



Report on School Year 2025-2026: Students Using CommonLit Saw Significant Reading Gains



Table of Contents

Executive Summary	4
Introduction	5
Research Questions	6
Method	6
Sample	6
Measures	9
CommonLit Usage	9
CommonLit Assessment Series	10
Results	12
How Did Students and Teachers Use CommonLit within the Usage Groups?	12
Within CommonLit Usage Groups in Each Grade Band	12
How Did Students Perform on the CommonLit Assessment Series?	14
Proficiency within Usage Groups (Grades 6-12)	14
Proficiency within Usage Groups (Middle School)	15
Proficiency within Usage Groups (High School)	16
Change in Proficiency Among Students who Scored Below Proficient (Grades 6-12)	17
Change in Proficiency Among Students who Scored Below Proficient (Middle School)	18
Change in Proficiency Among Students who Scored Below Proficient (High School)	19
Was CommonLit Usage Correlated With Student Reading Growth?	20
Reading Growth by Usage Group (Grades 6-12)	21
Reading Growth by Usage Group (Middle School)	22
Reading Growth by Usage Group (High School)	23
Percentile Point Gains	24
Robustness Checks	28
Discussion	29
Limitations	30
Post-Hoc Quasi-Experimental Study Design	30
Usage Group Criteria	31
No Recorded Usage Comparison Group	32

Ongoing Research	33
References	34
Appendices	35
Appendix A: Inclusion Criteria for the Student Sample	35
Appendix B: Results for the Elementary Grade Band (Grades 3-5)	36
How Did Students and Teachers Use CommonLit?	36
How Did Students Perform on the CommonLit Assessment Series?	36
Was CommonLit Usage Correlated with Student Reading Growth?	38
Appendix C: Additional Descriptives	41
Performance Categories (Elementary School)	41
Performance Categories (Grades 6-12)	42
Performance Categories (Middle School)	44
Performance Categories (High School)	45
Proficiency (Elementary School)	46
Proficiency (Grades 6-12)	47
Proficiency (Middle School)	48
Proficiency (High School)	49
Percent Increase in Proficiency Within Usage Groups (Elementary School)	50
Percent Increase in Proficiency Within Usage Groups (Grades 6-12)	50
Percent Increase in Proficiency Within Usage Groups (Middle School)	52
Percent Increase in Proficiency Within Usage Groups (High School)	52
Appendix D: Hierarchical Models	53
Elementary School	57
Grades 6-12	61
Middle School	66
High School	70
Appendix E: Baseline Equivalence Tests	74

Executive Summary

CommonLit conducts an annual study that examines the reading growth of students who took the CommonLit Pre-Assessment and Post-Assessment in a given school year. For the 2025-2026 school year, this study included 322,012 students in Grades 3-12 taught by 6,248 teachers in 1,243 schools in 46 states and Washington D.C. Across all grade bands, students with high CommonLit usage saw significantly higher reading gains compared to students with no recorded CommonLit usage or low to moderate CommonLit usage. Compared to their peers in the sample who would score at the median without using CommonLit, grade 6-12 students who engaged deeply with the CommonLit 360 curriculum (by accessing half or more of the reading or writing lessons in at least three units) saw reading gains of 7 percentile points. High school students who engaged deeply with the CommonLit 360 curriculum saw reading gains of 13 percentile points, compared to their peers in the sample who would score at the median without using CommonLit. Middle school students who engaged deeply with CommonLit 360 saw reading gains of 5 percentile points, compared to their peers in the sample who would score at the median without using CommonLit. Finally, elementary school students with high usage of CommonLit (36 or more lessons) saw reading gains of 4 percentile points, compared to their peers who would score at the median without CommonLit. The relationship between CommonLit usage and reading growth was positive across multiple analytic methods, including those that controlled for classroom and school effects and those that examined usage as a continuous variable. Additionally, the relationship was not significantly different among students in Title I eligible schools and other schools, suggesting similar effects across school contexts. In sum, this study demonstrated the effectiveness of CommonLit on student reading growth across several grade bands.

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. Future CommonLit studies will aim to meet ESSA Level 2 standards (“Moderate Evidence”) by examining the impact of CommonLit on outcomes beyond the Assessment Series, such as state English Language Arts (ELA) assessments, and 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 NAEP (National Assessment of Educational Progress) reading assessment showed significant declines in average reading scores and in the percent of students who scored at or above the NAEP Proficient Level (NAEP, 2024). 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, Lewis & Peltier, 2022; Kuhfeld & Lewis, 2025), with widening achievement gaps for economically disadvantaged students and traditionally underserved students (Fahle, Kane, Reardon, & Staiger, 2024).

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. To this end, CommonLit has developed a free, year-long core English Language Arts (ELA) curriculum for Grades 6-12, [recently rated as all-green for Grades 6-8 by EdReports](#), and offers wraparound support services through [its paid packages for schools and districts](#). CommonLit also continues to offer a free Text Library with thousands of free articles and reading passages for supplemental use.

In 2013, CommonLit began as a research project with 30 published online lessons (Brown, 2014). CommonLit has continued to be invested in conducting rigorous studies that evaluate the effectiveness of its core ELA program. CommonLit conducts an annual study that examines the reading growth of students who took the CommonLit Pre-Assessment and Post-Assessment in a given school year. This technical report describes the study that examined the relationship between CommonLit usage and reading growth on the CommonLit Assessment Series in the 2025-2026 school year.

Research Questions

- How did students and teachers use CommonLit during the 2025-2026 school year?
- How did students perform on the CommonLit Assessment Series during the 2025-2026 school year?
- Was CommonLit usage positively correlated with student reading growth during the 2025-2026 school year?

Method

Sample

In the 2025-2026 school year, 322,012 students in Grades 3-12 took the [CommonLit Assessment Series](#) at both the beginning and end of the year (see Appendix A for details about the inclusion criteria for this sample). Overall, these students were taught by 6,248 teachers and were enrolled in 1,243 schools across 46 states and Washington D.C. Based on classroom-level data provided by teachers when creating rosters on CommonLit, less than 1% of students were associated with rosters designated for English Language Learners and less than 1% of students were associated with rosters designated for Special Education students. Given the small number of students in this sample that were associated with these teacher-reported roster-level characteristics, this study did not include these variables in analyses.

Table 1 shows the number of students in each grade level and grade band. Grade bands were determined by the students' grade level on the Assessment Series. This report will primarily focus on the results for grades 6-12, reported both as a combined group and broken out by middle school and high school. For results on elementary school students, see Appendix B.

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

Grades	<i>N</i>	%
Elementary (3-5)	22,264	7%
Grade 3	6,398	2%
Grade 4	7,308	2%
Grade 5	8,558	3%
Middle (6-8)	119,276	37%
Grade 6	35,815	11%
Grade 7	42,269	13%
Grade 8	41,192	13%
High (9-12)	180,472	56%
Grade 9	63,448	20%
Grade 10	59,818	18%
Grade 11	37,790	12%
Grade 12	19,416	6%
Total	322,012	100%

Out of the 1,243 schools, 1,198 schools (96%) were public and 45 schools (4%) were private. In terms of [paid partnerships](#) with CommonLit, 575 schools (46%) had School Essentials (SE) PRO Plus, 620 schools (50%) had SE PRO, and 48 schools (4%) had SE with access to the Assessment Series (AS). Among schools with available Title I eligibility data, 679 schools (59%) received or were eligible to receive Title I funding in the 2025–2026 school year (or 55% of all schools in the study if schools with missing Title I eligibility data are counted as non-eligible). Additionally, 1,028 schools (83%) were eligible for the National School Lunch Program. Table 2 shows additional school-level characteristics for the entire sample and for each grade band. Note that some schools may be represented in more than one grade band if they served students in more than one grade band (e.g., a school with students in Grades 6-12).

Table 2. School characteristics for the overall sample and for each grade band.

