Paper ESSA Level III Study (2021 - 2022)

Prepared for:
Paper

Prepared by LearnPlatform:
Meetal Shah, Ph.D., Researcher, LearnPlatform
Mary Styers, Ph.D., Director of Research, LearnPlatform

March 2, 2023

EXECUTIVE SUMMARY

Paper contracted with LearnPlatform, a third-party edtech research company, to examine the relationship between student usage of Paper's tutoring platform and learning outcomes. LearnPlatform designed the study to satisfy Level III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA).

Study Sample, Measures, and Methods

This study occurred during the 2021-22 school year and analyses included 2,878 high school students (Grades 10-12) across four schools in one school district.

Researchers used multiple measures to provide insights into the implementation and evidence of Paper's potential impacts on student outcomes. Paper provided LearnPlatform with usage metrics. The school district provided student demographic data and standardized course assessment scores, which researchers used to measure students' math and English Language Arts (ELA) achievement outcomes.

LearnPlatform researchers used a variety of quantitative analytic approaches. Specifically, we used descriptive statistics to examine participant characteristics and support analyses of implementation, as well as regression analyses to investigate how use of Paper impacts students' achievement on course assessments. The analyses included student-level covariates to control for potential selection bias. In addition, researchers calculated standardized effect sizes to determine the magnitude of changes in math and ELA outcomes at mid- and end-of-year.

Key Takeaways

- Students who used Paper tutoring sessions had higher midyear ELA achievement compared to demographically similar students who did not use the program for the following courses:
    - American Literature
    - Studies in British Literature
- Students enrolled in the Studies in British Literature course who used Paper's essay review feature had higher midyear ELA achievement compared to demographically similar students who did not use the feature.
- Students enrolled in the American Literature course who used Paper's essay review feature had lower midyear ELA achievement compared to demographically similar students who did not use the feature.

Conclusions

Given positive outcome findings, this study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence). Specifically, this study met the following criteria for Level III:

- ✓ Comparative study with non-standardized outcome measure
- ✓ Proper design and implementation
- ✓ Statistical controls through covariates
- ✓ At least one statistically significant, positive correlation with statistical controls for selection bias

TABLE OF CONTENTS

Introduction

Paper contracted LearnPlatform, a third-party edtech research company, to examine the relationship between student usage of its tutoring platform and learning outcomes. LearnPlatform designed the study to satisfy Level III requirements (Promising Evidence) according to the Every Student Succeeds Act (ESSA).

Paper recognizes that many academic support options–especially in-person, on-demand tutoring–have equity, cost, and scalability limitations. However, on-demand tutoring can be a powerful tool for narrowing learning gaps exacerbated by the pandemic. Paper offers schools a cost-effective, 24/7 online tutoring option via their educational support system (see logic model in Appendix A; Shah & Styers, 2022).

The present study had the following research questions:

Implementation Questions

1. To what extent did high school students utilize Paper during the 2021–2022 school year:
    - How many math and ELA tutoring sessions were completed?
    - How much time was spent in math and ELA tutoring sessions?
    - How many ELA essays were submitted for review?

Outcome Questions

2. How did mid- and end-of-year math and ELA achievement outcomes of high school students who used Paper tutoring sessions during the 2021–22 school year compare to students who did not use the program? What was the magnitude of any observed difference?
3. How did mid- and end-of-year ELA achievement outcomes of high school students who used Paper's essay review feature during the 2021–22 school year compare to students who did not use the feature? What was the magnitude of any observed difference?

This report details the study design and methods, implementation, findings, conclusions, and recommended next steps.

Study Design and Methods

This section of the report briefly describes the study's design, setting, participants, measures, and analysis methods.

Study Design

This study used a quasi-experimental design with propensity score weighting (see Appendix B for more information about the propensity score weighting procedures used in this study) to allow comparisons of math and ELA outcomes among students who used Paper (i.e., tutoring sessions or essay review feature) during the 2021–2022 school year and students who did not use Paper, the study included comparison students.

Setting

This study included data from the 2021–22 school year and included 2,878[1] students enrolled in math and/or ELA courses across four high schools in one Illinois school district. Table 1 shows overall samples by Paper usage.

Participants

Table 1. Sample sizes by Paper use

	Number of Paper users	Number of non-users	Total
Math tutoring sessions	865	1,653	2,518
ELA tutoring sessions	611	1,696	2,307
ELA essay review	955	1,571	2,526

Measures

This study includes multiple measures to provide insights into Paper's implementation and evidence about the potential impacts of the learning solution on student achievement.

[1] Students used one or more of Paper's features, i.e., math tutoring sessions, ELA tutoring sessions, and ELA essay review, which is why the overall total in Table 1 does not equate to the overall sample of 2,878 students.

