CommonLit

Digital CommonLit Lesson Usage and MCAP ELA Achievement

Results from a Maryland District,
School Year 2025-2026



November 2026

Table of Contents

Executive Summary 3

Introduction 4

- Research Questions 5

Method 5

- Sample 5
- Measures 6
  - CommonLit Usage 6
  - Maryland Comprehensive Assessment Program (MCAP) 7
  - CommonLit Assessment Series 7
  - iReady Assessment 8

Results 8

- MCAP Performance in School Years 2024-2025 and 2025-2026 8
- CommonLit Usage in School Year 2025-2026 9
- CommonLit Usage and MCAP Performance 10

Discussion 15

- Limitations 15
  - Observational Study Design 15
  - Usage and Instructional Context 16
  - Outcome Scaling and Sample 16
- Future Research 17

References 18

Appendices 20

- Appendix A: Descriptive Statistics 20
- Appendix B: Robustness Checks 21
- Appendix C: CommonLit Assessment Series Positively Correlated with the MCAP 25

Executive Summary

This technical report describes a study that examined students' performance on the Maryland Comprehensive Assessment Program (MCAP) in the 2025-2026 school year. The study sample included 9,153 students in Grades 4-8 taught by 240 teachers across 43 schools in a district in Maryland. Results revealed that students who completed more CommonLit lessons than their classmates had significantly higher MCAP ELA scaled scores, even after controlling for prior achievement on the MCAP, grade level, and demographics. Each additional lesson was associated with a small gain, about 1 point per 6 additional lessons. Additionally, students who completed more CommonLit lessons than their classmates were more likely to score at the proficient levels on the MCAP (Levels 3 and 4). This translated into a 1 point difference in likelihood of scoring proficiently, per 2 additional lessons. In sum, this study provided support for the efficacy of CommonLit for students' ELA achievement.

This study was rated as meeting the Every Student Succeeds Act (ESSA) Level 3 standards ("Promising Evidence") by Instructure's Research team, a third-party edtech research group. Specifically, this study was a treatment-only study that documented sample characteristics and average CommonLit usage. Moreover, this study used appropriate analyses for the design, including correlative analyses with statistical controls for selection bias. Importantly, there were statistically significant, positive effects of the CommonLit program on students' reading comprehension growth. Ongoing CommonLit studies aim to meet ESSA Level 2 standards ("Moderate Evidence") by examining the impact of CommonLit by comparing students at schools that use CommonLit to a comparison group of students at schools that do not use CommonLit.



Introduction

In 2024, the National Assessment of Educational Progress (NAEP) reading assessment showed significant declines since 2019 in average reading scores and in the percent of students who scored at or above the NAEP Proficient Level (National Center for Education Statistics, 2025). Only 30% of eighth-grade students scored at or above the NAEP Proficient level, with the remaining students scoring at (37%) or below (33%) the NAEP Basic level. This was significantly lower than pre-pandemic results in 2019, where 34% of eighth-grade students scored at or above NAEP Proficient, with the remaining students scoring at (39%) or below (27%) NAEP Basic. Declines in reading proficiency have also been documented across several other reading assessments (Kuhfeld et al., 2023; Kuhfeld & Lewis, 2025), with widening achievement gaps for economically disadvantaged students and traditionally underserved students (Fahle et al., 2024). These national trends make it especially important for districts to understand whether the literacy tools they have adopted are associated with better student reading outcomes.

To help address these long standing literacy challenges, CommonLit provides instructional materials for teachers and students. CommonLit is a nonprofit education technology organization dedicated to ensuring that all students—especially those in economically disadvantaged schools—graduate with the reading, writing, communication, collaboration, and problem-solving skills they need to be successful. CommonLit's theory of change is that when schools—including those in underserved communities—have access to a coherent curriculum, aligned professional learning, meaningful assessments, and implementation support, all students will be able to read on grade level and succeed in school and beyond.

In 2013, CommonLit began as a research project that found increased student engagement in middle school classrooms in Boston that used CommonLit – an online library of 30 lessons at the time (Brown, 2014). CommonLit now offers a Text Library with thousands of supplemental texts for Grades 3-12 as well as a year-long English Language Arts (ELA) core curriculum for Grades 6-12 called CommonLit 360 that was rated as all-green by EdReports for middle school. CommonLit's evidence base has also grown over time. Several studies have shown positive outcomes for students learning with CommonLit, such as greater reading growth on the CommonLit Assessment Series (AdHoc

Analytics, 2018; CommonLit, 2023a; CommonLit, 2024a; CommonLit 2026a; CommonLit, 2026b; CommonLit & Mathematica, 2023), higher writing growth (CommonLit, 2024d), and higher performance on state ELA tests (CommonLit 2024b; CommonLit, 2025; Lee & Shah, 2023; Lee & Shah, 2024). Research has also shown that students' scores on the CommonLit Assessment Series were positively correlated with their scores on state ELA tests in New York, New Jersey, and Florida (CommonLit, 2023b, CommonLit, 2024b).

Studies involving students' performance on state ELA tests have only been possible through research partnerships between CommonLit and school districts. For the 2025-2026 school year, a district in Maryland entered into a research partnership and shared data from the Maryland Comprehensive Assessment Program (MCAP). This technical report details results from this research partnership. Beyond the research partnership, the district has been in a paid partnership with CommonLit since the 2021-2022 school year. In addition to the free resources from CommonLit's Text Library and CommonLit 360, the district has received access to paid resources such as a CommonLit Account Manager, the CommonLit Assessment Series, and virtual and on-demand professional development to support ELA instruction.

Research Questions

The study addressed three research questions:

- How did students perform on the MCAP during the 2024-2025 and 2025-2026 school years?
- How did students and teachers use CommonLit during the 2025-2026 school year?
- In the 2025-2026 school year, was there a significant relationship between students' CommonLit usage and their performance on the MCAP ELA assessment?

