Impact Evaluation of Flashlight360 in Mountain View Public Schools: Year 2

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

June 2024

JOHNS HOPKINS UNIVERSITY

Impact Evaluation of Flashlight360 in Mountain View Public Schools: Year 2

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

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

Preferred citation: Cook, M.A. and Ross, S.M. (2024) *Impact Evaluation of Flashlight360 in Mountain View Public Schools: Year 2*. Center for Research and Reform in Education, Johns Hopkins University.

Contents

EXECUTIVE SUMMARY

This report is a follow-up to CRRE's initial impact evaluation of Flashlight360 in Mountain View Public Schools[1] in a large western state. In this mixed-methods study, we continued to examine Flashlight360 implementation and impacts as Flashlight360 usage expanded to a larger number of district schools. The primary focus on this report is on one-year and two-year program impacts, as well as associations between Flashlight360 digital usage variables and achievement gains.

- The present study expanded on the quasi-experimental design (QED) used last year to examine the efficacy of Flashlight360.
- As in the Year One report, the present Year Two report was situated in a large suburban district in a large western state. The target student population was Grades 1-12 ELL students, with more schools implementing Flashlight360 in Year Two than the small number of MVPS schools that implemented in Year One.
- Data sources included WIDA ACCESS composite scores and proficiency levels, as well as Speaking and Writing subscale scores. Digital usage variables included counts of Flashlight360 student submissions, as well as ratings of student submissions, by content domain.
- Impact analyses showed significant positive impacts of Flashlight360 on WIDA ACCESS composite score and proficiency levels, both for two-year and one-year student participants.
- Additional analyses showed significant positive impacts of Flashlight360 on WIDA ACCESS Speaking and Writing subscale scores for one-year students, as well as directionally positive impacts for two-year students.
- Flashlight360 digital usage variables were not significantly associated with WIDA ACCESS scores or proficiency levels.

[1] District pseudonym.

INTRODUCTION

Overview of Flashlight360

As described by the provider, Flashlight360 is an asset and evidence-based formative assessment and practice tool to improve and increase students' language competencies through the deliberate integration of speaking and writing. It was designed as a tool for teachers to support them in assessing, developing, and accelerating student talk and student writing during English language development (ELD) and across content areas. Designed as an evidence-based assessment and practice tool for measuring productive language, Flashlight360 allows teachers to gather immediate information about a student's oral and written language. It informs teachers and students through reflective feedback to use the results to set goals for advancing students' language development. Flashlight360 employs a formative process as a cycle for continuous improvement, while intending to be sensitive to students' cultural and linguistic assets and voice in the learning process.[2]

Overview of the Evaluation

In 2023, Flashlight Learning contracted with CRRE to conduct a retrospective mixed-methods, quasi-experimental design (QED) to study the impact of Flashlight360, a formative practice and assessment tool. The specific research interest is to determine the effects of Flashlight360 on Grades 1-12 ELL (English Language Learners) students' oral and written language growth in the Mountain View Public Schools (MVPS)[3] located in a large western state. As MVPS expanded the use of Flashlight360 into more district schools, the present report examines program impacts on first-year Flashlight360 students in 2023-24, as well as longitudinal impacts of Flashlight360 on students who completed their second year of using Flashlight360 programming.

The present study used a QED design to examine these research questions:

1. What are the effects of Flashlight360 on student growth on the WIDA Access assessment for Speaking and Writing?
   a) Did Flashlight360 have differential impacts for grade levels and student subgroups?
   b) Are Flashlight360 impacts different for one-year vs. two-year participants?
2. Did the amount of teacher feedback in Flashlight360 relate to student growth on the WIDA assessment?
3. Did the amount of performance tasks students submitted in Flashlight360 relate to student growth on the WIDA assessment?

[2] Flashlight360 - Progress Monitoring for Multilingual Students

[3] District pseudonym.

