k12 Tutoring

Accelerating Math Growth Through High-Impact Tutoring: Evidence from a Multi-State Virtual School Study

ESSA Level II Report

August 2025

Contents

Executive Summary

To assess the impact of its tutoring services on student achievement, K12 Tutoring carried out a comprehensive evaluation. This study was specifically designed to satisfy the Moderate Evidence (Level II) requirements outlined by the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This study employed a quasi-experimental approach with propensity score matching to compare outcomes for 429 students in grades 2 through 6 enrolled in six full-time virtual schools across Arkansas, Arizona, Michigan, West Virginia, and Wisconsin during the 2024–2025 academic year. Among these students, 143 received tutoring services, while 286 were matched peers selected based on similar academic and demographic profiles. The sample was 52% female and racially diverse, including 48% White, 26% African American, 12% Hispanic, and 2% Asian students. Furthermore, 47% qualified for free or reduced-price lunch, 12% received special education support, and 3% were identified as English Language Learners.

Tutoring engagement was measured through total minutes of participation, and academic outcomes were assessed using the Renaissance Star Growth Assessment in math. The impact of tutoring was analyzed using regression models that controlled for key student characteristics. Baseline equivalence between the tutored and comparison groups was established in line with What Works Clearinghouse (WWC) standards to ensure the validity of the comparisons.

Results

Tutoring Participation. Among students who received K12 Tutoring, the median cumulative tutoring time over the 2024–2025 school year was 540 minutes. While median tutoring time varied slightly by grade level, ranging from 480 to 630 minutes, these differences were not statistically significant. Tutoring was delivered primarily in small-group formats, typically with fewer than four students per session.

Student Outcomes. Students who received tutoring had significantly higher end-of-year (EOY) Star Math scores than their matched peers. This difference was statistically significant (p = .02) and represented an effect size of 0.17 standard deviations, which translates to an additional gain of approximately 6.36 percentile points for tutored students.

Conclusion

Students who participated in K12 Tutoring demonstrated higher gains in mathematics achievement compared to their matched peers. These results provide evidence of the program's effectiveness in accelerating learning. The study meets ESSA Tier 2 (Moderate Evidence) criteria, supported by a well-executed quasi-experimental design that established baseline equivalence, applied appropriate statistical controls, included a sufficiently large and diverse sample across multiple schools, and yielded at least one statistically significant, positive outcome.

Introduction

Recent results from the National Assessment of Educational Progress (NAEP) indicate significant declines in fourth-grade mathematics achievement, highlighting urgent concerns about early math learning in the United States. Between 2019 and 2022, average fourth-grade math scores dropped by five points nationwide, representing one of the largest declines in the assessment's history (National Center for Education Statistics [NCES], 2022). These losses were especially pronounced among lower-performing students, exacerbating pre-existing achievement gaps related to income, race, and geography (NCES, 2022). Fourth grade represents a pivotal year in students' mathematical development, as it marks the transition from basic arithmetic to more complex concepts such as multi-digit operations, fractions, and early algebraic thinking (National Mathematics Advisory Panel, 2008). Deficits at this stage can have long-term consequences for students' academic trajectories, limiting their readiness for middle-grade math and beyond (Siegler et al., 2012). The scale and persistence of these declines underscore the need for evidence-based interventions that provide support as students approach this pivotal stage in mathematical learning and continue to receive guidance as they transition into more complex mathematical content.

High-impact tutoring has emerged as one of the most effective strategies for accelerating student learning and addressing persistent academic deficits in mathematics. Unlike traditional forms of remediation, high-impact tutoring is characterized by frequent sessions (at least three times per week), sustained implementation over time, alignment with classroom instruction, and strong tutor-student relationships (Robinson et al., 2021; Kraft & Falken, 2021). Experimental and quasi-experimental studies have demonstrated that well-designed tutoring programs can yield substantial gains in math achievement, often equivalent to several months of additional learning (Nickow et al., 2024). These effects are particularly pronounced for students who are furthest behind, making tutoring a critical tool for promoting educational equity in the wake of pandemic-related learning loss.