Method

Sample

The study sample included students who took the MCAP in the 2024-2025 and 2025-2026 school years and submitted at least one digital CommonLit lesson in the 2025-2026 school year. Overall, there were 9,153 students in Grades 4-8 taught by 240 teachers across 43 schools. The distribution of students across grades is represented in Table 1.

Table 1. Number of students in each grade level in the sample

Grade	Students	%
4	2,390	26%
5	2,464	27%
6	1,664	18%
7	936	10%
8	1,699	19%
Total	9,153	100%

All schools had access to the resources available in CommonLit's School Essentials PRO package. Based on National Center for Education Statistics (NCES) data, 13 schools (30%) were eligible for Title I funding. Student-level demographics were provided by the district. To protect student privacy, analyses only included groups with 20 or more students. This ensured that no individual student could be identified based on unique characteristics. The analysis sample included 51% male students and 49% female students. In terms of race and ethnicity, the sample included 63% White students, 16% Black students, 10% multiracial students, 7% Hispanic/Latino students, and 4% Asian students. Three percent of the sample qualified for English Language support. Fourteen percent of the sample received special education services. Thirteen percent of the sample were identified as gifted and talented.

Measures

CommonLit Usage

In this study, CommonLit usage was operationalized as the number of digital CommonLit lessons that were submitted by a student. CommonLit usage was used as a primary predictor variable in analyses that examined the relationship between CommonLit usage and student ELA achievement. To distinguish individual usage from classroom-level usage, each student's lesson count was separated into two components: (1) Relative usage was the student's usage relative to their classroom average, or how many more or fewer lessons they submitted than their classmates, and (2) Classroom average usage was the average usage of the student's classroom.

Classrooms were defined by each student's main teacher. The main teacher was the teacher through whom the student submitted the most digital lessons during the 2025-2026 school year, combined across CommonLit 360, Library, and Target lessons. Classroom average usage was calculated across students in the analytic sample who shared a main teacher, and relative usage was each student's lesson count minus that average.

Maryland Comprehensive Assessment Program (MCAP)

The district shared several measures from the MCAP with CommonLit: MCAP ELA scaled scores, MCAP reading-only scaled scores, and MCAP ELA performance levels (see Appendix A for descriptive statistics by grade level). MCAP ELA performance levels were as follows for each grade level: 1 = Beginning Learner, 2 = Developing Learner, 3 = Proficient Learner, and 4 = Distinguished Learner. For analyses that examined the relationship between CommonLit usage and student ELA achievement, MCAP ELA performance levels and MCAP ELA scaled scores from the 2025-2026 school year were used as the primary outcome variables. MCAP ELA scaled scores from the 2024-2025 school year were used as the baseline variable. MCAP reading-only scaled scores were also examined as an additional measure of students' end-of-year ELA achievement (see Appendix B for additional analyses). MCAP scaled scores (ELA and reading-only) were standardized within grade level prior to analysis, given that the MCAP is not vertically scaled.

CommonLit Assessment Series

In the 2025-2026 school year, each assessment in the CommonLit Assessment Series included three reading passages (two informational passages and one literary passage) and between 32-38 questions total, depending on the passage and the grade level. Scaled scores on the assessments ranged from 150 to 250. These assessments were field-tested with over a million students and items were calibrated for difficulty using item response theory (IRT). For more information about the Assessment Series, see the research report (CommonLit, 2024c).

Out of the 9,153 students in the sample who took the MCAP, 9% (786 students) also took the CommonLit Pre-Assessment at the beginning of the school year and the CommonLit Post-Assessment at the end of the school year. Previous research in Maryland demonstrated a positive correlation between the CommonLit Assessment Series and the MCAP with a larger sample of nearly 5,000 students (CommonLit, 2025). Given that the

number of students who took the CommonLit Assessment Series was much lower in this sample, we conducted exploratory correlational analyses and reported the results in Appendix C. Results replicated previous research and showed a positive correlation between the CommonLit Assessment Series and the MCAP.

iReady Assessment

The district also shared measures from the iReady Assessment: Lexile Levels for Fall, Winter, and Spring from the 2025-2026 school year (see Appendix A for descriptive statistics by grade level). Out of the 9,153 students in the sample who took the MCAP, 95% (8,718 students) also took the iReady Assessment in the Fall and the Spring. Students' Lexile levels from the Spring iReady Assessment were examined as an additional measure of students' end-of-year ELA achievement, controlling for baseline Lexile levels from the Fall iReady Assessment.

Results

MCAP Performance in School Years 2024-2025 and 2025-2026

First, we examined how students in this sample scored on the MCAP in school years 2024-2025 and 2025-2026. Across both school years, 66% of students in the sample scored at the proficient levels (Levels 3 and 4) on the MCAP. Table 2 provides a breakdown of MCAP performance by grade level across both school years. Figure 1 shows MCAP Performance Levels for the 2025-2026 school year.