Paper Usage Metrics. Researchers utilized 2021–22 student-level usage (i.e., total number of tutoring sessions, total amount of time in minutes spent in Paper tutoring sessions, and total number of essays submitted for review). Usage data informed the extent to which students used Paper during the school year and whether students' use of the product related to students' math and ELA outcomes. Tutoring sessions were coded by subject area, and therefore distinct usage data for math and ELA tutoring sessions were used in the analyses for math and ELA outcomes, respectively.

District Course Assessments. The district's course-specific assessments were used to assess students' math and ELA outcomes. These assessments do not have reliability or validity information, and as a result, are not eligible to meet ESSA Level II standards. The district has nine math courses and seven courses included assessment data from three different time points: beginning of year (i.e., fall 2021), midyear (i.e., winter 2021), and end-of-year (i.e., spring 2022). The district has six ELA courses and each course included assessment data from three different time points: beginning of year (i.e., fall 2021), midyear (i.e., winter 2021), and end-of-year (i.e., spring 2022). Researchers disaggregated the overall sample by course because course achievement measures were not measured on a vertical scale. Therefore, the analytic sample differed for each course.

Data Analysis

Researchers used a variety of quantitative analytic approaches to answer the research questions. First, researchers used descriptive statistics to examine participant characteristics and support analyses of implementation data. Then, researchers conducted regression analyses to investigate how use of Paper impacted student math and ELA achievement on course assessments in one school year. Analyses included student-level covariates to control for potential selection bias. In addition, researchers calculated standardized effect sizes to determine the magnitude of changes in student outcomes or the standardized difference between student groups' outcomes (i.e., difference in performance between Paper users and non-users).

Baseline Equivalence

To ensure the validity of the study's findings and to adhere to WWC quasi-experimental study standards, the researchers assessed the equivalence of student demographic characteristics (i.e., race, Individualized Educational Plan (IEP) status, and socioeconomic status (SES)) and standardized assessment scores between student groups (i.e., students who used Paper and students who did not use Paper). All baseline differences were below the 0.25 effect size

threshold, but above 0.05, and as a result, pretest scores were statistically controlled for in the final model. See Appendix B for more details regarding baseline equivalence.

Program Implementation

The charts below highlight Paper use during the 2021-22 school year based on internal usage data. Overall, students completed an average of three math tutoring sessions (SD = 4.5, range 1 – 41) and two ELA tutoring sessions (SD = 1.3, range 1 – 15). On average, students spent 101 total minutes on math tutoring sessions in Paper (SD = 217, range 1 – 2029 minutes) and 39 total minutes on ELA tutoring sessions in Paper (SD = 66, range 1 – 1,153 minutes). Students submitted two ELA essays (SD = 1.3, range 1 – 14) for review to Paper tutors. Detailed average use by math and ELA course can be found in Appendix C.

Average Total Paper Usage in 2021-22 School Year

	Average Use
Number of Paper Math tutoring sessions	3
Time spent (minutes) on Paper Math tutoring sessions	101
Number of Paper ELA tutoring sessions	2
Time spent (minutes) on Paper ELA tutoring sessions	39
Number of ELA essays submitted for review	2

Study Findings

To answer the study research questions, researchers conducted regressions analysis. Regressions included propensity score weighting. All findings are statistically significant at the $p < .05$ and include standardized effect sizes to assist with interpretation.

?

How did mid-year and end-of-year math and ELA achievement outcomes of high school students who used Paper tutoring sessions during the 2021–22 school year compare to students who did not use the program? What was the magnitude of any observed difference?

In the following section, researchers examined whether there were any differences between students who used Paper tutoring sessions during the 2021–22 school year and students who did not use the tutoring sessions. Of note, Appendix D provides additional information on these analyses and findings.

To determine whether there were differences between students who used Paper tutoring sessions and students who did not use them, researchers conducted regression analysis with propensity score weighting and included midyear or end-of-year course achievement scores as the outcome of interest. These models included IEP status and/or race as student-level covariates.

