Supporting Early Literacy: Evidence of K12 Tutoring's Impact on Young Learners

ESSA Level II Report

July 2025

Contents

Executive Summary

To evaluate the effectiveness of its programs and services, K12 Tutoring conducted an impact study on student outcomes. The study was also structured to meet the criteria for Moderate Evidence (Level II) under the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

Using a quasi-experimental design with propensity score matching, the analysis compared outcomes for 372 students in grades 1-3 from five full-time virtual schools across Arkansas, Arizona, Michigan, and West Virginia during the 2024-2025 school year. Of these, 124 students participated in the tutoring program, while 248 served as matched comparisons based on demographic characteristics and fall assessment performance. The full sample was 48% female and included 44% White, 29% African American, 16% Hispanic, and 3% Asian students. Additionally, 57% of students qualified for free or reduced-price lunch, 18% received special education services, and 3% were English Language Learners.

The study examined tutoring participation alongside academic performance outcomes on the Renaissance Star Growth Assessment. Regression models, controlling for student characteristics, were used to determine the effect of tutoring compared to non-users. Baseline equivalence between treatment and comparison groups was confirmed according to What Works Clearinghouse (WWC) criteria for acceptable equivalence.

Program Implementation and Student Findings

Student Usage. Students received approximately 540 minutes of tutoring in small groups of fewer than four students during the 2024-2025 school year. No statistically significant differences in total time across grades emerged.

Student Outcomes. Students who received K12 Tutoring were matched with non-participants based on pre-tutoring assessment scores and student-level factors, resulting in groups that were academically and demographically similar at the outset. Analysis showed that students who participated in tutoring achieved higher end-of-year reading scores compared to their matched peers. Among tutoring participants, increased time in tutoring was linked to better outcomes for students who began the year with the lowest reading scores.

Conclusion

Students who participated in K12 Tutoring demonstrated greater academic growth than their peers. Based on the positive findings, this study meets ESSA Level II (Moderate Evidence) standards. It was a well-designed quasi-experimental study that ensured baseline equivalence, applied statistical controls, included over 350 students from multiple schools, and demonstrated at least one significant positive outcome.

Introduction

High-impact tutoring (HIT) is a structured, research-based approach designed to provide individualized academic support through regular, sustained sessions integrated into the school day (Robinson et al., 2021). It typically consists of one-on-one or small-group instruction, with no more than four students, delivered by consistent, trained tutors using high-quality materials aligned with classroom content (Kraft & Falken, 2021). Sessions are held at least three times per week, often during school hours to promote equitable access and reinforce core classroom instruction. Studies have demonstrated that HIT can lead to significant improvements in reading achievement, particularly in the early elementary grades, with reported effect sizes as high as 0.39 in 1st grade (Kraft et al., 2024; Nickow et al., 2024). Its success is attributed to its focus on relationship-building, data-driven instruction, and frequent progress monitoring, all of which are essential for accelerating learning among students who are furthest behind (Hashim et al., 2024).

In reading, the early elementary years are a pivotal period for developing foundational skills that serve as the building blocks for later higher-order comprehension processes (Snow et al., 1998). Timely literacy-focused tutoring during this stage, when students are solidifying foundational reading skills that prepare them to transition from learning to read to reading to learn, may accelerate reading development and prevent long-term deficits (Hernandez, 2011; Lesnick et al., 2010). Tutoring in the early grades has been shown to significantly improve key components of early reading, including oral reading fluency (Cortes et al., 2025) and phonemic awareness (Allor & McCathren, 2004). These findings highlight the potential of early intervention to alter students' academic trajectories and reduce the risk of persistent reading struggles.

K12 Tutoring is a structured academic support program that offers personalized, curriculum-aligned instruction in core subjects through certified tutors. In early literacy, K12 implements the principles of the science of reading to build foundational skills and support reading development in the primary grades.

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 Questions

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

Impact Research Questions

2. To what extent did K12 Tutoring impact end-of-year (EOY) Star reading outcomes relative to non-participating students?
3. How did students' end-of-year (EOY) Star reading outcomes differ according to the total time they spent in tutoring?

Method

This section of the report provides a concise overview of the study's design, setting, participant details, measurement tools, analytical methods, and baseline equivalence procedures.

Study Design

To meet ESSA Level II evidence standards, this study used a quasi-experimental design with propensity score matching. The treatment group included students who received K12 Tutoring during the 2024-2025 school year, while the comparison group consisted of students with similar demographic characteristics and fall assessment performance who did not participate in the tutoring program.