Table 2. MCAP Performance by Grade Level in School Years 2024-2025 and 2025-2026

		SY 2024-2025		SY 2025-2026	
Grade (SY 2025-2026)	**Students**	**Average MCAP score**	**% who scored at Level 3 or 4**	**Average MCAP score**	**% who scored at Level 3 or 4**
All Students	9,153	–	66%	–	66%
Grade 4	2,390	759	64%	758	64%
Grade 5	2,464	756	61%	754	59%
Grade 6	1,664	755	65%	759	72%
Grade 7	936	761	74%	761	75%
Grade 8	1,699	760	74%	758	66%

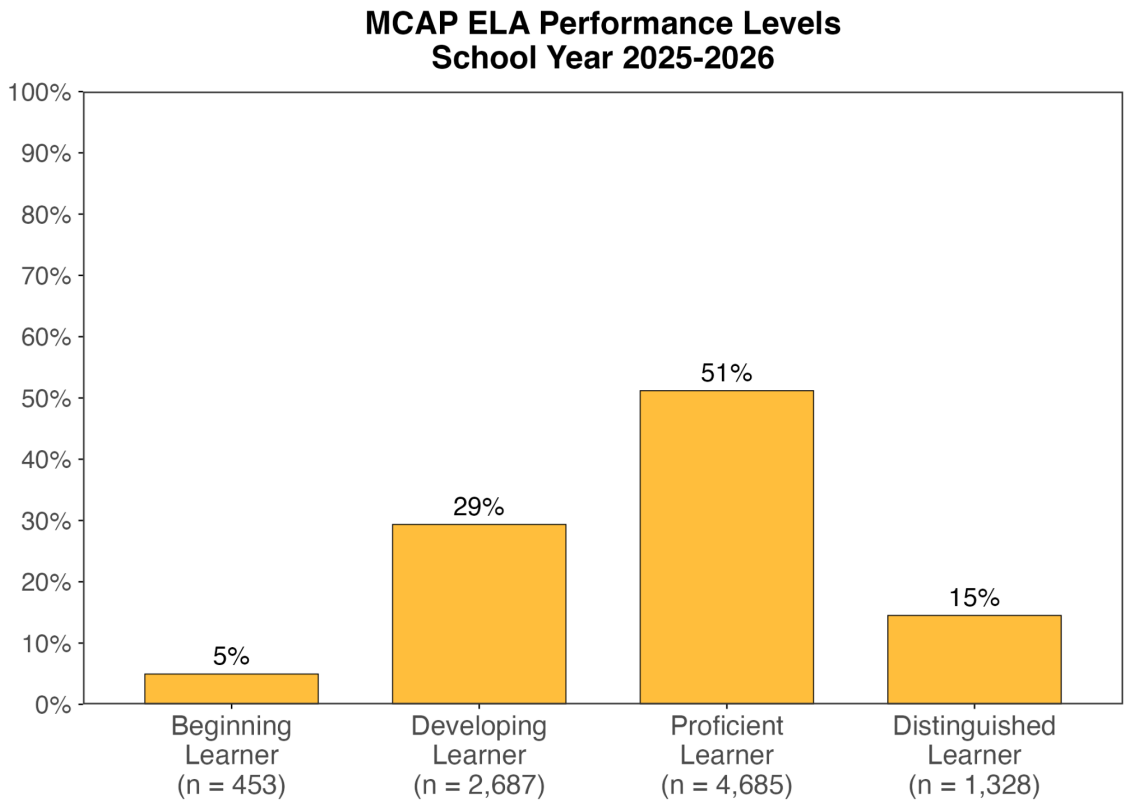


Figure 1. MCAP performance levels among students in the sample in school year 2025-2026.

CommonLit Usage in School Year 2025-2026

Next, we examined the average number of digital CommonLit lessons submitted by students. On average, students submitted 9 digital CommonLit lessons in the 2025-2026 school year. Within this average of 9 lessons, 8 lessons were from the CommonLit Text Library and 1 was a CommonLit Target Lesson. Table 3 shows the breakdown of digital CommonLit lessons submitted by grade level.

Table 3. CommonLit Usage by Grade Level (SY 2025-2026)

Grade	Students	Average Digital CommonLit Lessons	Average Digital CommonLit 360 Lessons	Average Digital CommonLit Text Library Lessons	Average Digital CommonLit Target Lessons
All Students	9,153	9	0	8	1
Grade 4	2,390	12	0	11	1
Grade 5	2,464	11	0	10	2
Grade 6	1,664	6	0	5	1
Grade 7	936	5	0	5	1
Grade 8	1,699	4	0	4	0

CommonLit Usage and MCAP Performance

To test the relationship between CommonLit usage and students' performance on the MCAP, we conducted two statistical models. The first model examined MCAP ELA scaled scores as a continuous outcome variable. The second model examined MCAP Performance Levels as a binary outcome variable, testing whether students were more likely to score at the proficient levels (Levels 3 and 4). Both models accounted for students being nested within classrooms and the possibility that students in the same classroom may perform more similarly than students in a different classroom. Both models also controlled for the following covariates: baseline MCAP ELA scores from the 2024-2025 school year, grade level (reference group: Grade 4), gender (reference group: male), race/ethnicity (reference group: White students), special education status (reference group: no), English Language Learner (ELL) status (reference group: no), gifted/talented status (reference group: no), and Title I eligibility (reference group: no). Scaled scores from the MCAP were standardized (z-scored) within grade prior to analysis to account for grade-level differences in scaled scores.

Table 4 displays the results of the model that examined MCAP ELA scaled scores. Baseline MCAP ELA scaled scores was the strongest predictor of students' MCAP ELA scaled scores, $b = 0.71$, $SE = 0.008$, $t = 94.33$, $p < .001$. Importantly, there was a statistically significant relationship between CommonLit usage and MCAP ELA scaled scores, $b = 0.01$, $SE = 0.002$, $t = 4.94$, $p < .001$. That is, even after controlling for baseline scores and demographic

variables, the number of digital CommonLit lessons a student submitted relative to their classmates was a significant predictor of their scaled scores on the MCAP ELA assessment. This result can be translated into about 0.3 points per 2 additional CommonLit lessons, or about 1 point per 6 additional CommonLit lessons. Figure 2 shows the relationship between CommonLit usage and MCAP performance.