METHOD

Research Design

This study examined perceptions and effectiveness of Flashlight360 by continuing a retrospective, mixed-methods quasi-experimental design of ELLs in Grades 1-12 during the 2023-24 school year in Mountain View Public Schools. Outcome measures for this study included composite, speaking, and writing achievement gains on the WIDA ACCESS assessment administered to students in schools that implemented Flashlight360, relative to those of students in schools that did not implement the program. Analyses were conducted with new Flashlight360 students in 2023-24, as well as for returning Flashlight360 students who completed their second year in the program.

Participants

Details about study participants are presented below.

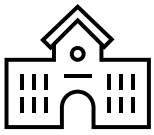
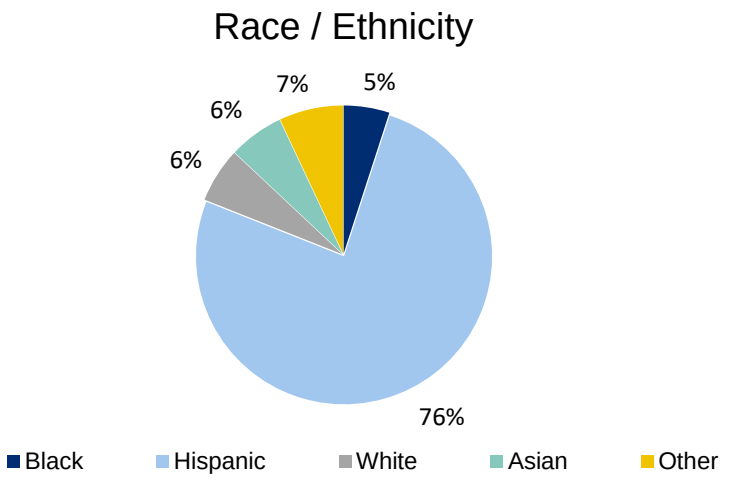
			
One-Year Sample	85 schools	4,141 Treatment students	4,552 Comparison students
Two-Year Sample	85 schools	306 Treatment Students	4,032 Comparison Students

Figure 1

Demographic snapshot of student participants



Mountain View Public Schools is a large western school district that operates 92 schools and serves approximately 64,000 students. MVPS expanded usage of Flashlight360 from limited implementation in 37 schools in 2022-23 to more expanded usage in nearly all of MVPS's 92 schools in the 2023-24 school year. Thus, two analytic samples are derived; one using two-year Flashlight360 students, in comparison to students that did not use Flashlight360 in either year, and another using new (one-year) 2023-24 Flashlight students only, in comparison to students that did not use Flashlight360 in either year.

Table 1
Demographics of Two-Year Analytic Sample

	Treatment %	Comparison (All) %	Matched Comparison %
Female	45.95	43.99	39.07
Black	6.76	5.61	6.29
White	7.09	5.06	5.96
Asian	8.45	5.88	9.93
Hispanic	78.35	77.50	77.81
Pacific Islander	3.38	2.88	2.32
American Indian	0.00	0.55	2.32
n	302	4,032	302

Race/Ethnicity demographics may not add up to 100%.

Table 2
Demographics of One-Year (2023-24 Only) Analytic Sample

	Treatment %	Comparison (All) %	Matched Comparison %
Female	45.81	44.43	44.82
Black	5.10	6.20	4.39
White	7.32	5.47	7.61
Asian	5.77	5.93	6.99
Hispanic	78.75	79.09	79.76
Pacific Islander	3.01	2.97	2.63
American Indian	0.49	0.53	0.47
n	4,141	4,552	1,115*

Race/Ethnicity demographics may not add up to 100%; *some comparison students were matched with more than one treatment student.

Across both the two-year and one-year samples, sample composition was similar across conditions. A majority of students (> 75%) were Hispanic, followed by small percentages of White, Black, and Asian students. No significant differences were observed across conditions on any of the provided demographic variables. It is important to note that MVPS did not provide data related to special education services or economic disadvantage; thus, we cannot report on differences on those variables between conditions or across samples.