	Overall		Grades 3-5 (Elementary School)		Grades 6-8 (Middle School)		Grades 9-12 (High School)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Schools	1,243	100%	152	100%	553	100%	665	100%
Institution Type								
Public	1,198	96%	149	98%	524	95%	641	96%
Private	45	4%	3	2%	29	5%	24	4%
Subscription Type								
SE PRO Plus	575	46%	28	18%	334	60%	275	41%
SE PRO	620	50%	80	53%	218	39%	387	58%
SE (with AS)	48	4%	44	29%	1	< 1%	3	< 1%
Title I Eligibility (Valid %)								
Eligible	679	59%	88	61%	315	64%	342	56%
Not Eligible	466	41%	57	39%	181	36%	267	44%
Title I Eligibility (% with Missing)								
Eligible	679	55%	88	58%	315	57%	342	51%
Not Eligible	466	37%	57	38%	181	33%	267	40%
Missing or Not Reported	98	8%	7	5%	57	10%	56	9%
National School Lunch Program								
Yes	1,028	83%	133	88%	450	81%	528	79%
No	77	6%	9	6%	37	7%	46	7%
Missing or Not Reported	138	11%	10	6%	66	12%	91	14%

	Overall		Grades 3-5 (Elementary School)		Grades 6-8 (Middle School)		Grades 9-12 (High School)	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Urban Categories								
City	216	17%	38	25%	88	16%	106	16%
Rural	276	22%	24	16%	142	26%	148	22%
Suburban	533	43%	73	48%	213	39%	289	43%
Town	122	10%	10	7%	52	9%	70	11%
Missing or Not Reported	96	8%	7	5%	58	10%	52	8%
Region								
Mid-Atlantic	225	18%	77	51%	66	12%	94	14%
Midwest	269	22%	23	15%	122	22%	143	22%
Northeast	105	8%	17	11%	47	8%	51	8%
South	221	18%	8	5%	97	18%	128	19%
West	286	23%	17	11%	156	28%	158	24%
Missing or Not Reported	137	11%	10	7%	65	12%	91	14%

Measures

CommonLit Usage

The primary predictor variable in this study was CommonLit usage, as measured by backend platform data. Students and teachers can access lessons across the CommonLit platform. Notably, the CommonLit 360 curriculum consists of six units per grade level across Grades 6-12. Each unit consists of Core Unit Texts that students engage in throughout the unit through Reading Lessons, Writing Lessons, and more. The CommonLit Text Library consists of thousands of articles and reading passages for Grades 3-12. Additionally, CommonLit's Target Lessons are skill-focused lessons that are designed to close specific learning gaps.

This study examined the relationship between CommonLit usage and reading growth in two ways. First, a categorical usage variable was created to identify three groups of students: 1) students with high usage that likely reflected CommonLit 360 as their core ELA curriculum, 2) students with low to moderate usage, and 3) students with No Recorded CommonLit Usage. This approach can shed light on the reading growth observed among students who engaged meaningfully with CommonLit 360 as a core ELA curriculum in comparison to students who had no recorded usage or low to moderate usage of CommonLit. Second, this study also examined the relationship between the overall quantity of CommonLit lessons accessed by students and teachers and reading growth to test whether reading growth was higher with greater CommonLit usage. This approach can reveal the relationship between overall CommonLit usage and reading growth and does not rely on researcher-created grouping variables. If the relationship between CommonLit usage and reading growth is robust, both approaches should lead to similar conclusions.

CommonLit Assessment Series

The main outcome measure in this study was reading comprehension growth, as measured by the [CommonLit Assessment Series](#). The CommonLit Assessment Series includes three benchmark reading assessments: Pre-Assessment, Mid-Year Assessment, and Post-Assessment. In the 2025-2026 school year, each assessment 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. Moreover, students' performance can be categorized into three performance categories in relation to end-of-year grade-level expectations: below or approaching grade level, on grade level, and above grade level. 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 development of the Assessment Series, see [this overview](#).

Previous research has documented positive relationships between students' scores on the CommonLit Assessment Series and state ELA tests in New York, Florida, New Jersey, and Maryland (CommonLit, 2023; CommonLit, 2024; CommonLit, 2025). Collectively, these studies provide support for the convergent validity of the Assessment Series as a measure

of student reading comprehension and suggest that educators may use these assessments in preparation for state ELA testing.

In this study, reading comprehension growth was calculated by subtracting students' Pre-Assessment scaled score from their Post-Assessment scaled score. In addition to scaled score growth, this study also examined growth in terms of meaningful change in performance categories. Students who scored "on grade level" or "above grade level" were considered to score at a proficient level, whereas students who scored "below or approaching grade level" were considered to score at a below proficient level. This study operationalized meaningful change in performance categories as scoring below proficient on the Pre-Assessment at the beginning of the year and scoring proficient on the Post-Assessment at the end of the year.

Results

How Did Students and Teachers Use CommonLit within the Usage Groups?

Within CommonLit Usage Groups in Each Grade Band

As mentioned earlier, a categorical usage variable was created to identify three groups of students. First, a “High CommonLit Usage” group consisted of students who accessed at least half of the grade-level CommonLit 360 Reading Lessons or Writing Lessons within a unit for at least three units¹. This group likely had teachers who implemented CommonLit 360 as the core ELA curriculum. Second, a “Low to Moderate CommonLit Usage” group included students who accessed some CommonLit lessons, but not at the level of depth of the first group. Finally, a “No Recorded CommonLit Usage” group consisted of students who did not submit any digital lessons and whose teachers did not download any lessons as observed using backend data. It is important to note that this usage variable relies on data captured by the CommonLit platform; it does not capture usage that occurs outside of the platform, such as resource sharing among teachers (e.g., a teacher or coach downloading materials and distributing them to colleagues who do not use the platform). As a result, some students may be categorized in a lower usage tier than their actual classroom experience.

Table 3 shows the number of students in each group within the secondary grade band (grades 6-12), as well as the middle school and high school grade bands. Across each of the grade bands, the Low to Moderate CommonLit Usage group represented the largest group of students. Moreover, the High CommonLit Usage group had the highest average number of CommonLit lessons by far, with the most notable differences seen for average CommonLit 360 lessons accessed.

¹ This definition of high usage differed from the one applied in the SY 2024-25 study. In the SY 2024-25 study, high usage was defined as accessing at least half of both the reading and writing lessons within a CommonLit 360 unit for at least three units. Moreover, there was no filtering based on grade-level lessons. The SY 2025-26 study applied an updated definition of high usage to address shifts in instructional delivery and resource sharing associated with writing lessons. That is, teachers were using writing lessons in a manner that was not trackable by the platform. The updated definition allows for high usage to be achieved through writing lessons, but does not require it.

Table 3. Usage Groups within Middle School and High School Grade Bands.

	Students	Average CommonLit Lessons	Average CommonLit 360 Lessons	Average CommonLit Library Lessons	Average CommonLit Target Lessons
Secondary Grade Band (Grades 6-12)					
High CommonLit Usage Group	33,406 (11%)	71	63	5	2
Low to Moderate CommonLit Usage Group	243,072 (81%)	17	10	7	1
No Recorded CommonLit Usage Group	23,270 (8%)	0	0	0	0
Middle School					
High CommonLit Usage Group	24,499 (21%)	66	59	5	2
Low to Moderate CommonLit Usage Group	90,504 (76%)	22	13	7	2
No Recorded CommonLit Usage Group	4,273 (4%)	0	0	0	0
High School					
High CommonLit Usage Group	8,907 (5%)	83	77	5	1
Low to Moderate CommonLit Usage Group	152,568 (85%)	15	8	6	1
No Recorded CommonLit Usage Group	18,997 (10%)	0	0	0	0

How Did Students Perform on the CommonLit Assessment Series?

Proficiency within Usage Groups (Grades 6-12)

We examined descriptive statistics of students' growth on the CommonLit Assessment Series within each of the usage groups (for additional descriptives, see Appendix C). First, we examined the percentage of students who scored at the proficient level on the Pre-Assessment and the Post-Assessment within each of the usage groups. Among students in grades 6-12 in the High CommonLit Usage group, 38% of students scored at the proficient level on the Pre-Assessment and 54% of students scored at the proficient level on the Post-Assessment, representing a 16 percentage point increase (see Figure 1). Within the Low to Moderate CommonLit Usage group, 39% of students in grades 6-12 scored at the proficient level on the Pre-Assessment and 52% of students scored at the proficient level on the Post-Assessment, reflecting a 13 percentage point increase. In the No Recorded CommonLit Usage Group, 47% of students in grades 6-12 scored at the proficient level on the Pre-Assessment and 54% of students scored at the proficient level on the Post-Assessment, which was a 7 percentage point increase. Note that students in the no recorded usage group still saw reading growth, as expected, with alternative instructional materials.