Setting and Participants

This quasi-experimental study was conducted during the 2024-2025 academic year and analyzed a matched sample of 372 students in grades 1 through 3 from five full-time virtual schools located in Arkansas, Arizona, Michigan, and West Virginia. Although each school operated under separate administrative oversight, the tutoring model was implemented consistently across sites. Of the total sample, 124 students received tutoring, while the remaining 248 served as matched comparisons. The full sample was 48% female. Racial and ethnic demographics included 44% white, 29% African American, 16% Hispanic, 3% Asian, and 3% identifying as multiracial; 6% of students did not report a racial or ethnic background. Additionally, 57% of students qualified for free or reduced-price lunch, 18% received special education services, and only 3% were identified as English Language Learners.

Measures

This study incorporated multiple measures to examine the effect of K12 Tutoring on student performance outcomes.

K12 Tutoring Participation Metrics.

Student-level data on participation were gathered, capturing the total number of minutes students attended tutoring. These metrics were analyzed to understand differences in student engagement with K12 Tutoring during the 2024–2025 school year and to explore potential relationships between participation levels and academic outcomes.

Student Outcomes.

Administrative records, including scale scores from the Renaissance Star Reading Assessment (Star), were used to evaluate academic outcomes. Students completed the assessment at both the beginning (BOY) and end (EOY) of the school year to measure growth.

Data Analysis

Descriptive statistics were used to summarize the characteristics of the study sample and to support the analysis of implementation patterns. To assess the impact of K12 Tutoring on student outcomes, regression models were employed, controlling for student-level variables and using nearest neighbor propensity score matching to reduce the risk of selection bias. Standardized effect sizes were also

computed to quantify the magnitude of outcome differences between students who received tutoring and those with comparable demographic and academic backgrounds who did not.

Baseline Equivalence

To support the study's validity and comply with ESSA Level II criteria, baseline equivalence between matched treatment and comparison groups was assessed. Details of these analyses are presented in the Results section and further documented in Appendix A.

Results

Implementation Research Findings

Utilization data were examined to assess how extensively students engaged with K12 Tutoring during the 2024–2025 academic year.

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

The overall median duration of tutoring received by students was 540 minutes. Although differences in total time spent in tutoring varied across grade levels, these differences were not statistically significant, $p > .05$ (1st grade = 615 minutes; 2nd grade = 465 minutes; 3rd grade = 525 minutes). Tutoring sessions were delivered throughout the 2024–2025 academic year in small group formats, consisting of fewer than four students per tutor.

Impact Research Findings

To address the impact research questions, regression analyses were performed comparing matched samples of K12 Tutoring participants and non-participants, as well as analyses restricted to K12 Tutoring users. The following section presents findings related to (a) the impact of K12 Tutoring on student learning outcomes, and (b) the relationship between time spent in tutoring and these outcomes. Statistical significance is reported only for results meeting the $p < .05$ criterion.

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

A matched sample was created by pairing students who participated in K12 Tutoring with those who did not, using pre-tutoring Star scores, school, grade, gender, ethnicity, and special education status as matching variables. Baseline equivalence between the groups was confirmed, with a Hedges' g of 0.04 and a mean difference of 5.22 scale score points, indicating similar Star scores prior to tutoring.

Using the matched sample, a linear regression analysis was conducted with tutoring status coded as a binary variable to evaluate the effect of K12 Tutoring on EOY Star scores. The model controlled for BOY Star scores, school, grade, gender, ethnicity, and special education status. Results indicated that students who participated in K12 Tutoring achieved significantly higher EOY Star scores compared to non-users.

This difference was statistically significant (Ω^2 = 0.01, p = .006; see Figure 1). The effect size indicates that tutored students scored, on average, 0.20 standard deviations higher than their matched peers, corresponding to a 7.53 percentile gain.