Differences Between High School Students who used Paper Tutoring Sessions and High School Students Who Did Not Use the Program at Midyear

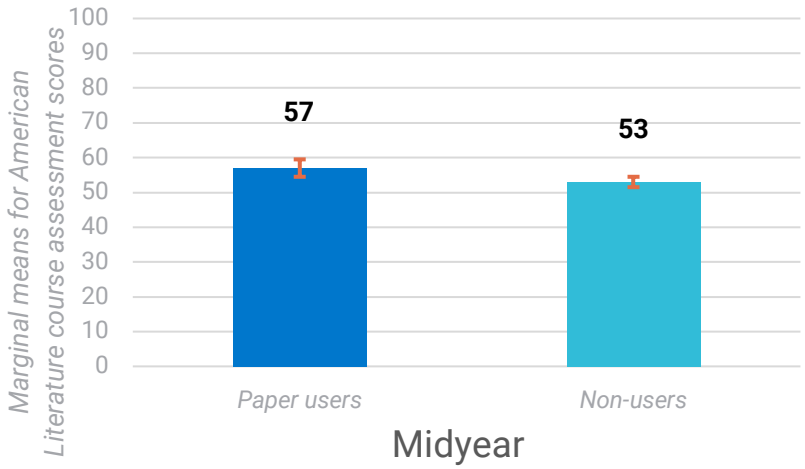
The first set of regression models, with propensity score weighting, included midyear course achievement scores as the outcome of interest.

Results show that students enrolled in the American Literature course (Key Finding 1) and the Studies in British Literature course (Key Finding 2) who used Paper tutoring sessions, had statistically significantly higher scores on midyear course assessments compared to demographically similar students who did not use them.

There were no other statistically significant differences in scores between students who used Paper tutoring sessions and those who did not use them on any other midyear math or ELA course assessments. In other words, Paper tutoring users and non-users had similar midyear performance in the following courses: Language Arts, Intro Extension English, Studies in English,

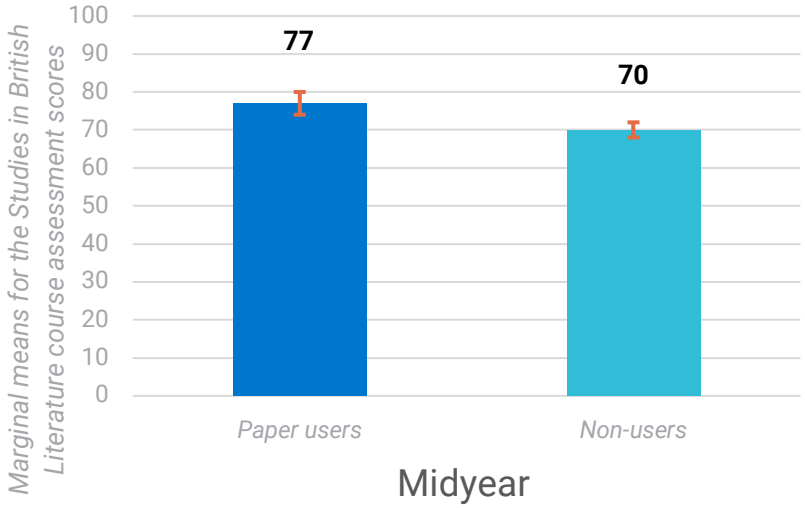
Studies in American Literature, Math 2, Math 3, Math 3A, and Honors Math 2 (see Appendix D for additional details).

Key Finding 1. Students enrolled in the American Literature course who used Paper ELA tutoring sessions had higher scores on the midyear course assessment than students who did not use the program (effect size = 0.19). If a comparison student, at the 50th percentile, had used Paper, they would have been expected to perform at the 58th percentile.



Note: The orange vertical lines at the top of each bar represent a 95% confidence interval.

Key Finding 2. Students enrolled in the Studies in British Literature course who used Paper ELA tutoring sessions had higher scores on the midyear course assessment than students who did not use the program (effect size = 0.54). If a comparison student, at the 50th percentile, had used Paper, they would have been expected to perform at the 71st percentile.



Note: The orange vertical lines at the top of each bar represent a 95% confidence interval.

Differences Between High School Students who used Paper Tutoring Sessions and Students Who Did Not Use the Program at the End of the Year

The second set of regression models, with propensity score weighting, included end-of-year course achievement scores as the outcome of interest.

There were no statistically significant differences in scores between students who used Paper tutoring sessions and those who did not use them on any end-of-year math or ELA course assessments. In other words, Paper tutoring users and non-users had similar end-of-year performance in the following courses: Language Arts, American Literature, Intro Extension English, Studies in British Literature, Studies in English, Studies in American Literature, Math 1, Math 2, Math 3, Math 3A, Honors Math, and Honors Math 2 (Appendix D).

How did mid- and end-of-year ELA achievement outcomes of high school students who used Paper's essay review feature during the 2021–22 school year compare to students who did not use the feature? What was the magnitude of any observed difference

In the following section, researchers examined whether there were any differences between students who used Paper's essay review feature during the 2021–22 school year and students who did not use the feature. Of note, Appendix D provides additional information on these analyses and findings.