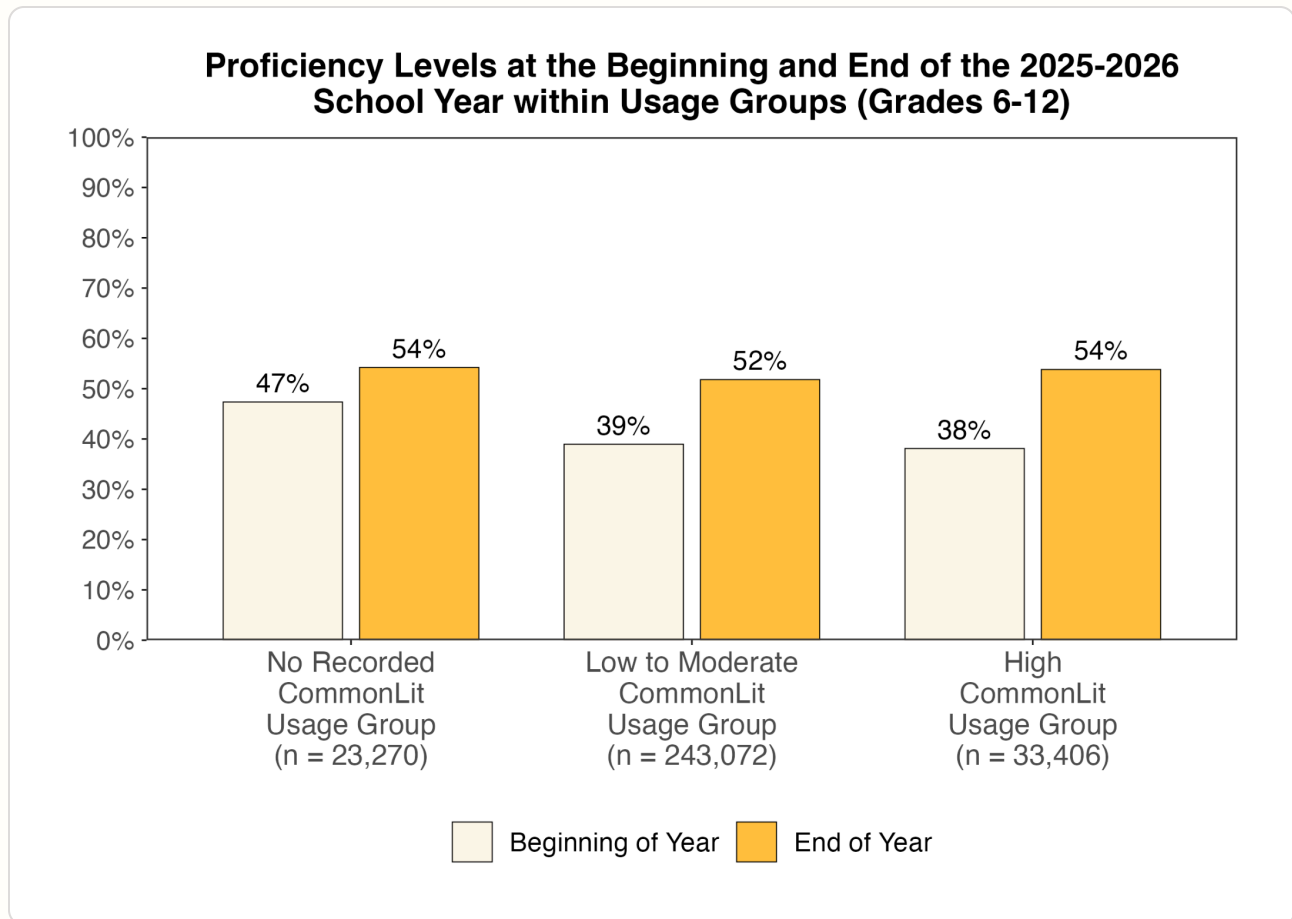


Figure 1. Percent of students in grades 6-12 (middle and high school) scoring at the proficient level at the beginning and end of the year within each usage group.

Proficiency within Usage Groups (Middle School)

Among middle school students in the High CommonLit Usage group, 40% of students were proficient on the Pre-Assessment and 53% of students were proficient on the Post-Assessment (a 13 percentage point increase; see Figure 2). In the Low to Moderate CommonLit Usage group, 41% of middle school students were proficient on the Pre-Assessment and 52% of students were proficient on the Post-Assessment (an 11 percentage point increase). Finally, in the No Recorded CommonLit Usage Group, 52% of middle school students were proficient on the Pre-Assessment and 59% of students were proficient on the Post-Assessment (a 7 percentage point increase).

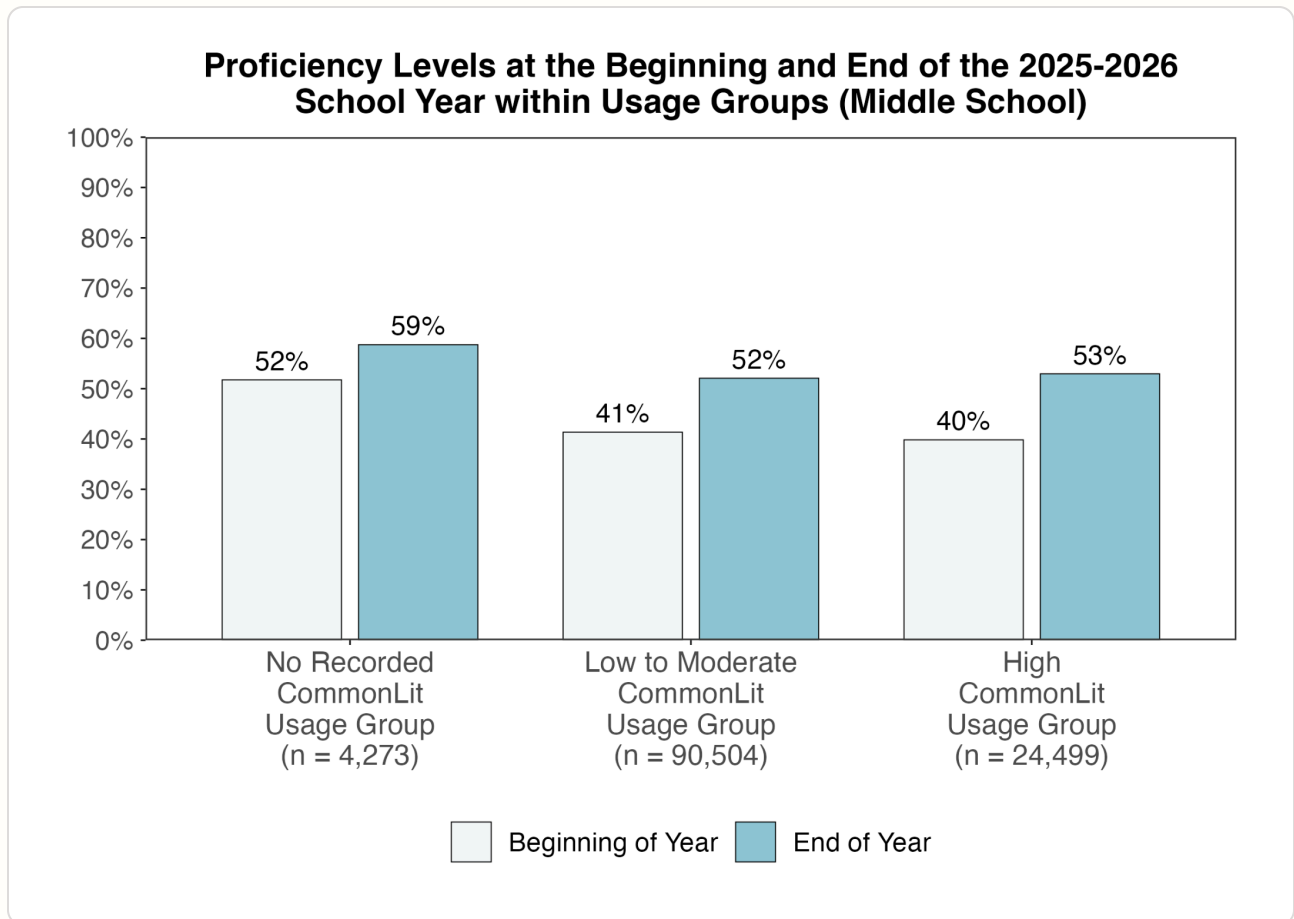


Figure 2. Percent of middle school students scoring at the proficient level at the beginning and end of the year within each usage group.