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

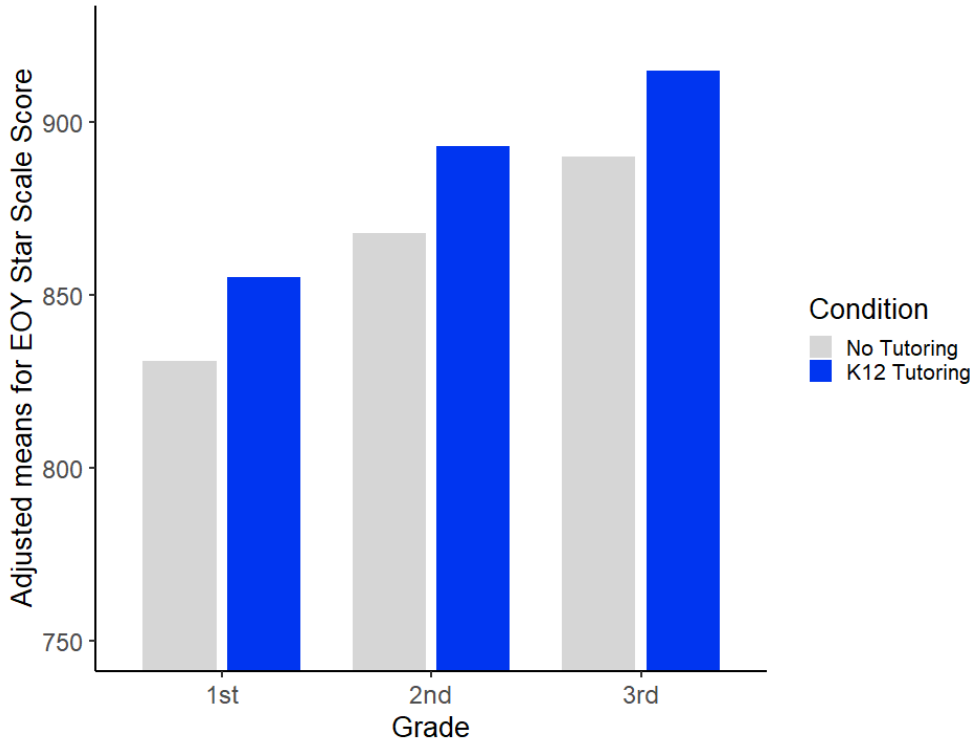

Figure 1. Adjusted means for end-of-year Star Growth reading assessment scale scores for K12 Tutoring and non-users. Mean difference was statistically significant (Ω^2 = 0.01, p = .006, n = 372).

How did students' end-of-year (EOY) Star reading outcomes differ according to the total time they spent in tutoring?

Regression analyses were conducted on K12 Tutoring students to examine how varying intensities of tutoring usage influenced EOY Star scores. This model used the same covariates as the primary analysis. To test whether outcomes varied by tutoring usage across different student groups, an interaction term between the hours spent in tutoring and BOY scale score was included. While time spent in tutoring was a significant predictor of EOY Star score for all students, t = 2.63, p = .01, this results was subsumed by the significant interaction, t = -2.53, p = .01. As shown in Figure 2, follow-up analyses indicated that spending more time in tutoring was associated with higher EOY scores for students with the lowest BOY

scores[1]. This suggests that consistent, sustained engagement in tutoring is especially critical for students who begin the year with the greatest academic needs.

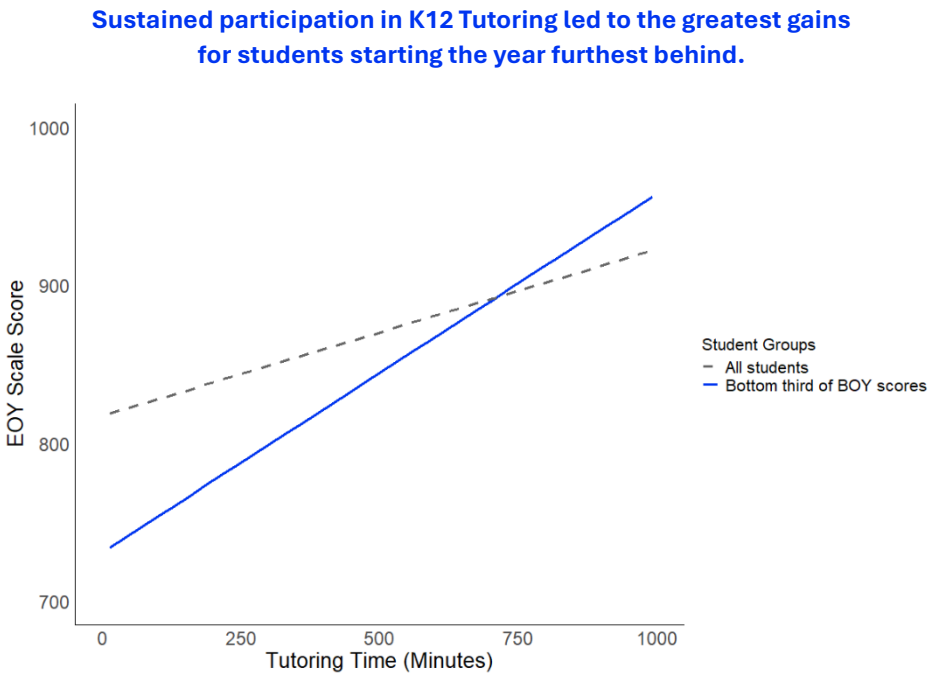

Figure 2. Influence of time spent in K12 Tutoring and BOY scale scores on EOY Star Growth reading assessment scale scores.