Prior research consistently demonstrates that tutoring can produce substantial improvements in math achievement, particularly when implemented with high dosage and fidelity. Meta-analyses of randomized controlled trials and quasi-experimental studies have found that tutoring interventions in mathematics yield effect sizes up to 0.30 standard deviations, depending on program design and student population (Nickow et al., 2024). Several large-scale evaluations have shown that high-dosage tutoring can produce the equivalent of four to ten months of additional learning in math over the course of an academic year (Robinson et al., 2021). These gains have been documented across grade levels, with particularly strong effects observed in upper elementary and middle school students who began the year performing below grade level. The consistent positive findings across diverse settings and student groups highlight tutoring as a highly scalable intervention for addressing math achievement gaps.

K12 Tutoring delivers individualized academic support through certified tutors, aligning instruction with school curricula to reinforce classroom learning. In mathematics, the program focuses on strengthening students' core understanding and confidence by addressing key concepts, building procedural fluency, and developing problem-solving strategies essential for long-term success.

To demonstrate effectiveness of their program and services, the impact of K12 Tutoring on student outcomes was examined. The study was also designed to satisfy Level II requirements (Moderate Evidence) of the Every Student Succeeds Act (ESSA).

The study had the following research questions:

Implementation Research Question

1. What is the cumulative duration of tutoring received by students participating in the K12 Tutoring program?

Impact Research Question

2. To what extent did participation in the K12 Tutoring program impact end-of-year Star math outcomes relative to a matched group of non-participating students?

Method

This section outlines the study's methodology, including its design, context, participant characteristics, outcome measures, analytical approach, and the procedures used to establish baseline equivalence.

Study Design

This study utilized a quasi-experimental design with propensity score matching to meet ESSA Tier 2 evidence requirements. The goal was to estimate the impact of K12 Tutoring by comparing outcomes between students who received tutoring during the 2024–2025 school year and a matched sample of peers who did not participate. Comparison students were selected based on similar demographic characteristics and comparable performance on fall benchmark assessments to help ensure baseline equivalence between groups.

Setting and Participants

The study was carried out during the 2024–2025 school year and focused on a sample of 429 students in grades 2 through 6, drawn from six full-time virtual schools located in five states: Arkansas, Arizona, Michigan, West Virginia, and Wisconsin. While each school operated under its own administrative structure, the tutoring program followed a standardized implementation model across all sites.

The analytic sample included 143 students who received tutoring and 286 matched peers who did not. Overall, 52% of the sample were female. In terms of racial and ethnic composition, the sample was 48% White, 26% African American, 12% Hispanic, 2% Asian, and 4% multiracial, with 8% of students not reporting race or ethnicity. Socioeconomic and program eligibility data showed that 47% qualified for free or reduced-price lunch, 12% received special education services, and 3% were classified as English Language Learners.

Measures

To evaluate the impact of K12 Tutoring on student achievement, the study drew on a range of assessment data that captured academic performance across key content areas. These measures provided the basis for analyzing differences between tutored and non-tutored students.

K12 Tutoring Participation Metrics. Participation data were collected at the student level, including the total number of minutes each student spent in tutoring sessions. These data were used to assess patterns of engagement with the K12 Tutoring program during the 2024–2025 academic year. In addition to describing overall participation, the metrics supported analyses examining whether the extent of tutoring was associated with differences in academic performance.

Student Outcomes. Academic achievement in mathematics was evaluated using administrative data from the Renaissance Star Math Growth Assessment (Star). Students completed the assessment at both the beginning (BOY) and end (EOY) of the 2024–2025 school year. Scaled scores from these two testing points were used to assess students' academic progress and served as the primary outcome measure for determining the impact of K12 Tutoring.

Data Analysis

Descriptive analyses were conducted to summarize student characteristics and examine patterns related to program implementation. To estimate the effect of K12 Tutoring on math achievement, the study utilized regression modeling techniques that included student-level covariates. Propensity score matching with nearest neighbor selection was applied to minimize potential selection bias and strengthen causal inference. In addition, standardized effect sizes were calculated to provide interpretable measures of the differences in outcomes between tutored students and their matched counterparts.