Proficiency within Usage Groups (High School)

Among high school students in the High CommonLit Usage group, 33% of students were proficient on the Pre-Assessment and 56% of students were proficient on the Post-Assessment (a 23 percentage point increase; see Figure 3). In the Low to Moderate CommonLit Usage group, 38% of high school students were proficient on the Pre-Assessment and 52% of students were proficient on the Post-Assessment (a 14 percentage point increase). Finally, in the No Recorded CommonLit Usage Group, 46% of high school students were proficient on the Pre-Assessment and 53% of students were proficient on the Post-Assessment (a 7 percentage point increase).

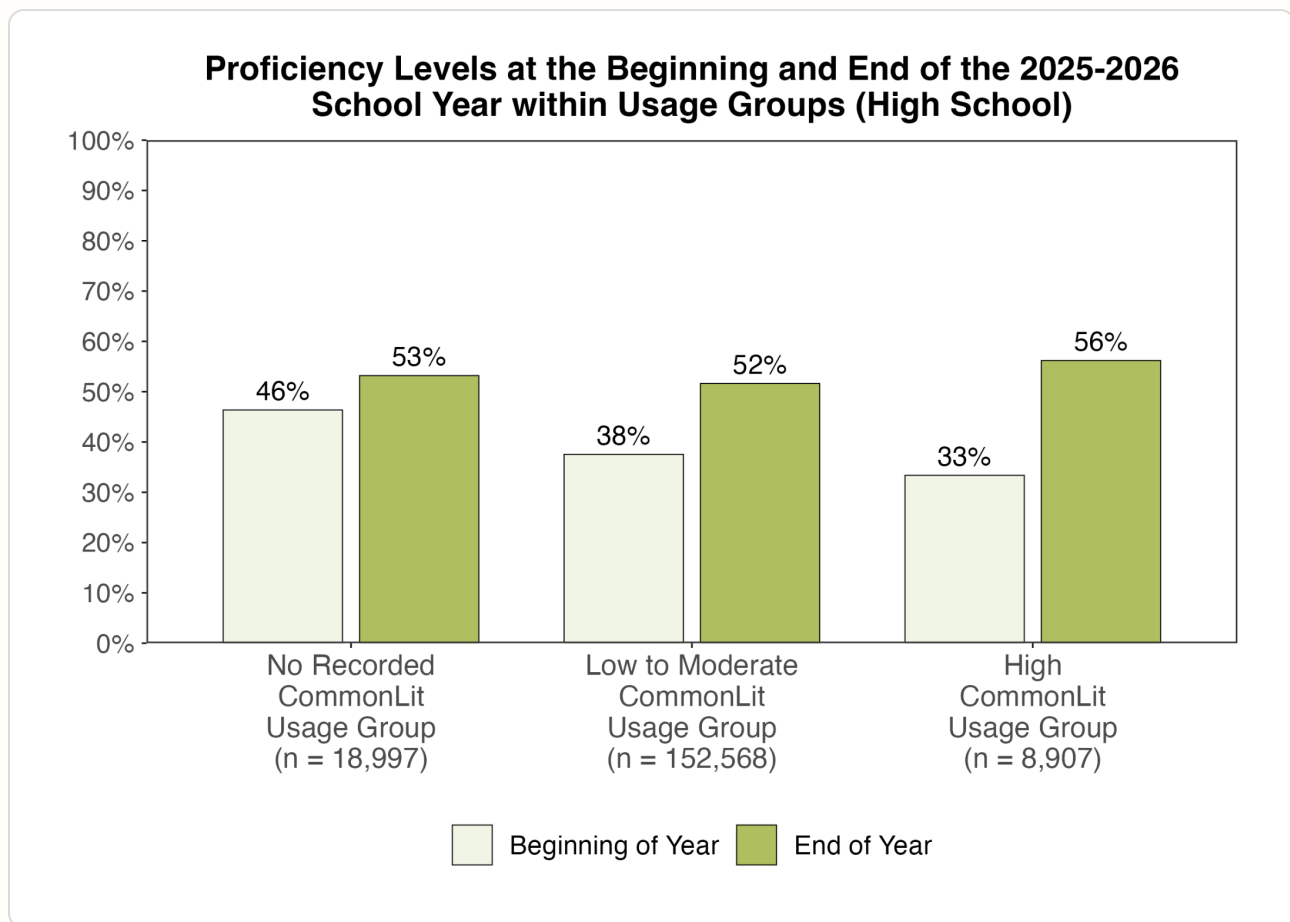


Figure 3. Percent of high school students scoring at the proficient level at the beginning and end of the year within each usage group.

Change in Proficiency Among Students who Scored Below Proficient (Grades 6-12)

Next, we examined the change in proficiency among students in grades 6-12 who scored below proficient at the beginning of the school year. Out of the 20,677 students in grades 6-12 in the High CommonLit Usage group who scored below proficient on the Pre-Assessment, 33% (6,784 students) were proficient on the Post-Assessment (see Figure 4). Out of the 148,371 students in grades 6-12 in the Low to Moderate CommonLit Usage group who scored below proficient on the Pre-Assessment, 30% (44,548) were proficient on the Post-Assessment. Finally, out of the 12,251 students in grades 6-12 in the No Recorded CommonLit Usage Group who scored below proficient on the Pre-Assessment, 27% (3,320) were proficient on the Post-Assessment.