To determine whether there were differences between students who used Paper's essay review feature and students who did not use the feature, researchers conducted regression analysis with propensity score weighting and included midyear or end-of-year ELA course achievement scores as the outcome of interest. These models included IEP status and/or race as student-level covariates.

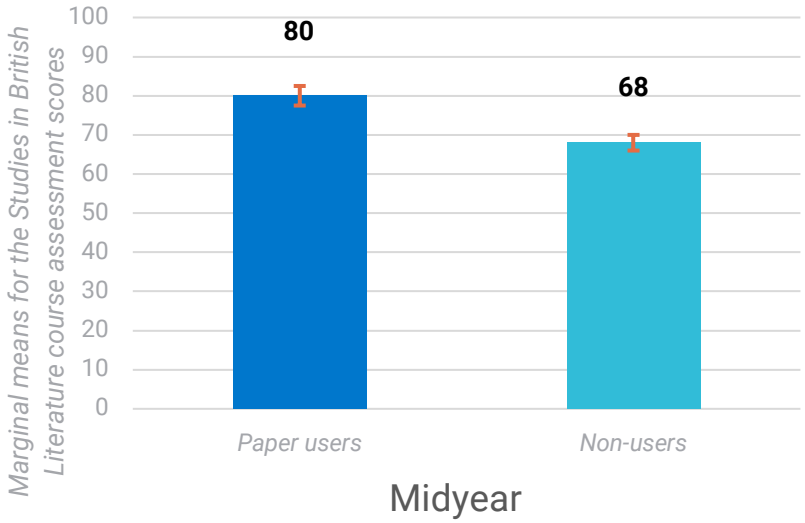
Differences Between High School Students who used Paper's Essay Review Feature and High School Students Who Did Not Use the Feature at Midyear

The first set of regression models, with propensity score weighting, included midyear ELA course achievement scores as the outcome of interest.

Results show that students enrolled in the Studies in British Literature course (Key Finding 3) who used Paper's essay review feature, had statistically significantly higher scores on midyear course assessments compared to demographically similar students who did not use the feature. However, students enrolled in the American Literature course (Key Finding 4) who used Paper's essay review feature, had statistically significantly lower scores on midyear course assessments compared to demographically similar students who did not use the feature.

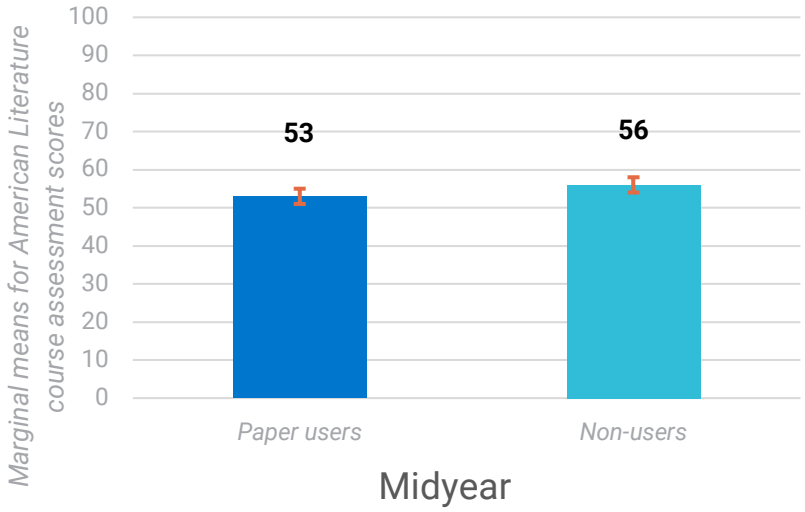
There were no other statistically significant differences in scores between students who used Paper's essay review feature and those who did not use the feature on any other midyear ELA course assessments. In other words, Paper essay review users and non-users had similar midyear performance in the following courses: Language Arts, Intro Extension English, Studies in English, and Studies in American Literature (see Appendix D for additional details).

Key Finding 3. Students enrolled in the Studies in British Literature course who used Paper's essay review feature had higher scores on the midyear course assessment than students who did not use the program (effect size = 0.94). If a comparison student, at the 50th percentile, had used Paper, they would have been expected to perform at the 83rd percentile.



Note: The orange vertical lines at the top of each bar represent a 95% confidence interval.

Key Finding 4. Students enrolled in the American Literature course who used Paper's essay review feature had lower scores on the midyear course assessment than students who did not use the program (effect size = - 0.16).



Note: The orange vertical lines at the top of each bar represent a 95% confidence interval.

Differences Between High School Students who used Paper's Essay Review Feature and Students Who Did Not Use the Feature at the End of the Year

The second set of regression models, with propensity score weighting, included end-of-year ELA course achievement scores as the outcome of interest.