Baseline Equivalence

Establishing baseline equivalence between the matched treatment and comparison groups was critical to ensure the validity of the findings and to meet ESSA Tier 2 standards. The methods and results of these equivalence assessments are detailed in the Results section and are further elaborated in Appendix A.

Results

Implementation Research Findings

Utilization data were analyzed to evaluate the level of student engagement with K12 Tutoring throughout the 2024–2025 academic year.

For students who participated in K12 Tutoring, what was the cumulative amount of tutoring time received?

Student utilization data were analyzed to understand the extent of engagement with the K12 Tutoring program during the 2024–2025 school year. Among those who participated, the median cumulative tutoring time was 540 minutes. While average total tutoring time differed slightly by grade: 570 minutes for 2nd grade, 480 minutes for 3rd and 4th grade, 630 minutes for 5th grade, and 510 minutes in 6th grade. These variations were not statistically different from one another, $p > .05$. Throughout the academic year, tutoring was delivered primarily in small groups, typically with fewer than four students assigned to each tutor.

Impact Research Findings

To investigate the study's impact questions, regression analyses were conducted comparing matched groups of students who participated in K12 Tutoring with those who did not, alongside analyses focused exclusively on tutoring participants. This section presents results pertaining to (a) the effect of K12 Tutoring on student learning outcomes. Statistical significance is noted only for findings with p-values less than .05.

To what extent did K12 Tutoring impact end-of-year (EOY) Star math outcomes relative to non-participating students?

A matched sample was developed by pairing students who received K12 Tutoring with non-participants, using baseline Star scores, school, grade, gender, ethnicity, and special education status as matching criteria. Baseline equivalence was established between the two groups, demonstrated by a Hedges' g of .02 and a mean difference of 1.95 scale score points, confirming comparable Star scores prior to tutoring.

Within this matched sample, a linear regression was performed to assess the impact of tutoring on end-of-year (EOY) Star scores. Tutoring status was treated as a binary predictor, while controlling for beginning-of-year (BOY) Star scores and demographic variables including school, grade, gender, ethnicity, and special education status. The analysis revealed that students who participated in K12 Tutoring had significantly higher EOY Star scores than their matched counterparts. This effect was statistically significant (Hedges' g = .17, p = .02; see Figure 1), and represents an additional gain of approximately 6.36 percentile points for tutored students compared with the control group.

Students who participated in K12 Tutoring demonstrated significantly higher end-of-year scale scores on the Renaissance Star Math Assessment compared to students who did not receive tutoring.
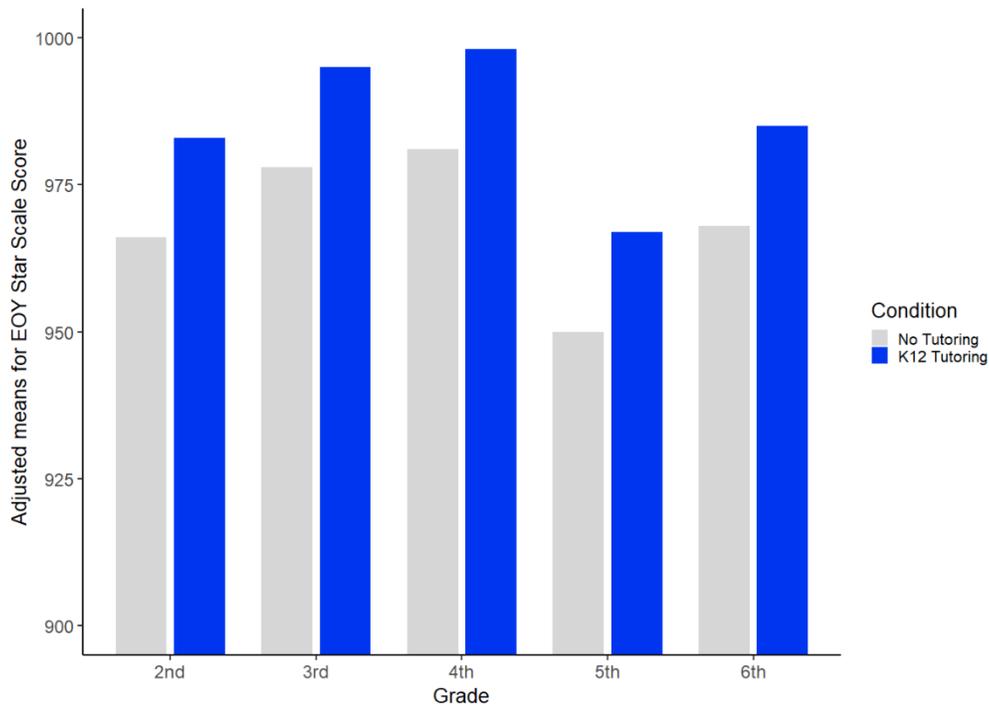