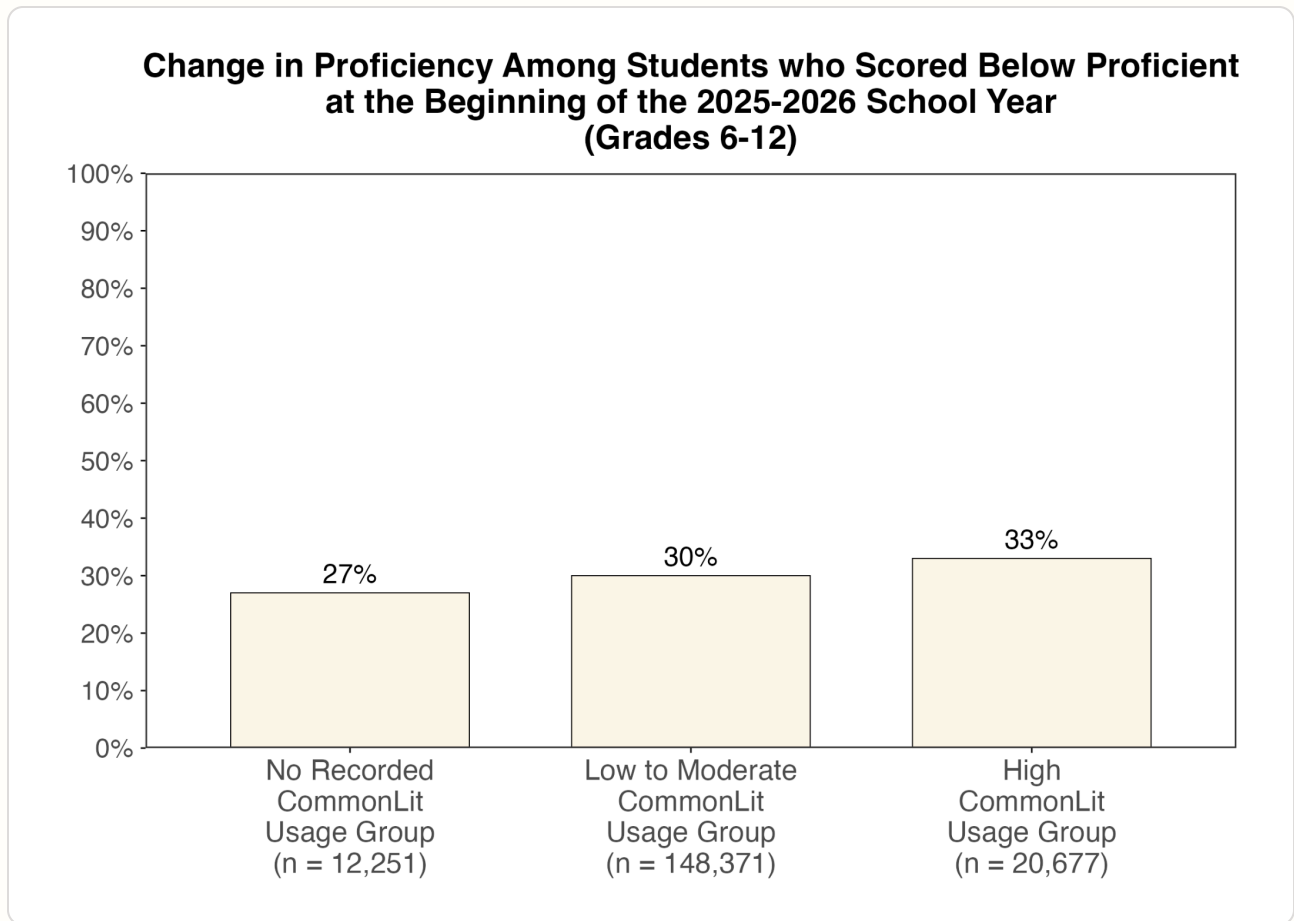


Figure 4. Percent of below proficient students in grades 6-12 (middle and high school) who scored at the proficient level on the Post-Assessment within each of the usage groups.

Change in Proficiency Among Students who Scored Below Proficient (Middle School)

Out of the 14,742 middle school students in the High CommonLit Usage group who scored below proficient on the Pre-Assessment, 30% (4,378 students) were proficient on the Post-Assessment (see Figure 5). Out of the 53,077 middle school students in the Low to Moderate CommonLit Usage group who scored below proficient on the Pre-Assessment, 27% (14,302) were proficient on the Post-Assessment. Finally, out of the 2,062 middle school students in the No Recorded CommonLit Usage Group who scored below proficient on the Pre-Assessment, 26% (535) were proficient on the Post-Assessment.

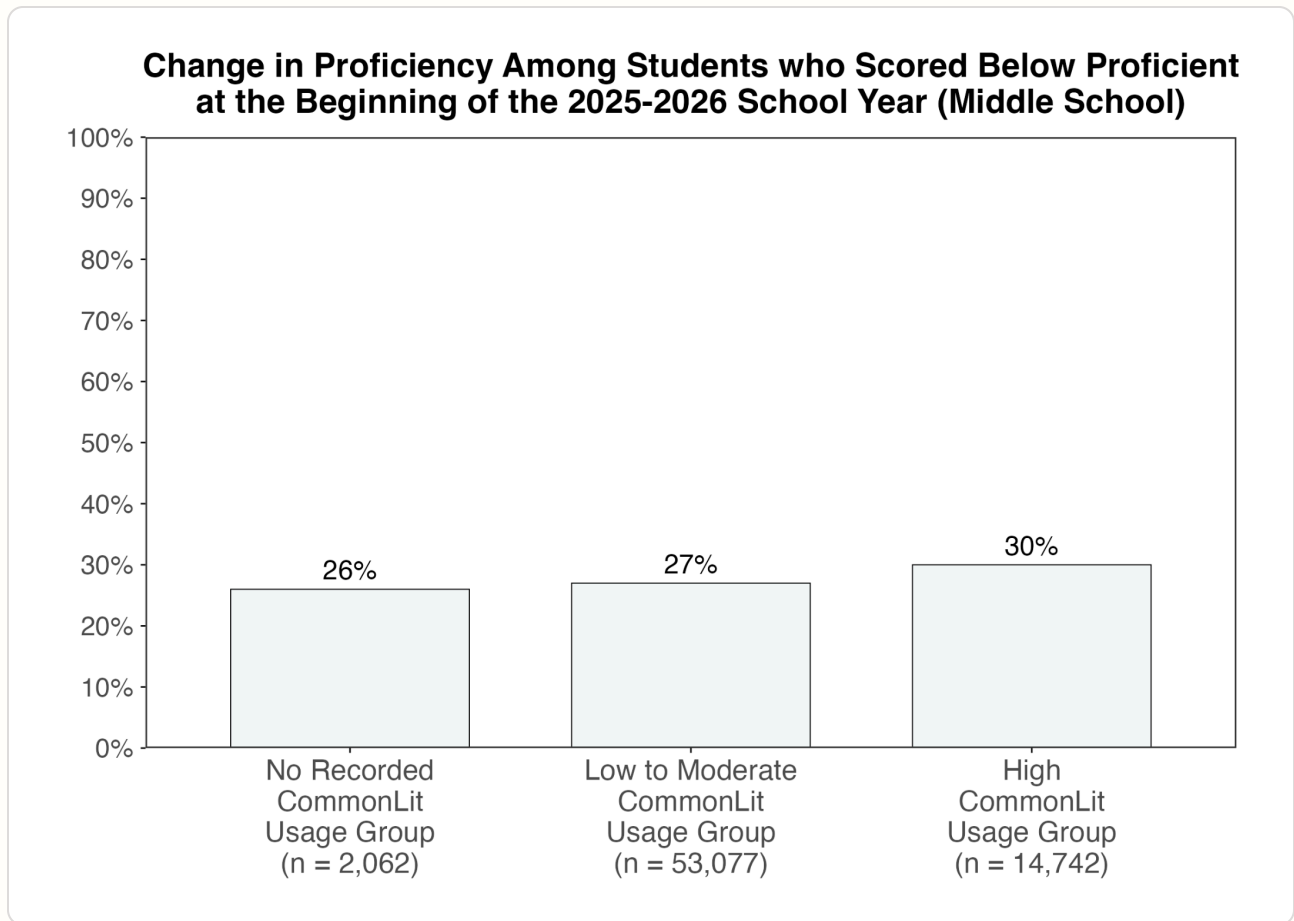


Figure 5. Percent of below proficient middle school students who scored at the proficient level on the Post-Assessment within each of the usage groups.

Change in Proficiency Among Students who Scored Below Proficient (High School)

Out of the 5,935 high school students in the High CommonLit Usage group who scored at the below proficient level on the Pre-Assessment, 41% (2,406 students) scored at the proficient level on the Post-Assessment (see Figure 6). Out of the 95,294 high school students in the Low to Moderate CommonLit Usage group who scored at the below proficient level on the Pre-Assessment, 32% (30,246) scored at the proficient level on the Post-Assessment. Finally, out of the 10,189 high school students in the No Recorded CommonLit Usage Group who scored at the below proficient level on the Pre-Assessment, 27% (2,785) scored at the proficient level on the Post-Assessment.