There were no statistically significant differences in scores between students who used Paper's essay review feature and those who did not use the feature on any end-of-year ELA course assessments. In other words, Paper tutoring users and non-users had similar end-of-year performance in the following courses: Language Arts, American Literature, Intro Extension English, Studies in British Literature, Studies in English, and Studies in American Literature (Appendix D).

Conclusions & Recommendations

Given positive outcome findings, this study provides results to satisfy ESSA evidence requirements for Level III (Promising Evidence).

Researchers recommend the following next steps:

- Increased usage of Paper by the district will enable third parties the ability to better understand Paper's impact on academic outcomes.
- There were statistically significant, positive differences in ELA outcomes between students who used Paper tutoring sessions and students who did not in some ELA courses. Further, there was a statistically significant, positive difference in ELA outcomes between students who used Paper's essay review feature and students who did not in one ELA course. Therefore, this information can be used to support implementation among other courses and subject areas.
- Paper should consider recruiting districts that implement standardized assessments with validity and reliability data to meet WWC's standards for assessment outcomes (ESSA Level II).

Acknowledgements

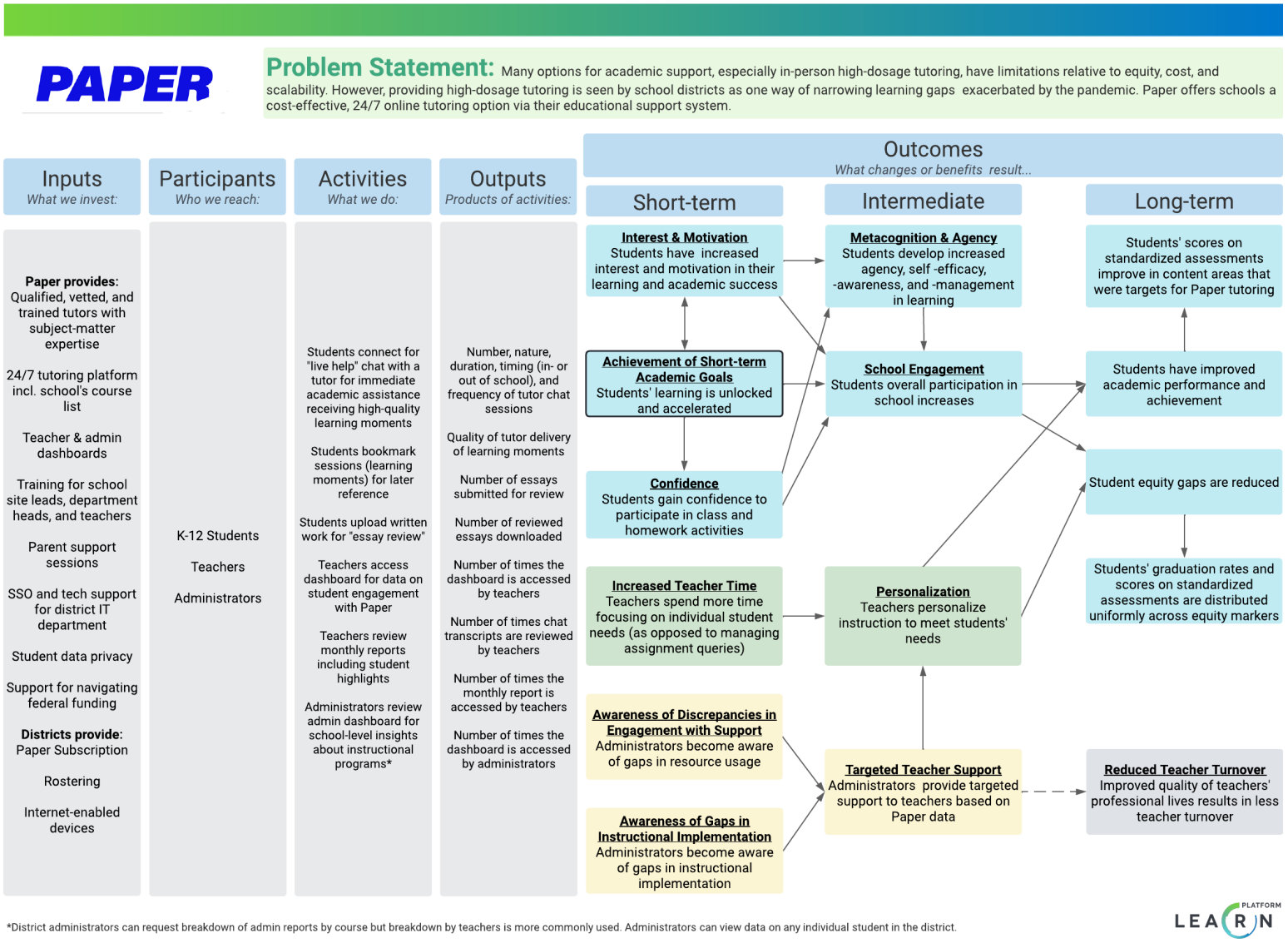
The authors would like to extend their deepest thanks to Avery Wall for her data preparation and editorial review.