Measures

In order to address the research questions, the study team analyzed WIDA ACCESS composite and subscale scores and data related to Flashlight360 student submissions and teacher feedback (see Table 3).

Table 3
Research questions with data sources and measures

Research questions	WIDA Access Scores	Flashlight360 submissions	Flashlight360 feedback
1. What are the effects of Flashlight360 on student growth on the WIDA Access assessment for Speaking and Writing?	✓		
2. Did the amount of teacher feedback in Flashlight360 relate to student growth on	✓	✓	✓

the WIDA assessment?			
3. Did the amount of performance tasks students submitted in Flashlight360 relate to student growth on the WIDA assessment?	✓	✓	

Data sources and measures for the current study included student achievement and demographics, as well as Flashlight360 digital usage variables as described below.

WIDA ACCESS for ELLs. The WIDA ACCESS for ELLs is a suite of summative English proficiency assessments that are administered annually in the spring of each school year to ELL students in Grades K-12 in WIDA Consortium states. In MVPS' home state, the spring 2023 testing window was from early January 2023 to early March 2023. The ACCESS assessment consists of the following domains: Reading, Writing, Listening, Speaking, Oral Language, Literacy, and Comprehension. Scale scores range from 100-600, while Proficiency Levels range from 1 to 6 (Entering, Emerging, Developing, Expanding, Bridging, and Reaching) for each domain. Overall composite scores are calculated as follows: 35% Reading + 35% Writing + 15% Listening + 15% Speaking. An overall score/proficiency level is only calculated when all four of these domains have been assessed. In the current evaluation, spring 2023 Overall scale scores and proficiency levels were used as the main outcome variables, while spring 2022 Overall scale scores and proficiency levels were used as prior achievement variables. Domain-level scale scores and proficiency levels were used in supplementary analyses.

In addition to WIDA ACCESS data, Flashlight360 also provided student-level demographic data for all MVPS students. These data files contained indicator variables for an array of demographic items including race/ethnicity, and various IEP variables. However, data were very sparse on all demographics except for the Hispanic indicator variable. Thus, we only included race/ethnicity variables in our models, as data on other demographic items were nearly completely missing across all students.

Digital Program Usage. As in the Year 1 study, Flashlight360 provided CRRE with digital usage data for MVPS treatment students. Based on Flashlight360 guidance, only submissions marked as "Submitted" or "Reviewed" were classified as completed digital program submissions and were included in subsequent analyses. Additionally, teachers provided scores and feedback to student submissions classified as "Reviewed;" scores were given on submissions across the following domains: Fluency, Grammar, Interpretations, Pronunciation, Vocabulary, Description, Points and Reason, and Vocabulary.

Analytical Approach

Data for students in Grades 1-12 were analyzed descriptively by examining patterns in WIDA ACCESS scores and usage. Hierarchical Linear Modeling (HLM) with students nested within schools was used to determine impacts of Flashlight360 on WIDA ACCESS score gains, as well as to determine relationships between Flashlight360 digital usage and WIDA ACCESS score gains. Demographic variables such as gender and race/ethnicity were included in all models, as well as dummy variables for student grade levels.

Initially, the two-year treatment sample scored over one standard deviation lower than the comparison pool of students, and one-year treatment students scored more than .65 SDs lower than comparison students. To adjust for these prior achievement and demographic differences, propensity score matching (PSM) was used to create comparison groups of students that were as similar as possible to treatment students (see Tables 1 and 2). Propensity scores were computed using the psmatch2 command in Stata (v 18.0); one-to-one matching without replacement was used for the longitudinal sample, while one-to-one matching with replacement was used for the one-year sample, as the expanded use of Flashlight360 in Year 2 restricted the size of the comparison pool of students in some grades. In other words, each treatment student was matched with one comparison student, but the same comparison student may have matched with multiple treatment students, due to small comparison group sample sizes in some grades. The result of these PSM procedures was that treatment students were matched with comparison students who were as similar as possible in terms of prior WIDA ACCESS achievement and demographic variables, allowing for a stronger contrast of treatment and comparison students. The matched samples demonstrated baseline equivalence overall and across all grade levels. Overall baseline equivalence for both samples can be found in Table 4. After PSM was applied to both samples, standardized mean differences for both samples were below 0.10 SDs, well below the WWC's 0.25 SD cutoff.