Table 4. Hierarchical Linear Model Results: MCAP ELA Scaled Scores

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.02	0.14
Residual	0.27	0.52

Predictor	Estimate	*SE*	*df*	*t*-value	*p*-value
(Intercept)	−0.02	0.03	285.25	−0.89	0.375
Baseline MCAP scores	0.71	0.01	9040.91	94.33	< .001
Digital Lessons Submitted (centered within classrooms)	**0.01**	**0.00**	**8930.27**	**4.94**	**< .001**
Grade 5	−0.02	0.03	244.37	−0.64	0.524
Grade 6	−0.02	0.04	216.00	−0.42	0.672
Grade 7	−0.02	0.04	238.43	−0.49	0.622
Grade 8	−0.02	0.04	214.41	−0.51	0.613
Female	0.10	0.01	9027.42	8.91	< .001
Asian	0.11	0.03	9061.53	3.65	< .001
Black	−0.04	0.02	9047.28	−2.37	0.018
Hispanic/Latino	−0.02	0.02	9081.39	−0.92	0.360
Two or More Races	−0.03	0.02	9102.74	−1.71	0.088
Special Education Status	−0.23	0.02	9127.00	−12.55	< .001
ELL Status	−0.10	0.04	9129.81	−2.50	0.013
Gifted/Talented Status	0.26	0.02	9125.49	13.90	< .001
Title I Eligibility	−0.05	0.03	271.82	−1.55	0.122

Predictor	Estimate	*SE*	*df*	*t*-value	*p*-value
Classroom Effect: Baseline MCAP scores	0.14	0.03	311.50	5.51	< .001
Classroom Effect: Digital Lessons Submitted	0.01	0.00	218.53	4.54	< .001

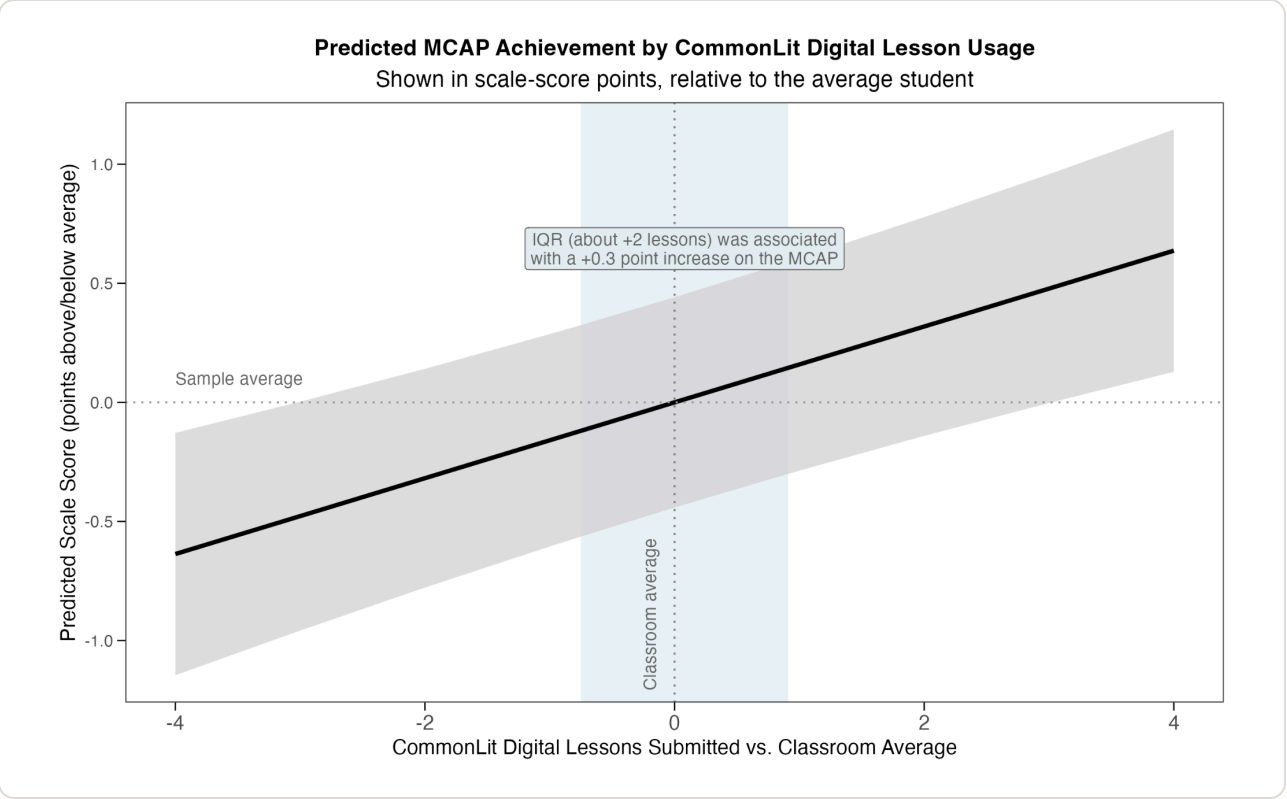


Figure 2. Predicted scaled scores on the MCAP, relative to the district average, based on CommonLit digital lessons submitted relative to one's classmates.

Note. Predicted scale score points relative to the average student, converted from z-scores using the overall standard deviation (19.74). The shaded blue band marks the 25th to 75th percentile of relative usage (-1 to +1 lessons), and the gray band is the 95% confidence interval.

Table 5 displays the results of the model that examined MCAP performance levels. Similar to the first model, baseline MCAP ELA scaled scores was the strongest predictor of students' MCAP performance levels, $b = 2.67$, $SE = 0.069$, $z = 38.43$, $p < .001$. Notably, the relationship between CommonLit usage and MCAP performance levels was statistically significant, $b = 0.03$, $SE = 0.010$, $t = 2.91$, $p = .004$. The number of digital CommonLit lessons a student submitted relative to their classmates was a significant predictor of their MCAP performance levels, even after controlling for baseline scores and

demographic variables. This result can be translated into a 1 point difference in likelihood, per 2 additional CommonLit lessons. Figure 3 shows the relationship between CommonLit usage and MCAP performance.