References

Shah, M. & Styers, M. 2022. *Paper Logic Model. Study Type: ESSA Evidence Level IV*. LearnPlatform.

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 (NCEE). This report is available on the What Works Clearinghouse website at https://ies.ed.gov/ncee/wwc/Handbooks.

Appendix A. Paper Logic Model



*District administrators can request breakdown of admin reports by course but breakdown by teachers is more commonly used. Administrators can view data on any individual student in the district.

Appendix B. Additional Information on Study Design and Methods

Additional Information on Participating Schools

The present study included four high schools in one large suburban public school district in Illinois. Table B1 documents NCES school-level demographic data for the participating high schools.

Table B1. Description of participating high schools

	School A	School B	School C	School D
Total students at the school	1,425	1,311	1,349	1,110
% American Indian/Alaskan Native	0	0	0	0
% Asian	1	0	5	3
% Black	24	91	9	27
% Hispanic or Latino	60	5	28	23
% Native Hawaiian/Pacific Islander	0	0	0	0
% White	13	0	54	42
% Two or more races	3	4	4	5

Source: 2020-2021 data retrieved from IES, NCES Common Core of Data https://nces.ed.gov/ccd/

Propensity Score Weighting

To help make the student groups (i.e., students who used Paper and students who did not use Paper) as comparable as possible, propensity score weights were calculated for each student. To calculate propensity scores, researchers conducted binary logistic regression with student group as the dependent variable and SES, IEP status, race, and fall course assessment scores as the covariates. The probability was saved as a new variable. Weights were calculated by finding the inverse of the probability (1/probability). Students without a weight were dropped from the final analytic sample.

Baseline Equivalence

Researchers conducted baseline equivalence analyses to determine whether there were baseline differences in characteristics between students who used Paper and students who did not use the program during the 2021–22 school year. Specifically, researchers used chi-square analyses on student-level demographics and regression on pretest scores. There were some statistically significant differences between groups for the two ELA courses, but the magnitude of the effect size was within the bounds for acceptable statistical adjustments in analyses (Tables B2 – B9). [2]

Table B2. Baseline Equivalence Analysis of Race by Student Group for American Literature

Race	Student who used Paper (*n* = 151)		Students who did not use Paper (*n* = 410)				
	%	*N*	%	*N*	Chi-Squared	*p*-Value	Effect Size
Hispanic	21	31	37	153	17.97	0.001**	0.18
American Indian or Alaska Native	-	-	-	-			
Asian	2	3	1	4			
African American	37	56	30	121			
Native Hawaiian or Other Pacific Islander	-	-	-	-			
White	34	52	30	123			
Two or more races	6	9	2	9			

[2] Baseline differences with an effect size between 0.05 and 0.25 must include acceptable statistical adjustments in analyses.

Table B3. Baseline Equivalence Analysis of SES by Student Group for American Literature

SES	Students who used Paper (*n* = 151)		Students who did not use Paper (*n* = 410)		Chi-Squared	*p*-Value	Effect Size
	Percent	*N*	Percent	*N*			
Low Income	28	43	28	113	0.05	0.830	0.01
Other	72	108	72	297			

Table B4. Baseline Equivalence Analysis of IEP status by Student Group for American Literature

IEP	Students who used Paper (*n* = 151)		Students who did not use Paper (*n* = 410)		Chi-Squared	*p*-Value	Effect Size
	Percent	*N*	Percent	*N*			
IEP	3	4	16	65	17.84	0.00***	-0.18
Other	97	147	84	345			

Table B5. Baseline Equivalence Analysis of the American Literature Beginning-of-Year Scores by Student Group

Outcome Variable	Coefficient	Standard Error	*t*-value	*p*-value	Effect Size
Beginning-of-Year Course Score	-0.043	0.013	-3.44	0.001**	0.02

Table B6. Baseline Equivalence Analysis of Race by Student Group for Studies in British Literature