Table 4
Baseline Equivalence, WIDA ACCESS Composite Scores

Analytic Sample	All students	Treatment			Comparison			Standardized Mean Difference
	n	*n*	*M*	*SD*	*n*	*M*	*SD*	*M*
Two Year	606	303	281.19	48.47	303	281.33	48.31	-0.02
One Year	8,282	4,141	287.60	54.57	4,141	287.46	54.56	0.09

RESULTS

This section of the report begins with findings related to Flaslight360 impacts on WIDA composite and subscale scores. Then, we present results pertaining to associations between teacher feedback on submissions and WIDA score gains. We conclude with a presentation of associations between counts of student submissions and WIDA score gains. Note that unadjusted descriptive analyses of WIDA ACCESS gains and Flashlight360 digital usage metrics can be found in Appendix A.

Flashlight360 WIDA Impacts

What are the effects of Flashlight360 on student growth on the WIDA Access assessment for Speaking and Writing?

Key Findings

- Two-year Flashlight360 students significantly outgained matched comparison students by more than 6 points on WIDA composite scores.
- Year two-only Flashlight360 students significantly outgained matched comparison students by more than 4 points on WIDA composite scores.
- Year two-only Flashlight360 students also significantly outgained matched comparison students by 6 points on the WIDA Writing assessment and by more than 8 points on the WIDA Speaking assessment.

Two-Year (Longitudinal) Analyses

This set of analyses focuses on the impact of Flashlight360 on spring 2024 WIDA ACCESS scale scores and proficiency levels for two-year Flashlight360 students. Table 5 shows the results for spring 2024 WIDA ACCESS scale scores as the outcome variable. Both main impact analyses were grand-mean centered to enable interpretation of the intercept. Additional impact analyses on WIDA subscale scores (Writing and Speaking) can be found in Appendix B.

Table 5
Impact Analysis of Flashlight360 on Spring 2024 Overall WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	6.188*	2.416	.010	0.13
Constant	323.553***	1.671	<.001	
Variance of Constant	20.504			
Residual	655.967			
Student *N*	598			
School *N*	68			

Note. * $p < .05$; *** $p < .001$.

Flashlight360 was found to have a significant positive impact on WIDA ACCESS scale score gains from spring 2022 to spring 2024. The regression estimate ("Flashlight360") can be interpreted as the difference in WIDA ACCESS scale score gains for treatment students in relation to matched comparison students. Thus, the results show that treatment students outgained matched comparison students by more than 6 points. The effect size was 0.13 SDs, indicating a medium-sized impact of Flashlight360.

We also conducted a similar analysis that examined WIDA ACCESS overall proficiency levels. Proficiency levels range from 1 as the lowest level to 6 as the higher level. Spring 2024 overall proficiency level was the outcome variable, while spring 2022 overall proficiency level was the prior achievement control variable. This model is otherwise identical to that used in the first impact analyses. Results are shown in Table 6.

Table 6
Impact Analysis of Flashlight360 on Spring 2024 Overall WIDA ACCESS Proficiency Levels

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	0.121*	0.056	.031	0.13
Constant	2.996***	0.039	<.001	
Variance of Constant	0.007			
Residual	0.387			
Student *N*	598			
School *N*	68			

Note. * $p < .05$; *** $p < .001$.