Table 5. Mixed-Effects Logistic Regression Results: MCAP ELA Performance Levels

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.26	0.51

Predictor	Estimate	*SE*	*z*-value	*p*-value
(Intercept)	1.22	0.12	10.35	< .001
Baseline MCAP scores	2.67	0.07	38.43	< .001
Digital Lessons Submitted (centered within classrooms)	**0.03**	**0.01**	**2.91**	**0.004**
Grade 5	−0.58	0.13	−4.42	< .001
Grade 6	0.58	0.17	3.47	0.001
Grade 7	0.95	0.20	4.83	< .001
Grade 8	0.12	0.18	0.71	0.480
Female	0.36	0.07	5.31	< .001
Asian	0.06	0.21	0.27	0.788
Black	−0.11	0.10	−1.05	0.292
Hispanic/Latino	0.09	0.14	0.64	0.523
Two or More Races	−0.03	0.12	−0.25	0.802
Special Education Status	−0.54	0.10	−5.22	< .001
ELL Status	−0.54	0.25	−2.20	0.028
Gifted/Talented Status	0.95	0.19	4.96	< .001
Title I Eligibility	−0.03	0.16	−0.20	0.844
Classroom Effect: Baseline MCAP scores	0.51	0.13	3.93	< .001
Classroom Effect: Digital Lessons Submitted	0.02	0.01	2.72	0.007

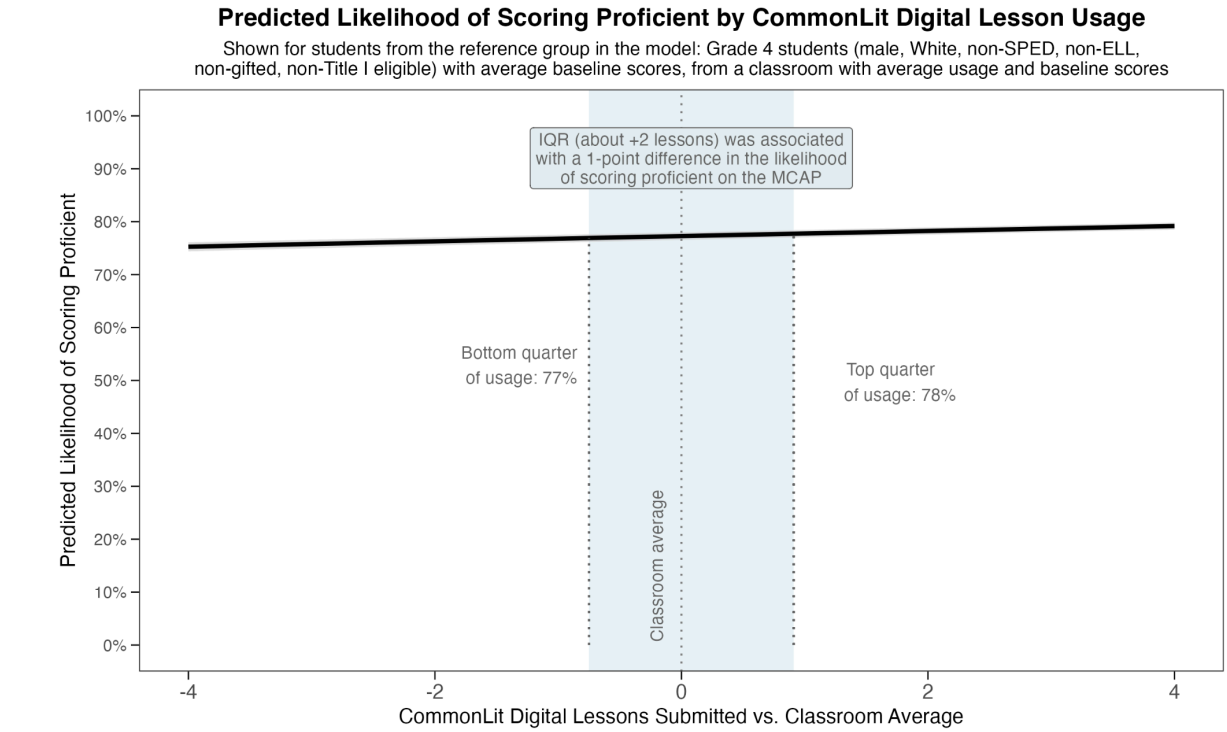


Figure 3. Predicted likelihood of scoring proficient on the MCAP, based on CommonLit digital lessons submitted relative to one's classmates.

Note. Reference-group student: Grade 4, male, White, non-SPED, non-ELL, not gifted and talented, with average prior achievement in a classroom with average usage and average prior achievement. Predictions are population-level (fixed-effects) predictions. The shaded blue band marks the 25th to 75th percentile of relative usage.

We conducted several additional analyses to test whether these results were robust (see Appendix B). The first set of analyses tested whether the results of both models held under different scenarios: excluding small teacher clusters ($n < 10$), excluding usage outliers (> 80 lessons), and excluding influential clusters (Cook's D). Under all three scenarios, CommonLit usage remained a significant predictor of MCAP ELA scaled scores and performance levels, demonstrating that results were robust. The second set of analyses tested whether CommonLit usage predicted performance on alternative measures of students' end-of-year ELA achievement: the reading-only component of the MCAP and the iReady assessment. Across both outcome measures, CommonLit usage was a significant predictor of students' performance, suggesting that results were not limited to the overall MCAP ELA performance measures. The third set of analyses tested whether the relationship between CommonLit usage and students' end-of-year ELA

achievement was moderated by student demographics or schools' Title I eligibility. In all but one analysis, the relationship between CommonLit usage and students' ELA achievement was not significantly moderated by these variables. This suggested that the effect of CommonLit was robust across student demographic groups and schools. In sum, all three sets of additional analyses provided support for the positive relationship between CommonLit usage and student achievement.