Conclusions

Findings from this study demonstrate that students who participated in K12 Tutoring outperformed their matched peers on an EOY Star Reading Assessment, even when accounting for prior achievement and demographic characteristics. In addition to this overall benefit, results showed that the amount of time spent in tutoring played a meaningful role in student outcomes. Specifically, students who began the year with lower reading proficiency experienced greater gains when they engaged in tutoring more consistently over time. Overall, these findings highlight that targeted and consistent tutoring can be an effective strategy for closing achievement gaps and supporting struggling learners.

[1] BOY Star scores were included in the model as a continuous variable. For visualization purposes, Figure 2 highlights the effect of tutoring time on end-of-year scale scores for all students (dotted line), with the solid blue line representing only students in the bottom tertile (bottom third of BOY scale scores). A simple slopes analysis indicated that the slope for students in the lowest tertile was significantly different from zero (p = .04).

Given multiple positive outcome findings, this study provides results to satisfy ESSA evidence requirements for Level II (Moderate Evidence). Specifically, this quasi-experimental study met the following criteria for Level II:

- ✓ 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

Allor, J., & McCathren, R. (2004). The efficacy of an early literacy tutoring program implemented by college students. *Learning Disabilities Research & Practice*, *19*(2), 116-129.

Cortes, K. E., Kortecamp, K., Loeb, S., & Robinson, C. D. (2025). A scalable approach to high-impact tutoring for young readers. *Learning and Instruction*, *95*, 102021.

Hashim, A. K., Davison, M., Postell, S., & Isaacs, J. (2024). *High-dosage tutoring for academically at-risk students*. NWEA.

Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Annie E. Casey Foundation.

Kraft, M. A. & Falken. G. T. (2021). A blueprint for scaling tutoring and mentoring across public schools. *AERA Open, 7*. https://doi.org/10.1177/23328584211042858

Kraft, M. A., Schueler, B. E., & Falken. G. (2024). *What impacts should we expect from tutoring at scale? Exploring meta-analytic generalizability.* EdWorkingPaper, 24-1031. Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/zygj-m525

Lesnick, J., Goerge, R., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How is it related to high school performance and college enrollment?* Chapin Hall at the University of Chicago.

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.

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

Robinson, C. D., & Loeb, S. (2021). High-impact tutoring: State of the research and priorities for future learning. *National Student Support Accelerator, 21*(284), 1-53.

Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children* (pp. 207-210). National Academy Press.

Appendix A. Additional Information on Study Design and Methods

A total of 124 K12 Tutoring students in grades 1-3 had complete demographic, achievement, and usage data. In addition, a sample of comparison students with complete demographic and achievement data was obtained. Using this dataset of K12 Tutoring students and non-users, propensity score matching (i.e., nearest neighbor matching without replacement, 2:1 ratio) was used to create a matched study sample of K12 Tutoring students and 248 non-users with similar academic and demographic profiles. Matching variables included pre-tutoring Star scores, school, grade, gender, ethnicity, and special education status. Consequently, the matched study sample included 372 students. Additional demographic information on participating students is included in Table A1.

Table A1. Student demographics by group for matched sample.

Characteristic	K12 Tutoring (n = 124)		Non-users (n = 248)		Total Sample (N = 372)	
	Percent	*n*	Percent	*n*	Percent	*n*
Gender						
Female	47%	58	49%	122	48%	180
Male	53%	66	51%	126	52%	192
Ethnicity						
White	46%	57	43%	107	44%	164
Black	27%	33	30%	74	29%	107
Hispanic	15%	18	16%	40	16%	58
Asian	3%	4	2%	6	3%	10
Multi-racial	3%	4	3%	8	3%	12
Unreported	7%	8	5%	13	6%	21
English Learner (EL Status)						
Yes	2%	3	3%	8	3%	11
No	98%	121	97%	240	97%	361
Special Education Status						
Yes	9%	11	23%	56	18%	67
No	91%	113	77%	192	75%	305
Economically Disadvantaged						
Yes	48%	60	61%	151	57%	211
No	52%	64	39%	97	39%	161

Note. English Learner (EL) status was excluded from the matching model and subsequent analyses due to its low prevalence in the sample; testing its inclusion showed no improvement in the regression model's fit. Economically Disadvantaged was excluded from the matching model and subsequent analyses due to accuracy of reporting in virtual schools (Pondiscio, 2024).