The proficiency-level results are very similar to those of scale scores. The regression estimate can be interpreted as the difference in proficiency level gain for treatment students, in relation to matched comparison students. Treatment students averaged nearly 0.12-point gains in WIDA ACCESS Overall proficiency level from spring 2022 to spring 2024, and this difference was statistically significant ($p < .05$). The practical significance of Flashlight360 in this analysis was also similar to that observed in the prior analysis, with an observed effect size of 0.13 SDs. Taken together, these results

provide evidence supporting the efficacy of two years of Flashlight360 participation in relation to the achievement of ELLs.

One-Year Analyses

Table 7 shows the results of impact analyses examining Flashlight360 impacts on WIDA ACCESS composite score gains from spring 2023 to spring 2024 for students who entered Flashlight360 programming in 2023-24. As in the previous section, all impact analyses were grand-mean centered to enable interpretation of the intercept.

Table 7
Impact Analysis of Flashlight360 on Spring 2024 Overall WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	4.283**	1.370	.002	0.10
Constant	313.048***	1.376	<.001	
Variance of Constant	39.091			
Residual	550.782			
Student *N*	5,144			
School *N*	81			

Note. ** $p < .01$; *** $p < .001$.

Flashlight360 was found to have a significant positive impact on WIDA ACCESS scale score gains from spring 2023 to spring 2024. As in the previous analyses, the regression estimate can be interpreted as the difference in WIDA ACCESS scale score gains for treatment students in relation to matched comparison students. Thus, the results show that treatment students outgained matched comparison students by more than 4 points. The effect size was 0.10 SDs, indicating a medium-sized practical effect of Flashlight360 on WIDA ACCESS composite scores.

We also conducted a similar analysis that examined WIDA ACCESS overall proficiency levels. Spring 2024 overall proficiency level was the outcome variable, while spring 2023 overall proficiency level was the prior achievement control variable. This model is otherwise identical to that used in the prior impact analyses. Results are shown in Table 8.

Table 8
Impact Analysis of Flashlight360 on Spring 2024 Overall WIDA ACCESS Proficiency Levels

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	0.103**	0.032	.001	0.11
Constant	3.087***	0.033	<.001	
Variance of Constant	0.245			
Residual	0.316			
Student *N*	5,144			

School *N*	81

Note. ** $p < .01$; *** $p < .001$.

The proficiency-level results are very similar to those of scale scores. Treatment students averaged approximately 0.10-point gains in WIDA ACCESS Overall proficiency level from spring 2023 to spring 2024, and this difference was statistically significant ($p < .01$). The practical significance of Flashlight360 in this analysis was also similar to that observed in the prior analysis, with an observed effect size of 0.11 SDs.

Below we report the results of impact analyses examining the Flashlight360 impacts on WIDA Writing and Speaking subscale scores for Year Two-only students. These analyses control for spring 2023 Writing and Speaking scores, respectively, but are otherwise similar to the main one-year analyses. The results of these analyses are shown in Tables 9 and 10.

Table 9

Impact Analysis of Flashlight360 on Spring 2024 Writing WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	6.022**	2.078	.004	0.11
Constant	294.274***	1.946	<.001	
Variance of Constant	62.828			
Residual	1166.346			
Student *N*	5,144			
School *N*	81			

Note. ** $p < .01$; *** $p < .001$.

Table 10

Impact Analysis of Flashlight360 on Spring 2024 Speaking WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	8.936**	3.437	.009	0.13
Constant	277.640***	2.907	<.001	
Variance of Constant	136.510			
Residual	2975.926			
Student *N*	5,144			
School *N*	81			

Note. ** $p < .01$; *** $p < .001$.

Significant positive impacts of Flashlight360 were observed on both Writing and Speaking subscale scores, with treatment students averaging 6-point larger gains on Writing and nearly 9-point larger gains on the Speaking subscale. Effect sizes ranged between 0.11-0.13 SDs, indicating medium-sized practical effects on Writing and Speaking scale score gains.