Discussion

This study provided support for the efficacy of CommonLit for students' ELA achievement among a sample of 9,153 students across Grades 4-8 in 43 schools. Across the 2024-2025 and 2025-2026 school years, 66% of students in the sample scored at the proficient levels (Levels 3 and 4) on the MCAP. During the 2025-2026 school year, students in the sample submitted an average of 9 digital CommonLit lessons. Several statistical models tested the relationship between students' CommonLit usage and performance on the MCAP and related measures of ELA achievement. Across all models, CommonLit significantly predicted students' performance. The relationship between CommonLit usage and students' performance was robust under several scenarios (e.g., excluding outliers), across multiple types of ELA achievement measures (e.g., the reading-only component of the MCAP, iReady assessment), and when testing for potential differences across student groups and school types. In practical terms, for MCAP ELA scaled scores, results can be translated into an increase of about 1 scaled score point per 6 additional CommonLit lessons. For MCAP performance levels, results can be translated into an increase of 1 percentage point in likelihood of scoring proficient (Levels 3 and 4) per 2 additional CommonLit lessons. In sum, this study supported the effectiveness of CommonLit for students' ELA achievement.

Limitations

Observational Study Design

This was an observational study. Students were not randomly assigned to levels of CommonLit usage, so students who submitted more lessons may differ from their classmates in ways beyond usage, such as motivation, attendance, or access to technology. This study attempted to address potential selection bias by comparing students who were taught by the same teacher and by statistically controlling for prior-

year MCAP achievement, classroom mean prior achievement, grade, and student characteristics. Because unmeasured differences between students (such as differences in motivation or attendance) cannot be ruled out, claims from this study should be made with caution, as this design did not support causal conclusions. However, because random assignment of instructional materials may not be feasible in practice, this design provides a way to examine the association between CommonLit usage and a state assessment outcome while balancing demands on district and school leaders.

Usage and Instructional Context

Across several different assessments, this study documented a positive relationship between CommonLit usage and students' ELA achievement. However, this study was not a comprehensive test of the efficacy of all instructional resources used in classrooms in the 2025-2026 school year. In this sample, CommonLit appeared to be used as a supplemental resource, with a majority of the usage associated with CommonLit's Text Library. In this way, the potential impact that CommonLit could have had on students' scores may have been relatively limited compared to a resource that was implemented as the core ELA curriculum. A limitation of this study is the lack of insight into students' level of usage of the core curricular resource as well as what other resources they may have received (e.g., Tier 2 intervention resources). Future research may attempt to conduct a more comprehensive study of all literacy resources used in a classroom and estimate their independent and combined effects on students' ELA achievement. This kind of research could enable educators to make more informed decisions about how to maximize the benefits of different literacy resources in the classroom.

Outcome Scaling and Sample

MCAP scale scores are not vertically scaled, so scores were standardized within grade using the means and standard deviations of this sample. As a result, findings describe students' relative standing within this sample, not absolute change or standing relative to the state. The sample included only students in Grades 4-8 who submitted at least one digital lesson and had MCAP scores in both years, so the study cannot compare CommonLit users with non-users. The analyses assumed a linear relationship between usage and outcomes. Nonlinearity was not tested, although a larger CommonLit study found that reading growth increased more gradually at the highest levels of usage

(CommonLit, 2026). Finally, the sample came from a single district, so findings may not generalize to other districts, grades, or school years.

Future Research

CommonLit regularly evaluates the effectiveness of its program on student achievement. Including this study, CommonLit has published 12 ESSA Tier 3 studies that provided "promising evidence" and one ESSA Tier 2 study that provided "moderate evidence" for the efficacy of CommonLit. CommonLit is currently undertaking a large-scale quasi-experimental study of the impact of the CommonLit 360 curriculum on middle school students' performance on state ELA tests. This study is being supported by an early-stage Education Innovation and Research (EIR) Grant from the U.S. Department of Education and will be externally evaluated by researchers at CPRE (Consortium for Policy Research in Education) at Teachers College, Columbia University. With positive findings, this EIR study would meet ESSA Tier 2 standards for "moderate evidence" and represent the largest and most rigorous evaluation of CommonLit to date. Beyond EIR, future studies may examine the impact of CommonLit across additional states and on measures beyond student achievement. Through research partnerships and grants, CommonLit aims to evaluate the effectiveness of its programs for students and teachers.

References

AdHoc Analytics. (2018). *CommonLit Digital: Evidence of effectiveness*. CommonLit. [Link]

Brown, M. (2014). *Investigating the effects of CommonLit on student engagement in middle school reading classrooms* (ED662965). ERIC. [Link]

CommonLit. (2023a). *Report on school year 22-23: Students using CommonLit saw significant reading gains*. [Link]

CommonLit. (2023b). *CommonLit assessments are strongly predictive of state test results in New York and Florida*. CommonLit Blog. [Link]

CommonLit. (2024a). *Report on school year 23-24: Students using CommonLit saw significant reading gains*. [Link]

CommonLit. (2024b). *Case Study Report: CommonLit 360 at Lakewood Middle School in New Jersey*. [Link]

CommonLit. (2024c). *Overview of assessment series test design & specifications*. [Link]

CommonLit. (2024d). *CommonLit usage and writing growth in school year 2022-2023*. [Link]

CommonLit. (2025). *Results from Harford County Public Schools (MD): School year 2023-2024*. [Link]

CommonLit. (2026a). *Report on school year 2024-2025: Students using CommonLit saw significant reading gains*. [Link]