Figure 1. Adjusted means for end-of-year Star math assessment scale scores for K12 Tutoring and non-users. Mean difference was statistically significant (p = .02, Hedges' g = .17, n = 429).

Conclusions

The results of this study indicate that students who participated in K12 Tutoring achieved higher scores on the end-of-year (EOY) Star Math Assessment compared to matched peers, even after controlling for baseline achievement and demographic factors. These findings underscore the value of targeted, sustained tutoring as a promising approach for narrowing achievement gaps and supporting learners with greater academic needs.

Based on these positive outcomes, this study meets the ESSA Tier 2 (Moderate Evidence) standards. Specifically, the quasi-experimental design satisfied key criteria, including:

- ✓ Proper design and implementation
- ✓ Baseline equivalence for treatment and comparison groups
- ✓ Statistical controls through covariates
- ✓ At least 350 students in the analysis sample
- ✓ Representative, multi-site study
- ✓ At least one statistically significant, positive finding

References

Kraft, M. A., & Falken, G. (2021). A blueprint for scaling tutoring across public schools. *Brown University Annenberg Institute.*

National Center for Education Statistics. (2022). *The Nation's Report Card: Mathematics 2022*. U.S. Department of Education, Institute of Education Sciences. https://www.nationsreportcard.gov

National Mathematics Advisory Panel. (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. U.S. Department of Education.

Nickow, A., Oreopoulos, P., & Quan, V. (2024). The promise of tutoring for PreK–12 learning: A systematic review and meta-analysis of the experimental evidence. *American Educational Research Journal, 61*(1), 74-107.

Robinson, C. D., Kraft, M. A., Loeb, S., Schueler, B. E., & Stuart, E. A. (2021). *Accelerating student learning with high-dosage tutoring*. EdResearch for Recovery Project.

Pondiscio, R. (2024). Is it time for a second look at virtual education? *The Thomas B. Fordham Institute.* Retrieved from https://fordhaminstitute.org/national/commentary/it-time-second-look-virtual-education

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691–697.

Appendix A. Additional Information on Study Design and Methods

A total of 143 K12 Tutoring students in grades 2 through 6 had complete records for demographics, achievement, and program participation. A corresponding group of comparison students with complete demographic and achievement data was also identified. Propensity score matching was then applied—using nearest neighbor matching without replacement at a 2:1 ratio—to pair K12 Tutoring students with non-users who had similar academic and demographic characteristics. Matching variables included baseline Star scores, school, grade, gender, ethnicity, and special education status. This process resulted in a matched sample comprising 143 tutored students and 286 matched non-users, totaling 429 students. Further details on participant demographics can be found in Table A1.

Table A1. Student demographics by group for matched sample.

Characteristic	K12 Tutoring (n = 143)		Non-users (n = 286)		Total Sample (N = 429)	
	Percent	n	Percent	n	Percent	n
Gender						
Female	46.9%	67	48.9%	140	48.3%	207
Male	53.1%	76	51.1%	146	51.7%	222
Ethnicity						
White	49.0%	70	47.9%	137	48.3%	207
Black	25.9%	37	26.2%	75	26.1%	112
Hispanic	12.6%	18	12.2%	35	12.4%	53
Asian	2.1%	3	1.4%	4	1.6%	7
Multi-racial	3.5%	5	3.5%	10	3.5%	15
Unreported	7.0%	10	8.7%	25	8.2%	35
English Learner						
Yes	2.1%	3	3.5%	10	3.0%	13
No	97.9%	140	96.5%	276	97.0%	416
Special Education Status						
Yes	11.2%	16	12.6%	36	12.1%	52
No	88.8%	127	84.2%	250	85.0%	377
Economically Disadvantaged						
Yes	41.3%	59	49.7%	142	46.9%	201
No	58.7%	84	50.3%	144	53.1%	228