Flashlight360 Feedback and WIDA Score Associations

Did the amount of teacher feedback in Flashlight360 relate to student growth on the WIDA assessment?

Key Findings

- Pearson correlations between Flashlight360 submission scores and spring 2024 WIDA Composite scores were generally weak in magnitude and inconsistent in direction.
- Similarly, regression analyses generally found non-significant associations between Flashlight360 submission scores and spring 2024 WIDA Composite scores.

Associations Between Flashlight360 Submissions and WIDA Scores

In this section, we discuss the results of correlational analyses examining the relationships between Flashlight360 submissions and WIDA overall scale scores. Descriptive statistics overviewing average submission scores by domain and for each sample (two-year and one-year) can be found in Appendix C.

Table 11 shows the unadjusted Pearson correlations between Flashlight360 submission scores, by domain, and spring 2024 WIDA Composite scale scores. Results were calculated separately for two-year and one-year students. It is important to note that not all student submissions were scored; thus, sample sizes were considerably smaller in these analyses than in prior impact analyses.

Table 11
Correlations Between Flashlight360 Scores and Spring 2024 WIDA ACCESS Scores

Flashlight360 Score Domain	Two-Year Students (*n* = 61)	One-Year Students (*n* = 772)
Fluency	+.03	-.02
Grammar	+.03	+.01
Interpretation	-.11	-.01
Pronunciation	-.01	-.04
Vocabulary	+.01	.00
Description	-.09	+.05
Points and Reasons	-.07	+.08
Vocabulary	-.21	+.02

No significant correlations between Flashlight360 domain scores and WIDA ACCESS composite scores were observed in either treatment sample. Correlations were generally very small in magnitude and inconsistent in direction.

Regression analyses similar to the main impact analyses were also conducted to further examine associations between Flashlight360 submission scores and WIDA ACCESS composite scores, controlling for prior WIDA scores and demographic variables. Domain scores that were found to be significantly associated with WIDA ACCESS composite scores in these analyses are shown in Table 12; full regression tables containing results from all analyses can be found in Appendix C.

Table 12
Selected Associations Between Flashlight360 Scores and Spring 2024 WIDA ACCESS Scores

Flashlight360 Score	Estimate	Standard Error	*p*-value
Two Year (*n* = 61)			
Grammar	11.065*	4.878	.023
Vocabulary	10.017*	4.563	.028
One Year (*n* = 772)			
Points and Reasons	2.265*	1.077	.035

Note. * $p < .05$.

In the two-year sample, significant associations were found between Flashlight360 Grammar and Vocabulary domain submission scores and WIDA score gains, with each 1-point increase in a student's submission score in Grammar or Vocabulary associated with a more than 10-point increase in WIDA ACCESS composite score. Similarly, Points and Reasons domain submission scores were significantly associated with WIDA score gains, with each 1-point increase in Points and Reasons domain submission score associated with more than a 2-point gain in WIDA ACCESS composite score. Across all eight domains and both samples, results of these analyses were generally inconsistent in direction and small in magnitude, similar to the results of the Pearson correlation analyses. Thus, there is only limited evidence of Flashlight360 submission scores relating to WIDA ACCESS composite scores in the 2023-24 school year.

Associations Between Student Submissions and WIDA Scores

Did the amount of performance tasks students submitted in Flashlight360 relate to student growth on the WIDA assessment?

Key Findings

- Two-year students completed more Flashlight360 submissions, on average, than did one-year students.
- No statistically significant associations between counts of Flashlight360 submissions and WIDA ACCESS composite scores were observed.

Student Flashlight360 Submissions and WIDA ACCESS Scores

We begin by descriptively examining student Flashlight360 submissions by sample. Table 13 shows descriptive statistics from Flashlight360 submissions for students in both the two-year and one-year samples. Note that these counts include all student submissions that were classified as "Submitted" or "Reviewed," while analyses in the previous section only included submissions classified as "Reviewed."