Race	Student who used Paper (*n* = 85)		Students who did not use Paper (*n* = 167)		Chi-Squared	*p*-Value	Effect Size
	%	*N*	%	*N*			
Hispanic	24	20	35	58	14.53	0.006**	0.24
American Indian or Alaska Native	-	-	-	-			
Asian	7	6	0	0			
African American	26	22	22	36			
Native Hawaiian or Other Pacific Islander	-	-	-	-			
White	41	35	41	69			
Two or more races	2	2	2	4			

Table B7. Baseline Equivalence Analysis of SES by Student Group for Studies in British Literature

SES	Students who used Paper (*n* = 85)		Students who did not use Paper (*n* = 167)				
	Percent	*N*	Percent	*N*	Chi-Squared	*p*-Value	Effect Size
Low Income	18	15	11	19	1.90	0.168	0.09
Other	82	70	88	148			

Table B8. Baseline Equivalence Analysis of IEP status by Student Group for Studies in British Literature

IEP	Students who used Paper (*n* = 85)		Students who did not use Paper (*n* = 167)				
	Percent	*N*	Percent	*N*	Chi-Squared	*p*-Value	Effect Size
IEP	0	0	0	0	-	-	-
Other	100	85	100	167			

Table B9. Baseline Equivalence Analysis of Studies in British Literature Beginning-of-Year Scores by Student Group for Studies in British Literature

Outcome Variable	Coefficient	Standard Error	*t*-value	*p*-value	Effect Size
Beginning-of-Year Course Score	-0.031	0.020	-1.54	0.126	0.01

Appendix C. Additional Information on Implementation

Table C1. Total Number of Paper Tutoring Sessions by ELA Course

ELA Course	*N*	mean	SD	Min	Max
Language Arts	95	1.5	1.1	1	8
American Literature	151	1.5	.83	1	6
Intro Extension English	131	1.9	1.8	1	15
Studies in British Literature	85	2.0	1.3	1	7
Studies in English	123	2.0	1.3	1	7
Studies in American Literature	25	1.8	1.7	1	8

Table C2. Total Time in Minutes Spent on Paper Tutoring Sessions by ELA Course

ELA Course	*N*	mean	SD	Min	Max
Language Arts	95	42	118	2	1153
American Literature	151	31	29	1	215
Intro Extension English	131	42	58	1	392
Studies in British Literature	85	40	59	5	403
Studies in English	123	42	49	4	395
Studies in American Literature	25	50	74	4	270

Table C3. Total Number of Paper Tutoring Sessions by Math Course

Math Course	*N*	mean	SD	Min	Max
Math 1	278	2.6	3.6	1	31
Math 2	150	3.3	5.5	1	39
Math 3	83	2.6	4.5	1	38
Math 2A	67	2.3	2.7	1	16
Math 3A	51	2.6	3.5	1	22
Honors Math 1	71	3.2	5.8	1	41
Honors Math 2	62	4.0	5.1	1	29
Honors Math 3	80	4.1	5.1	1	34
Math 1 Prep	21	3.7	5.2	1	22

Table C4. Total Time in Minutes Spent on Paper Tutoring Sessions by Math Course

Math Course	*N*	mean	SD	Min	Max
Math 1	278	85	224	1	2,029
Math 2	150	105	219	2	1,960
Math 3	83	84	140	1	953
Math 2A	67	62	105	3	682
Math 3A	51	117	226	7	1,459
Honors Math 1	71	102	219	1	1,491

Math Course	*N*	mean	SD	Min	Max
Honors Math 2	62	184	306	6	1,496
Honors Math 3	80	113	176	3	1,169
Math 1 Prep	21	145	353	2	1,535

Table C5. Total Number of Essays Submitted for Review by ELA Course

ELA Course	*N*	mean	SD	Min	Max
Language Arts	640	0.4	0.9	0	6
American Literature	650	0.7	0.9	0	6
Intro Extension English	602	0.6	0.9	0	5
Studies in British Literature	257	1.2	2	0	12
Studies in English	243	1	1.6	0	14
Studies in American Literature	134	0.6	1.1	0	4

Appendix D. Additional information on Outcome Findings

Table D1. Unadjusted, raw course achievement scores for ELA courses across the school year

ELA Course	*N*	Fall 2021 Mean	Winter 2022 Mean	Spring 2022 Mean
Language Arts (Paper users)	95	35	54	53
Language Arts (Paper non-users)	513	34	54	53
American Literature (Paper users)	151	31	56	64
American Literature (Paper non-users)	410	35	54	62
Intro Extension English (Paper users)	131	43	69	61
Intro Extension English (Paper non-users)	415	42	66	58
Studies in British Literature (Paper users)	85	44	77	62
Studies in British Literature (Paper non-users)	167	47	70	64
Studies in English (Paper users)	123	57	83	71
Studies in English (Paper non-users)	113	55	81	70
Studies in American Literature (Paper users)	25	52	67	77
Studies in American Literature (Paper non-users)	73	47	62	82