Note. English Learner (EL) status was excluded from the matching model and subsequent analyses due to its low prevalence in the sample. Economically Disadvantaged was excluded from the matching model and subsequent analyses due to accuracy of reporting in virtual schools (Pondiscio, 2024). Both variables were tested for inclusion and, hence, were excluded because they did not improve the regression model's fit.

Following matching, regression analyses were performed on the propensity score–matched sample to verify the absence of statistically significant differences between the treatment and comparison groups. The matched sample satisfied WWC baseline equivalence criteria, with a Hedges' *g* effect size of .02, indicating comparable group characteristics prior to the intervention.

Table A2. Baseline Equivalence on Renaissance Star Math Assessment Scale Score (Completed Prior to Tutoring).

Outcome Variable	K12 Tutoring Adjusted Mean (*SD*)	*n*	Non-users Adjusted Mean (*SD*)	*n*	Hedges' *g*	*p*-value
Scale Score	954.15 (74.40)	143	952.20 (79.82)	286	0.02	.81

Appendix B. Additional Information on Results

The following section provides additional details on the analyses conducted for the Impact Research Questions.

To what extent did K12 Tutoring impact end-of-year (EOY) Star math outcomes relative to non-participating students?

A linear regression analysis was performed on the full sample across grade levels to examine the impact of K12 Tutoring participation on end-of-year (EOY) Star math scale scores. Tutoring status was included as a binary predictor, with non-participants serving as the reference group. The model accounted for baseline STAR math scores (continuous), as well as categorical covariates including school, grade, gender, ethnicity, and special education status. Although grade was included as a factor in the model, analyses were not conducted separately for each grade level due to insufficient sample sizes. These covariates were selected based on prior research and theoretical considerations to account for potential sources of variability in student outcomes and to ensure that the estimated effect of tutoring reflects the contribution of tutoring itself rather than differences in student or school characteristics. Models were also estimated without these covariates, and the patterns of results for the primary predictor remained consistent, but their inclusion ensures that the estimated effect of tutoring time on EOY math scores was not biased by differences in these student characteristics. Results showed that tutoring participation significantly predicted higher EOY math scores. The overall model explained 47% of the variance in growth scores, $F(18, 410) = 21.94$, $p < .001$. Full details of the regression model are presented in Table B1.

Table B1. End-of-year Star math scores as predicted by participation in K12 Tutoring.

	Unstandardized Beta (B)	Standard Error	t-value	p-value
Intercept	174.28	51.17	3.41	< .001
Tutoring	17.18	7.26	2.37	.02
BOY Score	0.85	0.06	15.24	< .001
School				
School 1-2	14.90	14.82	1.01	.32
School 1-3	4.22	15.38	0.28	.78
School 1-4	15.27	9.47	1.61	.11
School 1-5	6.87	20.87	0.33	.74
School 1-6	-102.49	41.50	-2.47	.01
Grade				
$2^{nd} - 3^{rd}$	11.84	17.58	0.67	.50
$2^{nd} - 4^{th}$	15.38	17.40	0.88	.38
$2^{nd} - 5^{th}$	-15.87	17.40	0.88	.37
$2^{nd} - 6^{th}$	1.97	17.34	0.11	.91
Gender	4.60	7.06	0.65	.51
Ethnicity				
African American	-11.98	8.74	-1.37	.17
Asian	16.76	28.82	0.58	.56
Hispanic/Latino	-4.36	12.57	-0.35	.73
Multi-racial	25.40	17.07	1.49	.14
Unreported	-29.10	18.32	-1.59	.11
SE	-16.84	11.11	-1.52	.13

Note. Reference groups were as follows: students who did not participate in tutoring (*Tutoring*), female students (*Gender*), White students (*Ethnicity*), and students not enrolled in special education programs (*SE*).