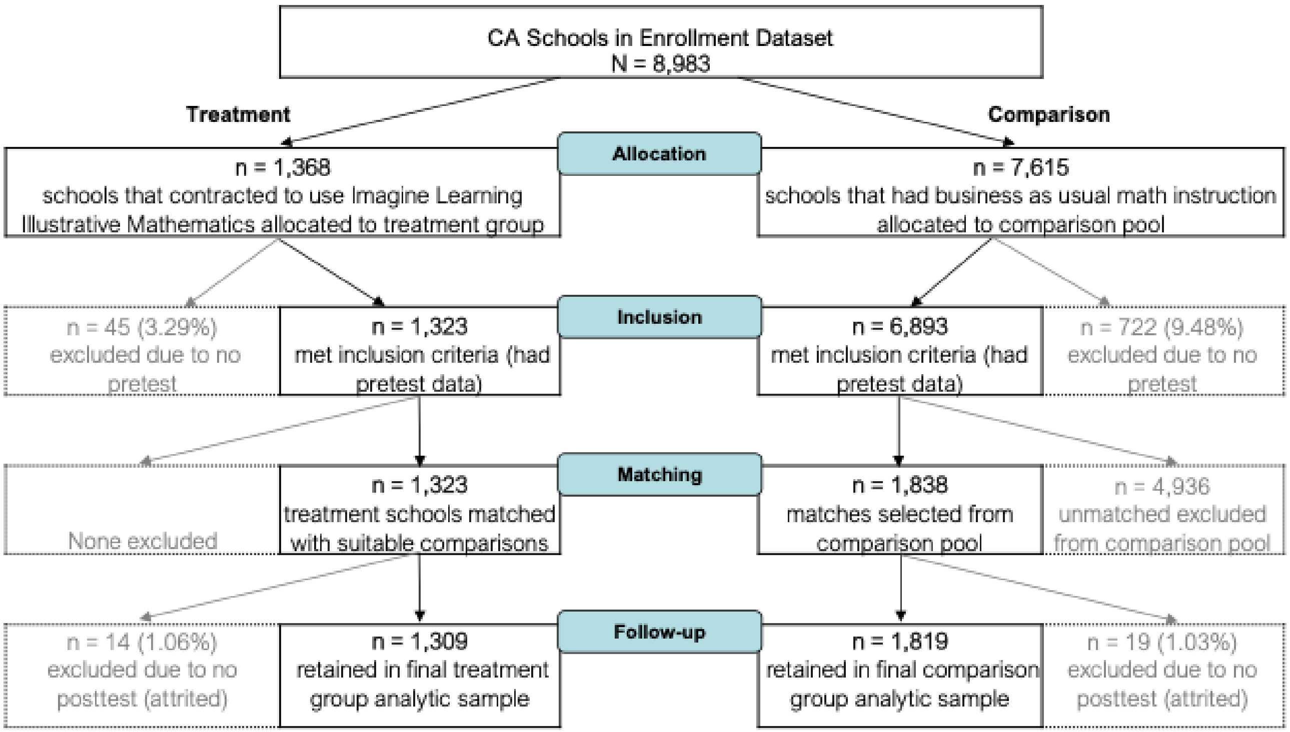
What Works Clearinghouse. (WWC; 2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

APPENDIX A:

Attrition and Representativeness

Figure A1

Flow of participating schools



Note. Overall attrition (33 with no posttest / 3,161 matched) = 1.04%
Differential attrition (1.06% treatment attrition – 1.03% comparison attrition) = 0.03%

Table A1

Characteristics of Schools Included and Excluded from Final Analytic Sample

	School population (Before matching) *N* = 8,216	Excluded from analytic sample *n = 5,088*	Included in analytic sample *n = 3,128*
Spring 2024 Math, % met or above (M)	33.88	33.84	33.94
% White	20.74	24.44	14.72
% Hispanic	56.88	53.24	62.79
% Disability	15.26	15.24	15.28
% Homeless	5.21	5.57	4.61
% Socioeconomically disadvantaged	65.87	61.51	72.96
% Eligible free/reduced lunch	65.16	60.73	72.34
% English language learners	20.10	18.94	21.98
% Female	48.10	48.00	48.26
# Students (M)	617.51	621.69	610.86
# Grade K-5 students (M)	255.56	242.36	277.01
# Grade 6-8 students (M)	135.51	138.67	130.37
# Grade 9-12 students (M)	193.83	209.40	168.49

APPENDIX B:

Supplementary Hierarchical Linear Models to Assess Impact

In addition to the planned main analyses, we conducted a supplementary hierarchical linear model (HLM) to account for the potential non-independence of observations (i.e., the possibility that school scores within a district were more similar to each other than students' scores across districts). Unlike traditional regression models, HLMs allow for separate regression lines for each cluster (i.e., district), estimating not only relationships that are consistent across all schools (fixed effects), but also the contribution of unobserved between-district differences (random effects). By calculating an intraclass correlation coefficient (ICC), HLMs estimate the percent of the variance in the outcome that is attributable to the clustering (i.e., district-level differences). Results of this model are presented in Table B1.

Table B1

Hierarchical Linear Model of Imagine Learning Illustrative Mathematics Impact on Math Achievement (N = 3,128)

	Coefficient	Std. err.	95% conf. interval		Sig.
Imagine Learning Illustrative Mathematics	0.58	0.33	-0.06	1.23	0.076
Constant	7.92	0.88	6.19	9.66	<0.001
Residual intraclass correlation	ICC	Std. err.	95% conf. interval		
District	0.05	0.01	0.03	0.08	