Table D2. Comparison between matched samples of Paper users (tutoring sessions) and non-users on midyear ELA assessments (beginning-of-year performance was a covariate in all models)

ELA Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p>|t|*
Language Arts	513	95	608	-	-.004	.018	-0.21	0.836
American Literature	410	151	561	IEP, race	**.033**	.017	2.01	0.045*
Intro Extension English	415	131	546	IEP	.024	.017	1.40	0.162
Studies in British Literature	167	85	252	race	**.073**	.018	3.99	0.0***
Studies in English	113	123	236	-	.007	.012	0.60	0.551
Studies in American Literature	73	25	98	-	.036	.026	1.39	0.166

Table D3. Comparison between matched samples of Paper users (tutoring sessions) and non-users on end-of-year ELA assessments (beginning-of-year performance was a covariate in all models)

ELA Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p>|t|*
Language Arts	513	95	608	-	-.003	.018	-0.20	0.844
American Literature	410	151	561	IEP, race	.014	.015	0.97	0.330
Intro Extension English	415	131	546	IEP	.018	.016	1.21	0.229
Studies in British Literature	167	85	252	race	-.013	.015	-0.96	0.339
Studies in English	113	123	236	-	-.006	.015	-0.41	0.681
Studies in American Literature	73	25	98	-	-.059	.034	-1.75	0.083

Table D4. Comparison between matched samples of Paper users (tutoring sessions) and non-users on midyear math assessments (beginning-of-year performance was a covariate in all models)

Math Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p*>|*t*|
Math 1	553	278	831	No midyear assessment data				
Math 2	197	150	347	SES	.001	.006	0.21	0.836
Math 3	198	83	281	-	.000	.005	0.00	0.998
Math 2A	213	67	280	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				
Math 3A	133	51	184	-	-.008	.005	-1.53	0.128
Honors Math 1	128	71	199	No midyear assessment data				
Honors Math 2	125	62	187	SES	.002	.004	0.47	0.640
Honors Math 3	44	80	124	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				
Math 1 Prep*	51	21	72	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				

Table D5. Comparison between matched samples of Paper users (tutoring sessions) and non-users on end-of-year math assessments (beginning-of-year performance was a covariate in all models)

Math Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p>|t|*
Math 1	553	278	831	Race, SES	.020	.016	1.25	0.212
Math 2	197	150	347	SES	.005	013	0.41	0.680
Math 3	198	83	281	-	.014	.024	0.60	0.551
Math 2A	213	67	280	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				
Math 3A	133	51	184	-	.008	.027	0.31	0.755
Honors Math 1	128	71	199	-	-.022	.033	-0.66	0.512
Honors Math 2	125	62	187	SES	-.011	.024	-0.47	0.638
Honors Math 3	44	80	124	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				
Math 1 Prep*	51	21	72	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				

Table D6. Comparison between matched samples of Paper users (essay review feature) and non-users on midyear ELA assessments (beginning-of-year performance was a covariate in all models)

ELA Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p*>|*t*|
0122 Language Arts	478	161	639	race, IEP	-0.002	0.015	-0.15	0.884
0132 American Lit	345	305	650	-	-.030	0.014	-2.10	0.037
0113 Intro-Ext English	345	257	602	race	0.018	0.014	1.29	0.197
0130 Studies in Brit Lit	169	87	256	race	0.120	0.017	7.25	0.000
0114 Studies in English	139	104	243	race	0.022	.011	1.91	0.058
0129 Studies in American Lit	92	41	133	Could not run the comparison because, the differences in race between Paper users and non-users is statistically significant **and** the effect size is beyond allowable bounds of WWC standards				

Table D7. Comparison between matched samples of Paper users (essay review feature) and non-users on end-of-year ELA assessments (beginning-of-year performance was a covariate in all models)

ELA Course	Non-users	Users	Total *N*	Covariates	Coefficient	Std .Error	*t*-value	*p*>|*t*|
0122 Language Arts	478	161	639	race, IEP	0.026	0.014	1.91	0.057
0132 American Lit	345	305	650	-	-.011	0.120	-0.93	0.351
0113 Intro-Ext English	345	257	602	race	-0.008	0.012	-0.68	0.498
0130 Studies in Brit Lit	169	87	256	race	0.003	0.014	0.22	0.823
0114 Studies in English	139	104	243	race	-0.011	0.016	-0.68	0.498
0129 Studies in American Lit	92	41	133	-	Could not run the comparison because, the differences in race between Paper users and non-users is st. sig. and also the ES is beyond allowable bounds of WWC standards			