CommonLit. (2026b). *Report on school year 2025-2026: Students using CommonLit saw significant reading gains*. [Link]

CommonLit, & Mathematica. (2023). *Effects of CommonLit 360 on reading comprehension, student engagement, and cultural awareness*. [Link]

Fahle, E., Kane, T. J., Reardon, S. F., & Staiger, D. O. (2024). *The first year of pandemic recovery: A district-level analysis*. Education Recovery Scorecard. [Link]

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

Kuhfeld, M., & Lewis, K. (2025, March 18). *5 years after COVID-19 hit: Test data converge on math gains, stalled reading recovery*. Brookings. [Link]

Kuhfeld, M., Lewis, K., & Peltier, T. (2023). Reading achievement declines during the COVID-19 pandemic: Evidence from 5 million U.S. students in grades 3–8. *Reading and Writing, 36*(2), 245–261. [Link]

Lee, A., & Shah, M. (2023). *CommonLit: Examining impacts among urban public school students* [ESSA Evidence Level III study]. LearnPlatform. [Link]

Lee, A., & Shah, M. (2024). *CommonLit: Examining impacts among Florida middle school students (2022-23)* [ESSA Evidence Level II study]. LearnPlatform. [Link]

National Center for Education Statistics. (2025). *2024 NAEP reading assessment: Results at grades 4 and 8*. U.S. Department of Education, Institute of Education Sciences. [Link]

Truckenmiller, A. J., Cho, E., Bourgeois, S., & Friedman, E. (2024). Uses and misuses of commercial reading assessment: An applied framework for decision making in grades K through 6. *The Reading Teacher, 77*(5), 609–623. [Link]

Appendices

Appendix A: Descriptive Statistics

Table A1. Descriptive Statistics for the MCAP by Grade (School Year 2024-2025)

	MCAP ELA Scaled Score				MCAP Reading Scaled Score			
Grade	**Mean**	***SD***	**Median**	**Range**	**Mean**	***SD***	**Median**	**Range**
Grade 3	759	22	759	694 – 836	53	8	53	29 – 87
Grade 4	756	20	756	689 – 812	52	8	52	25 – 79
Grade 5	755	17	756	679 – 807	52	7	52	19 – 78
Grade 6	761	18	761	708 – 820	54	8	54	31 – 81
Grade 7	760	16	760	708 – 819	54	7	54	30 – 81

Note. Grade levels reflect SY 2024-2025 enrollment. The few students who repeated grade 8 in SY 2025-2026 were omitted from this table of descriptives to protect student anonymity.

Table A2. Descriptive Statistics for the MCAP by Grade (School Year 2025-2026)

	MCAP ELA Scaled Score				MCAP Reading Scaled Score			
Grade	**Mean**	***SD***	**Median**	**Range**	**Mean**	***SD***	**Median**	**Range**
Grade 4	758	21	757	704 – 828	53	8	52	29 – 85
Grade 5	754	19	754	698 – 825	51	8	51	28 – 85
Grade 6	759	18	759	707 – 821	54	8	54	31 – 81
Grade 7	761	18	761	710 – 812	55	8	55	32 – 80
Grade 8	758	21	758	685 – 821	53	7	53	26 – 77

Table A3. Descriptive Statistics for the iReady Assessment by Grade

	Fall iReady Lexile Level				Spring iReady Lexile Level			
Grade	**Mean**	***SD***	**Median**	**Range**	**Mean**	***SD***	**Median**	**Range**
Grade 4	718	238	735	−365 – 1300	823	235	865	−260 – 1300
Grade 5	827	234	865	−320 – 1400	924	225	960	−245 – 1400
Grade 6	928	223	965	−190 – 1455	988	238	1025	−255 – 1500
Grade 7	990	238	1015	−155 – 1575	1037	255	1080	−205 – 1600
Grade 8	1045	219	1075	−110 – 1675	1067	258	1114	−135 – 1700

Appendix B: Robustness Checks

We conducted three sets of analyses to test whether the key findings in the main text were robust. The first set of analyses tested whether the results of both models held under different scenarios: excluding small teacher clusters ($n < 10$), excluding usage outliers (> 80 lessons), and excluding influential clusters (Cook's D). Under all three scenarios, CommonLit usage remained a significant predictor of MCAP ELA scaled scores and performance levels, demonstrating that results were robust. Tables A4 and A5 show the results of these robustness checks, focusing on the within-classroom estimate of CommonLit usage and reporting the number of students and teachers retained in the analysis for context.

Table A4. Robustness checks for the model that examined MCAP ELA scores

Check	**Number of students**	**Number of teachers**	**Within-classroom estimate of CommonLit usage**	***p*-value**
Model examining MCAP ELA scores	9,153	240	0.01	< .001
Excluding small teacher clusters ($n < 10$)	9,075	220	0.01	< .001
Excluding usage outliers (> 80 lessons)	9,150	240	0.01	< .001
Excluding influential clusters (Cook's D)	9,050	238	0.01	< .001

Table A5. Robustness checks for the model that examined MCAP ELA performance levels

Check	Number of students	Number of teachers	Within-classroom estimate of CommonLit usage	*p*-value
Model examining MCAP ELA performance levels	9,153	240	0.03	.004
Excluding small teacher clusters (n < 10)	9,075	220	0.03	.004
Excluding usage outliers (> 80 lessons)	9,150	240	0.03	.002
Excluding influential clusters (Cook's D)	8,858	234	0.04	.002

The second set of analyses tested whether CommonLit usage predicted performance on alternative measures of students' end-of-year ELA achievement: the reading-only component of the MCAP and Spring iReady Lexile Levels. For the analysis examining the reading-only component of the MCAP, the baseline measure was the reading-only component of the MCAP in the 2024-2025 school year. For the analysis examining Spring iReady Lexile Levels, the baseline measure was Fall iReady Lexile Levels. Across both outcome measures, CommonLit usage was a significant predictor of students' performance, suggesting that results were not limited to the overall MCAP ELA performance measures. Table A6 shows results of the two models, focusing on the effect of CommonLit usage.