After matching across groups, regression analyses were conducted using the propensity score matched sample to confirm there were no statistically significant differences between student groups. The matched sample met WWC baseline equivalence standards with Hedges' *g* effect size = 0.04.

Table A2. Baseline Equivalence on Renaissance Star Reading 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	823.97 (118.88)	124	829.19 (122.58)	248	0.04	.70

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 reading outcomes relative to non-participating students?

A linear regression model was conducted on the full sample across grade levels to evaluate the impact of K12 Tutoring participation on EOY Star reading scale score. Tutoring participation was included as a dummy-coded predictor, with non-users serving as the reference group. The model also controlled for pre-tutoring STAR reading scores (continuous predictor), as well as school, grade, gender, ethnicity, and special education status (nominal predictors), though analyses were not disaggregated by grade level. Tutoring was a significant positive predictor of EOY Star scores, with the full model accounting for 59% of the variance in growth scores, $F(15, 356) = 36.19$, $p < .001$. See Table B1 for the full model.

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

	Unstandardized Beta (B)	Standard Error	*t*-value	*p*-value
Intercept	335.85	36.35	9.24	< .001
Tutoring	24.80	8.93	2.78	.01
BOY Score	0.63	0.04	14.71	< .001
School				
School 1-2	-29.58	17.13	-1.73	.09
School 1-3	-3.71	11.93	-0.31	.76
School 1-4	28.13	23.83	1.18	.24
School 1-5	-31.83	23.84	-1.34	.18
Grade				
1st-2nd	36.96	10.98	3.37	< .001
1st-3rd	58.57	12.34	4.74	< .001
Gender	-15.63	8.72	-1.79	.07
Ethnicity				
African American	-13.38	10.82	-1.24	.22
Asian	23.88	24.65	0.97	.33
Hispanic/Latino	28.48	13.33	2.14	.03
Multi-racial	48.46	22.42	2.16	.03
Unreported	-22.19	23.70	-0.94	.35
SE	-49.40	11.52	-4.29	< .001

Note. The reference group for Tutoring was non-users. The reference group for Ethnicity was White.

How did students' end-of-year (EOY) Star reading outcomes differ according to the total time they spent in tutoring?

A linear regression model was conducted to examine the impact of K12 Tutoring usage on EOY Star reading scale score among tutoring participants. Tutoring usage was included as a continuous predictor in hours. The model also controlled for pre-tutoring STAR reading scores (continuous predictor), as well as school, grade, gender, ethnicity, and special education status (nominal predictors). As shown in Table B2, the significant interaction indicates that spending more time in tutoring was associated with higher EOY scores for students with the lowest BOY scores, with the full model explaining 56% of the variance in growth scores, $F(16, 107) = 10.87$, $p < .001$.

Table B2. End-of-year Star reading scores as predicted by tutoring usage.

	Unstandardized Beta (B)	Standard Error	*t*-value	*p*-value
Intercept	303.54	70.04	4.33	< .001
Time * BOY Score	-.005	.002	-2.53	.01
Time	4.47	1.70	2.63	.01
BOY Score	0.69	0.08	8.22	< .001
School				
School 1-2	-92.64	36.08	-2.57	.01
School 1-3	-12.21	22.04	-0.55	.58
School 1-4	-105.00	50.72	-2.07	.04
School 1-5	-102.96	42.59	-2.42	.02
Grade				
1st-2nd	35.58	20.22	1.76	.08
1st-3rd	58.57	12.34	4.74	< .001
Gender	68.30	22.88	2.99	.003
Ethnicity				
African American	10.63	18.73	0.57	.57
Asian	-10.95	54.27	-0.20	.84
Hispanic/Latino	80.64	23.75	3.40	< .001
Multi-racial	0.43	42.51	0.01	.99
Unreported	60.70	44.52	1.36	.18
SE	-17.14	26.24	-0.65	.52