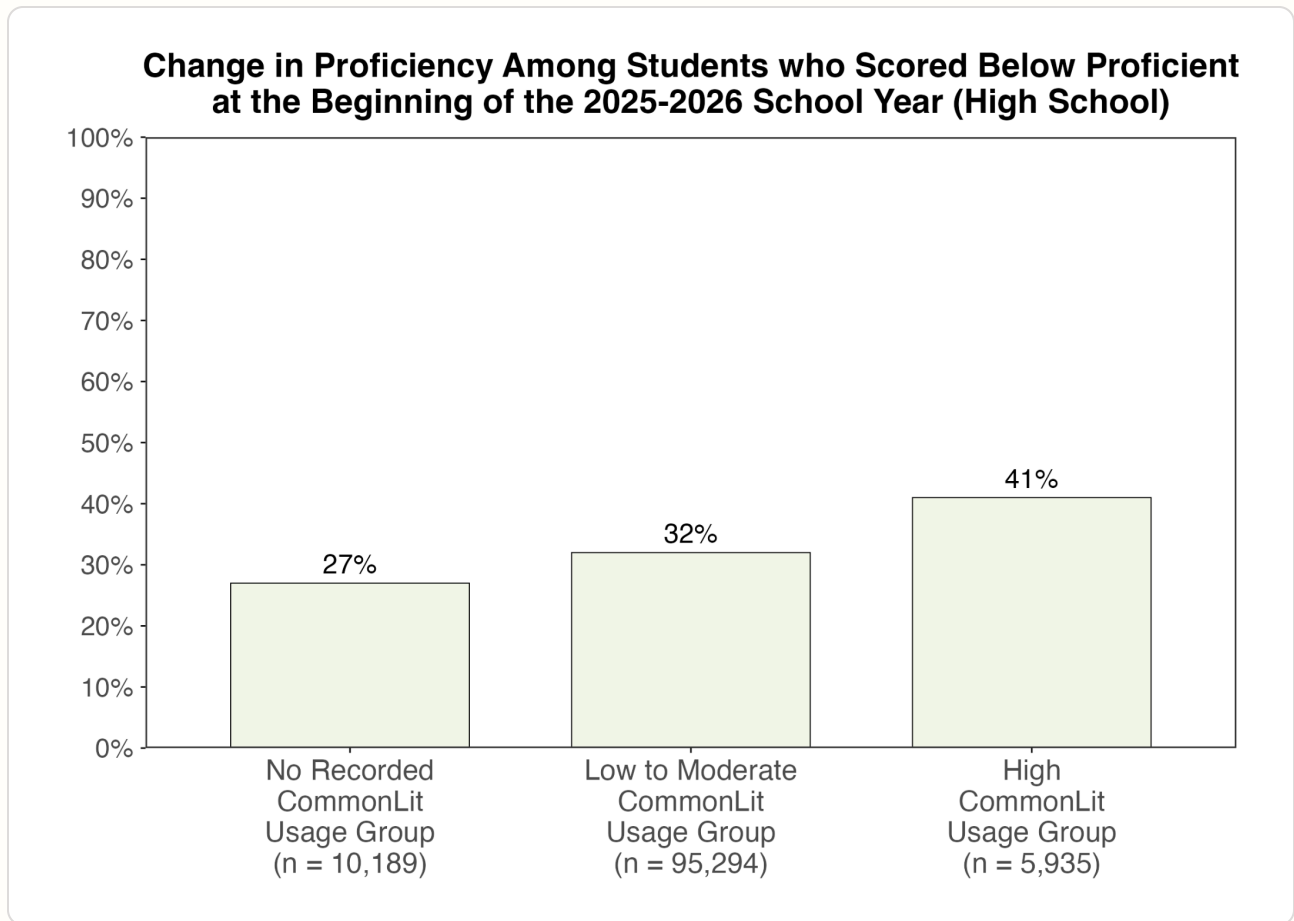


Figure 6. Percent of below proficient high school students who scored at the proficient level on the Post-Assessment within each of the usage groups.

Was CommonLit Usage Correlated With Student Reading Growth?

Next, we conducted several statistical analyses to test whether students' reading growth was significantly different across the three usage groups. First, we conducted one-way analyses of covariance (ANCOVA) that predicted students' scaled score growth from the categorical CommonLit usage group variable (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage). Importantly, these analyses controlled for the following variables that may impact students' reading growth: students' baseline scores on the CommonLit Pre-Assessment, students' grade level, and schools' Title I eligibility. The outcome variable (reading growth) and students' baseline scores on the CommonLit Pre-Assessment were standardized (z-scored) by grade level prior to analysis, in order to place students on a common scale of relative performance across

grades. Note that the values in Figures 7-9 show standardized reading growth, where 0 represents average growth within each grade level (as opposed to a lack of growth). Positive values reflect growth estimates above the group's average growth, and negative values reflect growth estimates below the group's average growth (as opposed to a decline in scores).

Reading Growth by Usage Group (Grades 6-12)

Grades 6-12 students' reading growth significantly differed across usage groups, $F(2, 299737) = 621.30, p < .001$. Post-hoc analyses showed that the differences between each of the usage groups were statistically significant. That is, students in grades 6-12 in the High CommonLit Usage group had higher reading growth than those students in both the Low to Moderate CommonLit Usage group, $t(299737) = 15.18, p < .001$, and the No Recorded CommonLit Usage group, $t(299737) = 24.25, p < .001$. Moreover, students in grades 6-12 in the Low to Moderate CommonLit usage group had higher reading growth than those students in the No Recorded CommonLit Usage group, $t(299737) = 17.79, p < .001$. In sum, greater CommonLit usage was consistently associated with greater reading growth among grades 6-12 students. Figure 7 shows the model-adjusted averages for standardized reading growth for each of the usage groups among students in grades 6-12.

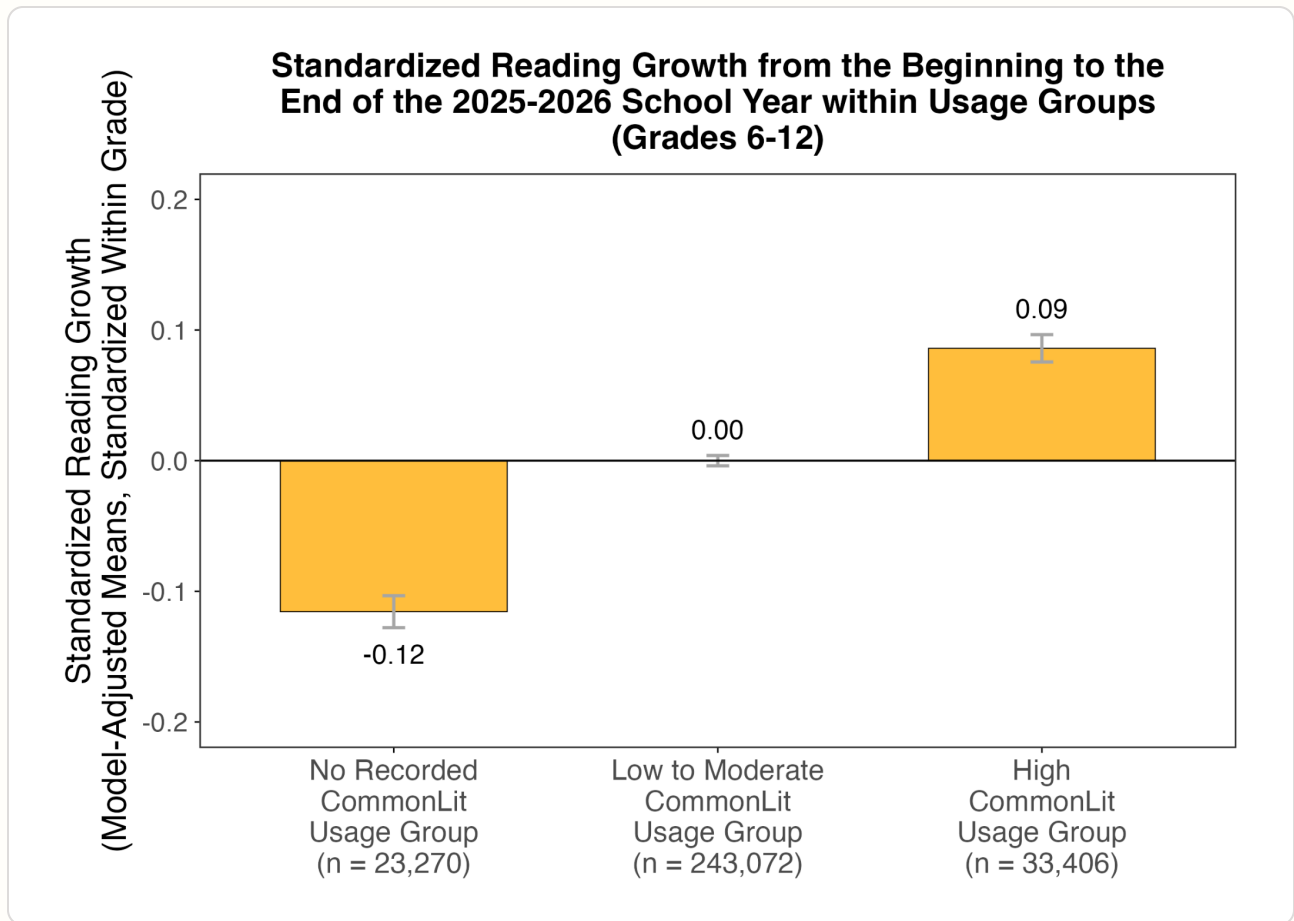


Figure 7. Model-adjusted standardized reading growth averages for each of the three usage groups (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage) among students in grades 6-12 (middle and high school). Differences between groups were statistically significant.