Table A6. Summary of HLM Results for Alternative Measures of ELA Achievement

	Estimate	*SE*	*df*	*t*-value	*p*-value
Effect of CommonLit Usage on MCAP Scaled Scores (Reading-Component Only)	0.01	0.002	8937.68	4.35	< .001
Effect of CommonLit Usage on Spring iReady Lexile Levels	0.01	0.002	8454.01	2.86	.004

The third set of analyses tested whether the relationship between CommonLit usage and students' end-of-year ELA achievement was moderated by student characteristics or schools' Title I eligibility. To examine this, statistical interactions were modeled between CommonLit usage and all covariates:

1. CommonLit Usage x Baseline Scores
2. CommonLit Usage x Gender
3. CommonLit Usage x Race
4. CommonLit Usage x Special Education Status
5. CommonLit Usage x ELL Status
6. CommonLit Usage x Gifted/Talented Status
7. CommonLit Usage x Title I Eligibility

Given the number of statistical interactions being tested, we applied the Holm correction (Holm, 1979) to address the issue of multiple comparisons and reduce the possibility of false positives (Type I errors). We examined whether CommonLit was moderated by these variables across four outcome measures: MCAP ELA scaled scores, MCAP performance levels, MCAP reading-only component scores, and iReady Lexile levels. Across the 28 interactions examined, the relationship between CommonLit usage and student performance was not significantly moderated in 27 instances and only significantly moderated in one instance (see Table A7). This suggested that results were largely robust across student demographic groups and school types. Notably, there were no statistically significant interactions from the three models that examined MCAP performance. Table A7 provides a high-level summary of the results, focusing on the statistical significance of the interactions across models.

Table A7. Summary of Moderation Analyses

	p-values (with Holm corrections)			
Interaction	**MCAP ELA Scaled Score**	**MCAP Performance Levels**	**MCAP Reading Component**	**Spring iReady Lexile Level**
CommonLit Usage x Baseline Scores	1.000	1.000	1.000	0.035
CommonLit Usage x Gender	1.000	1.000	1.000	1.000
CommonLit Usage x Race	0.956	1.000	1.000	0.535
CommonLit Usage x SPED Status	1.000	1.000	1.000	1.000
CommonLit Usage x ELL Status	0.956	0.339	1.000	1.000
CommonLit Usage x Gifted Status	1.000	0.270	1.000	1.000
CommonLit Usage x Title I Eligibility	1.000	1.000	1.000	0.904

The only significant interaction was between baseline iReady Lexile Levels and CommonLit usage on Spring iReady Lexile Levels. We further examined this significant interaction through simple slopes analyses. Specifically, the relationship between CommonLit usage and Spring iReady Lexile Levels was examined at different levels of Fall iReady baseline scores. Among students with relatively lower Fall iReady baseline scores (1 standard deviation below the average), the relationship between CommonLit usage and Spring iReady Lexile Levels was positive and significant, $b = 0.009$, $SE = 0.003$, $t = 3.10$, $p = .002$. By contrast, among students with relatively higher Fall iReady baseline scores (1 standard deviation above the average), the relationship between CommonLit usage and Spring iReady Lexile Levels was not significant, $b = -0.003$, $SE = 0.003$, $t = -1.03$, $p = .303$. This showed that the relationship between CommonLit usage and Spring iReady Lexile Levels was stronger for students with relatively lower Fall iReady Lexile Levels compared to students with relatively higher Fall iReady Lexile Levels.

Overall, the relationship between CommonLit usage and students' ELA achievement was largely not significantly moderated by student characteristics or school types. This suggested that the effect of CommonLit was robust. In sum, all three sets of additional analyses provided support for the positive relationship between CommonLit usage and student achievement.

Appendix C: CommonLit Assessment Series Positively Correlated with the MCAP

Exploratory correlation analyses were conducted with the subset of students who took the Assessment Series at the beginning and end of the year (*n* = 786 students). Table A8 shows the correlations between the students' scores on the CommonLit Assessment Series (Pre-Assessment and Post-Assessment) and their scores on the MCAP (ELA scaled scores and reading-only scaled scores).

Table A8. Correlations between the MCAP and the CommonLit Assessment Series (*n* = 786)

	CommonLit: Pre-Assessment		CommonLit: Post-Assessment	
MCAP measure	**Correlation**	**95% CI**	**Correlation**	**95% CI**
MCAP ELA Scaled Score	.75	(.72, .78)	.76	(.73, .79)
MCAP Reading Scaled Score	.76	(.73, .79)	.76	(.73, .79)

All correlations between the Assessment Series and the MCAP were positive and statistically significant, *p's* < .001. Correlations ranged from *r* = .75 to *r* = .76, indicating moderately-strong, positive correlations. For moderate-stakes decisions, such as using assessment data to differentiate instructional time, guidelines from Truckenmiller and colleagues (2024) state that correlations with state tests should be positive and above *r* = .60. These results indicate that students' performance on the CommonLit Assessment Series were highly related to their performance on the MCAP, providing support for the concurrent validity of the Assessment Series. These findings replicated previous research that documented a positive relationship between students' scores on the CommonLit Assessment Series and the MCAP (CommonLit, 2025). These findings are also consistent with the positive correlations found between the CommonLit Assessment Series and state ELA tests in New York, Florida, and New Jersey (CommonLit, 2023b; CommonLit 2024b). Overall, these results are promising and suggest that CommonLit's assessments may be useful resources in state test preparation.