Table 13
Flashlight360 Submission Averages, by Sample

Analytic Sample	Mean	SD	*N*
Two-Year	3.48	3.24	303
One-year	2.87	3.23	4,141

Two-year students averaged nearly 3.5 Flashlight360 submissions, while one-year students averaged less than three submissions per student. This suggests that two-year students may have been participating more in Flashlight360 than the one-year students who just started in the program last year.

We also conducted regression analyses similar to the main impact analyses to examine the associations between counts of Flashlight360 submissions and WIDA ACCESS score gains. The results of these analyses, which were conducted by sample, are shown in Table 14.

Table 14
Associations Between Flashlight360 Submissions and Spring 2024 WIDA ACCESS Scores

Flashlight360 Score	Estimate	Standard Error	*p*-value
Two-Year (*n* = 296)	0.203	0.513	.692
One-Year (*n* = 4,056)	0.008	0.126	.950

The results of these analyses did not show any significant associations between counts of Flashlight360 submissions and WIDA score gains. Regression estimates in both analyses were positive, but very small in magnitude.

DISCUSSION

The purpose of the present study was to examine Year 2 impacts of Flashlight360 programming on ELL students across Mountain View Public Schools. This report focused on quantitative research questions only; the prior qualitative report focused on teacher survey responses in Year 2 of program implementation. To address the primary research questions, we extended the QED design used in the Year 1 evaluation to examine both two-year and one-year impacts of Flashlight360 on WIDA ACCESS score gains, as well as to again examine associations between digital program usage variables and WIDA ACCESS score gains. Two separate analytic samples (two-year usage and one-year usage) were identified, with propensity-score matching used to identify students similar to treatment students in each analytic sample. Flashlight360 impacts were generally positive in Year 2, with the main impact analyses finding significant positive two-year and one-year impacts. Correlational analyses examining relationships between digital program usage variables and WIDA score gains generally did not find significant associations, with most associations being very small in magnitude. Collectively, the results of this evaluation continue to support the efficacy of Flashlight360 on language achievement for ELL students.

Achievement Impacts

Results of the main impact analyses showed positive and statistically significant impacts of Flashlight360 on ELL learners' WIDA ACCESS scores, for both two-year and one-year users. Two-year treatment students outgained matched comparison students by more than 6 points, while one-year treatment students outgained matched comparison students by more than 4 points. Findings were similar when examining proficiency level gains, with treatment students in both samples outgaining matched comparison students by more than .10 proficiency levels. When examining WIDA ACCESS Writing and Speaking subscale scores, one-year treatment students significantly outgained matched comparison students by more than 10-points on both subscales, while two-year treatment students directionally outgained comparison students. As Flashlight360 is specifically designed to address speaking and writing skills, it was expected that program impacts may be more pronounced on these two WIDA subscales.

Digital Usage Findings

Usage analyses showed that two-year students averaged nearly 3.5 Flashlight360 submissions across the 2023-24 school year, while one-year students averaged slightly less than three submissions that year. Counts of Flashlight360 submissions were not significantly associated with WIDA ACCESS scores or proficiency level gains in either sample. Similarly, Flashlight360 submission scores were not significantly related to WIDA ACCESS scale scores and proficiency levels, with correlations generally small in magnitude and inconsistent in direction. Regression analyses similar to the main impact analyses showed that these associations remained generally nonsignificant, even after controlling for prior WIDA achievement and demographic variables. This is a different finding from Year 1 analyses, which found significant positive associations between Flashlight360 submission scores and WIDA ACCESS score gains.

Conclusion

In all, the results of this evaluation build on the findings of the initial Year One evaluation that found significant positive impacts of Flashlight360 on WIDA ACCESS scores by significantly supporting the efficacy of Flashlight360 for students who have used the program for one year or two years. Program usage was more frequent, in terms of counts of submissions, for two-year students compared to one-year students. One limitation of this study was that it included only one school district and therefore may not generalize to other contexts or student populations. Additional study is thus encouraged in additional school districts in other educational settings to continue to examine Flashlight360's impact on ELL language and speaking achievement.