Reading Growth by Usage Group (Middle School)

For the middle school grade band, students' reading growth significantly differed across usage groups, $F(2, 119269) = 77.08, p < .001$. Post-hoc analyses showed that the differences between each of the usage groups were statistically significant. That is, middle school students in the High CommonLit Usage group had higher reading growth than middle school students in both the Low to Moderate CommonLit Usage group, $t(119269) = 7.81, p < .001$, and the No Recorded CommonLit Usage group, $t(119269) = 6.51, p < .001$. Moreover, middle school students in the Low to Moderate CommonLit usage group had higher reading growth than middle school students in the No Recorded CommonLit Usage group, $t(119269) = 3.29, p = .003$. In sum, greater CommonLit usage was consistently

associated with greater reading growth among middle school students. Figure 8 shows the model-adjusted standardized reading growth averages for each of the usage groups among middle school students.

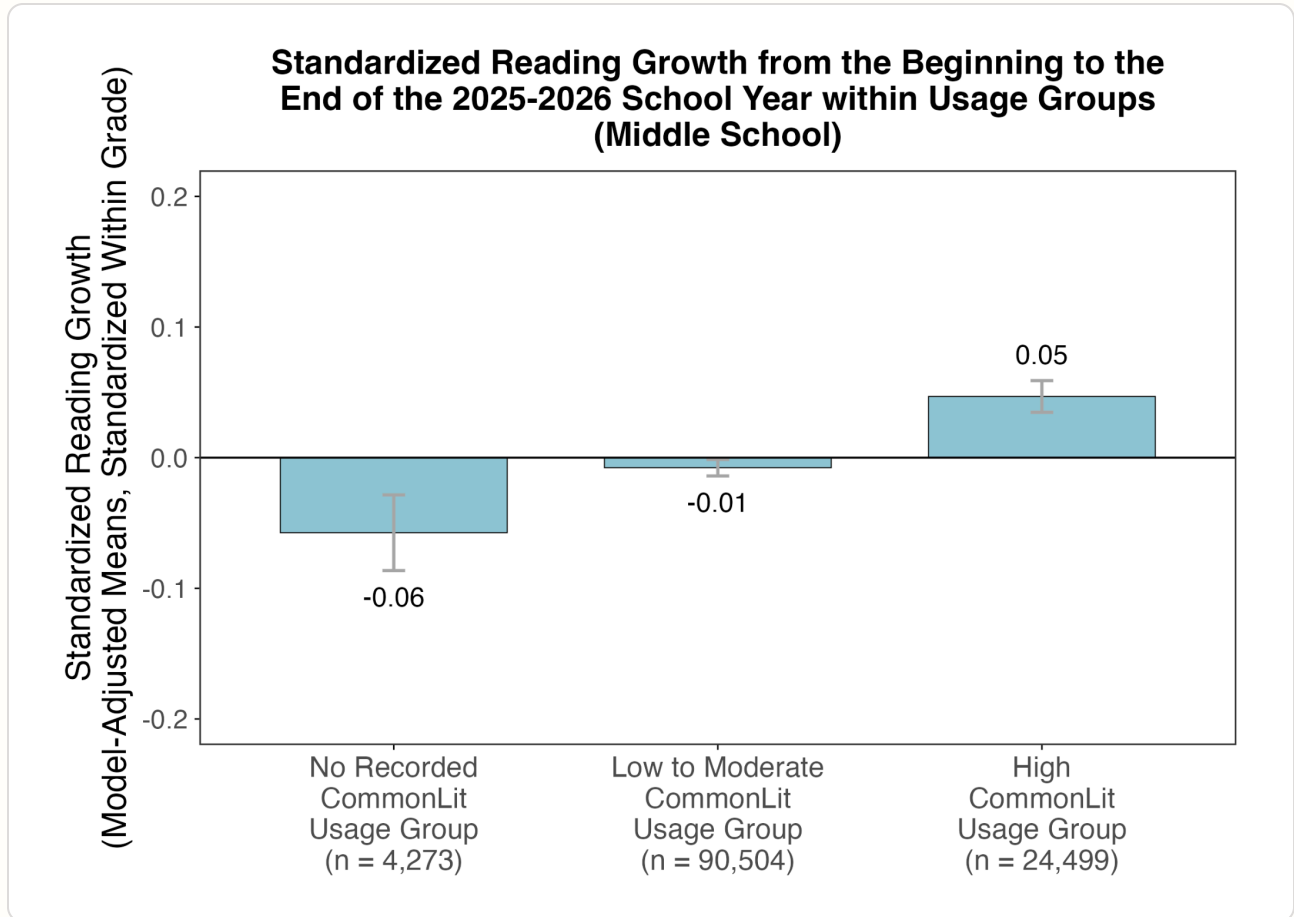


Figure 8. Model-adjusted standardized reading growth averages for each of the three usage groups (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage) among middle school students. Differences between groups were statistically significant.

Reading Growth by Usage Group (High School)

Results for the high school grade band were consistent with results among the middle school grade band. Reading growth was significantly different across usage groups among high school students, $F(2, 180464) = 698.58, p < .001$. Follow-up analyses revealed that the differences between each of the three usage groups were statistically significant. High school students in the High CommonLit Usage group had greater reading growth than high school students in both the Low to Moderate CommonLit Usage group, $t(180464) =$

16.70, $p < .001$, and the No Recorded CommonLit Usage group, $t(180464) = 24.29$, $p < .001$. High school students in the Low to Moderate CommonLit usage group also had greater reading growth than high school students in the No Recorded CommonLit Usage group, $t(180464) = 17.04$, $p < .001$. Overall, this demonstrates that greater CommonLit usage was consistently associated with greater reading growth among high school students. Figure 9 shows the model-adjusted standardized reading growth averages for each of the usage groups among high school students.