APPENDIX A: Descriptive Analyses

Table A1

WIDA ACCESS Pretest and Spring 2024 Scores, Across All Grade Levels

Group	Pretest*	Posttest	Change
Two-Year			
Treatment (*n* = 296)	281.50	329.49	47.99
Matched Comparison (*n* = 302)	281.25	323.75	42.50
One-Year			
Treatment (*n* = 4,141)	287.60	316.57	28.97
Matched Comparison (*n* = 1,115)	287.46	313.29	25.83

Note. Pretest is Spring 2022 WIDA ACCESS for Two-Year students and Spring 2023 WIDA ACCESS for One-Year students.

Table A2

Average Flashlight360 Submission Scores, by Domain, Two-Year Students

Domain	Mean	SD	N
Fluency	2.15	0.82	64
Grammar	1.91	0.72	64
Interpretation	2.02	0.83	64
Pronunciation	2.34	0.90	64
Vocabulary	2.03	0.76	64
Description	1.92	0.73	64
Points and Reasons	1.77	0.65	64
Vocabulary	2.10	0.67	64
Completed Submissions	3.48	3.24	303

Table A3

Average Flashlight360 Submission Scores, by Domain, One-Year Students

Domain	Mean	SD	N
Fluency	2.25	0.91	782
Grammar	1.89	0.76	781
Interpretation	2.13	0.85	781
Pronunciation	2.46	0.94	781
Vocabulary	2.12	0.82	780
Description	1.88	0.78	781
Points and Reasons	1.78	0.78	781
Vocabulary	2.04	0.74	780
Completed Submissions	2.87	3.23	4,141

APPENDIX B: Two-Year WIDA Subscale Score Analyses

Table B1

Impact Analysis of Flashlight360 on Spring 2024 Writing WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	2.612	3.264	.424	0.05
Constant	310.310***	2.261	<.001	
Variance of Constant	26.295			
Residual	1285.370			
Student *N*	598			
School *N*	68			

Note. *** $p < .001$.

Table B2

Impact Analysis of Flashlight360 on Spring 2024 Speaking WIDA ACCESS Scale Scores

Variable	Estimate	Standard error	*p*-value	Effect Size
Flashlight360	11.155	5.730	.052	0.16
Constant	282.825***	3.996	<.001	
Variance of Constant	180.358			
Residual	3263.901			
Student *N*	598			
School *N*	68			

Note. *** $p < .001$.

APPENDIX C: Full Digital Usage Regression Analyses

Table C1

Adjusted Associations Between Flashlight360 Scores and Spring 2024 WIDA ACCESS Scores, Two-Year Students (n = 61)

Flashlight360 Score Domain	Estimate	Standard Error	*p*-value
Fluency	6.485	4.331	.134
Grammar	11.065*	4.878	.023
Interpretation	-0.613	4.055	.880
Pronunciation	6.747	4.239	.111
Vocabulary	10.017*	4.563	.028
Description	9.397	5.196	.071
Points and Reasons	-1.038	5.634	.854
Vocabulary and Grammar	3.447	5.930	.561

Note. * $p < .05$.

Table C2

Adjusted Associations Between Flashlight360 Scores and Spring 2024 WIDA ACCESS Scores, One-Year Students (n = 772)

Flashlight360 Score Domain	Estimate	Standard Error	*p*-value
Fluency	-0.492	0.928	.596
Grammar	0.578	1.120	.606
Interpretation	0.272	0.994	.784
Pronunciation	-0.582	0.907	.521
Vocabulary	0.160	1.024	.876
Description	1.636	1.077	.129
Points and Reasons	2.265*	1.077	.035
Vocabulary and Grammar	1.532	1.141	.179

Note. * $p < .05$.