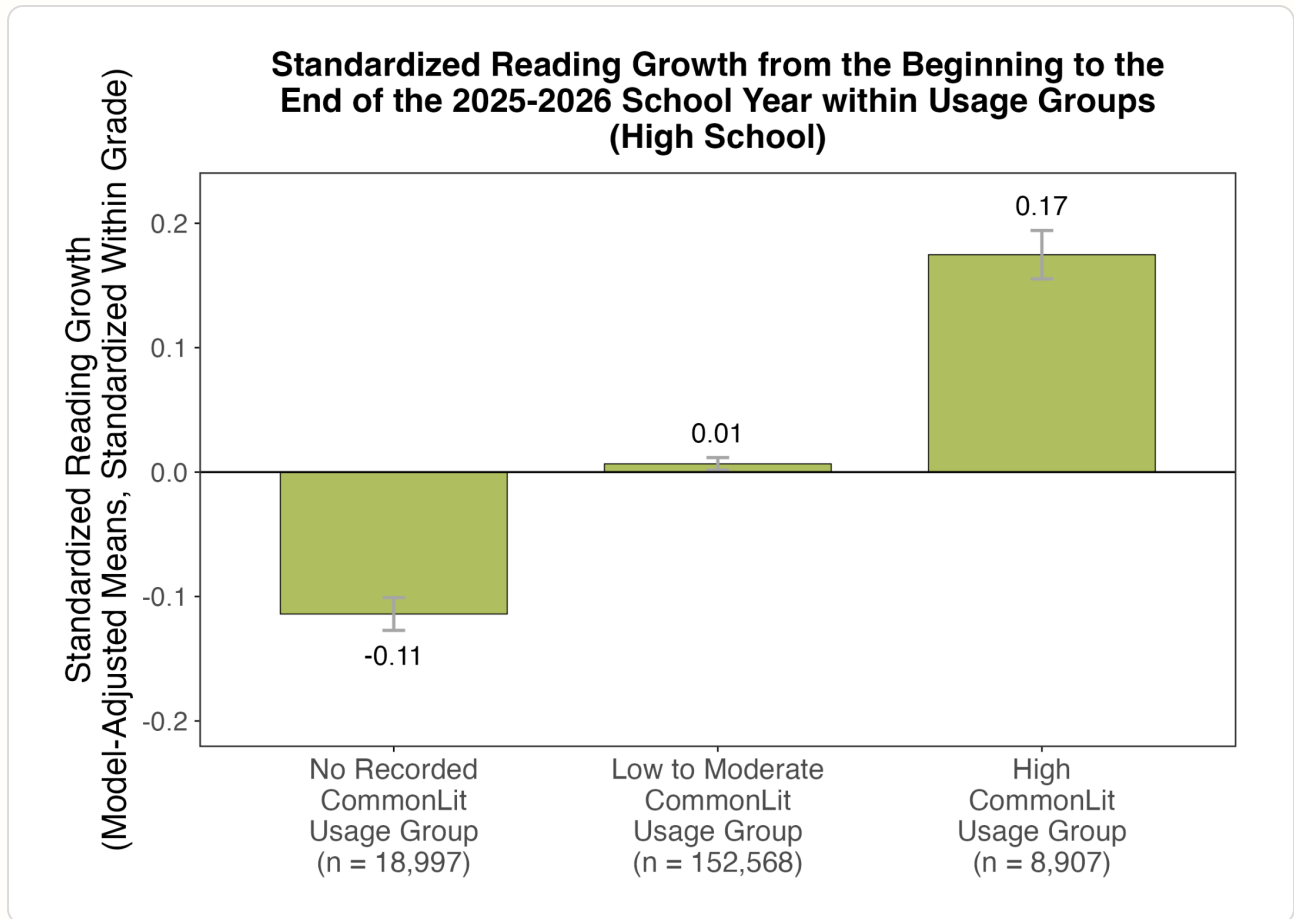


Figure 9. Model-adjusted standardized reading growth averages for each of the three usage groups (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage) among high school students. Differences between groups were statistically significant.

Percentile Point Gains

Next, we translated these results into percentile point gains to convey the practical significance of these findings. First, we calculated the effect size (Cohen's d) of students'

reading comprehension growth by grade within each usage group, and then weighted the effect size by grade level sample sizes within each grade band. This resulted in the following weighted effect sizes for grades 6-12: $d = .12$ for the No Recorded CommonLit Usage group, $d = .27$ for the Low to Moderate CommonLit Usage group, and $d = .33$ for the High CommonLit Usage group; the middle school grade band: $d = .14$ for the No Recorded CommonLit Usage group, $d = .22$ for the Low to Moderate CommonLit Usage group, and $d = .27$ for the High CommonLit Usage group; and the high school grade band: $d = .12$ for the No Recorded CommonLit Usage group, $d = .30$ for the Low to Moderate CommonLit Usage group, and $d = .48$ for the High CommonLit Usage group. Next, we calculated the difference in effect sizes and multiplied each effect size difference by 37 to approximate the translation from standard deviations to percentile points (see von Hippel, 2024).

Within grades 6-12, there was an effect size difference of .20 standard deviations between students in the High CommonLit Usage group and the No Recorded CommonLit Usage group (see Figure 10). For a student in grades 6-12 who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 7 percentile points – from the 50th to the 57th percentile. There was an effect size difference of .15 standard deviations between students in the Low to Moderate CommonLit Usage group and the No Recorded CommonLit Usage group. For a grades 6-12 student who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 5 percentile points – from the 50th to the 55th percentile. Overall, the reading growth observed among this particular sample of grades 6-12 students can be translated into 5-7 percentile point gains, when compared to grades 6-12 students who would score at the median (50th percentile) without using CommonLit.

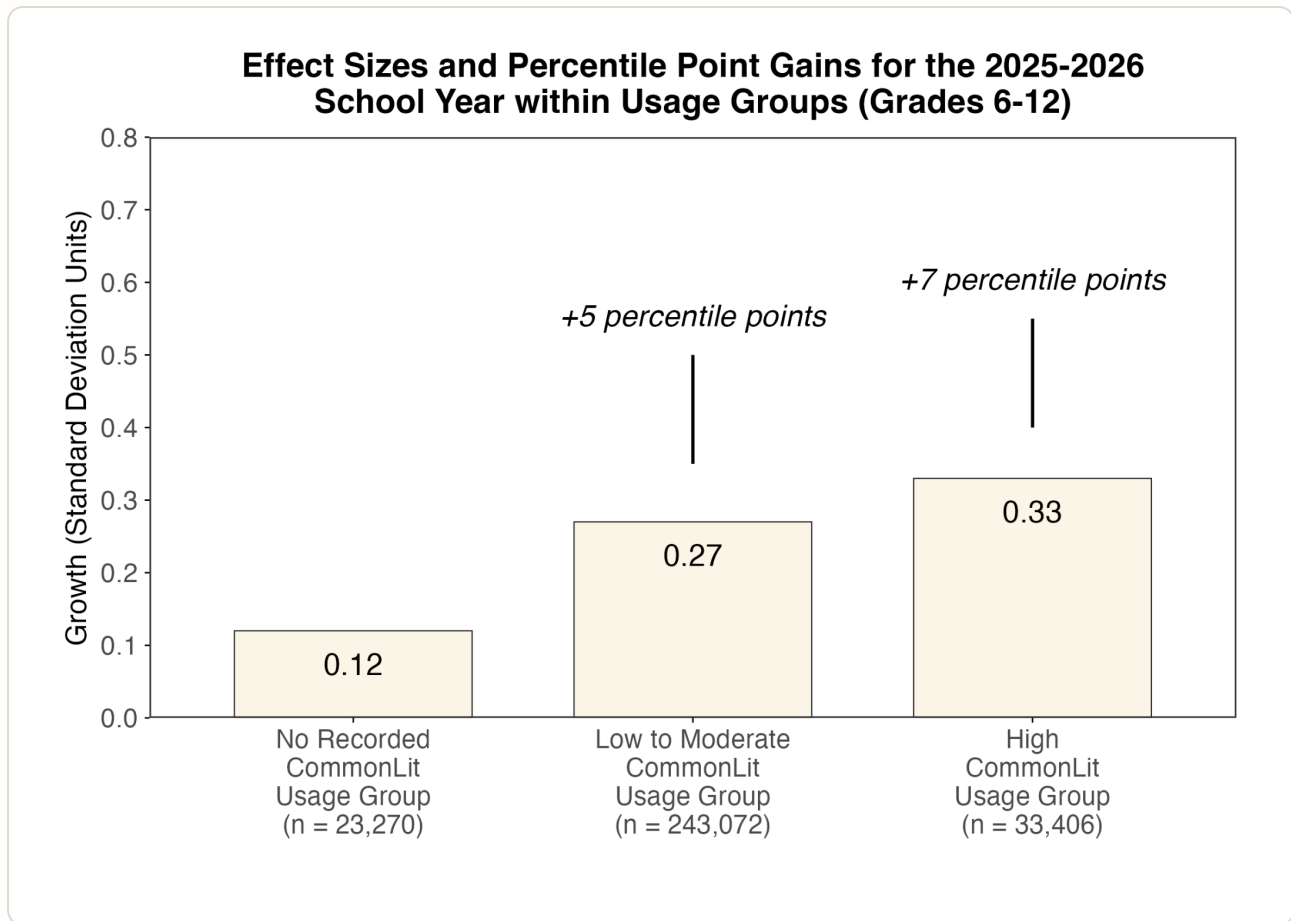


Figure 10. Effect sizes and percentile point gains for each usage group in grades 6-12 (middle and high school).

Within the middle school grade band, there was an effect size difference of .13 standard deviations between students in the High CommonLit Usage group and the No Recorded CommonLit Usage group (see Figure 11). For a middle school student who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 5 percentile points – from the 50th to the 55th percentile. There was an effect size difference of .08 standard deviations between students in the Low to Moderate CommonLit Usage group and the No Recorded CommonLit Usage group. For a middle school student who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 3 percentile points – from the 50th to the 53rd percentile. Overall, the reading growth observed among this particular sample of middle school students can be translated into 3-5 percentile point gains, when compared

to middle school students who would score at the median (50th percentile) without using CommonLit.

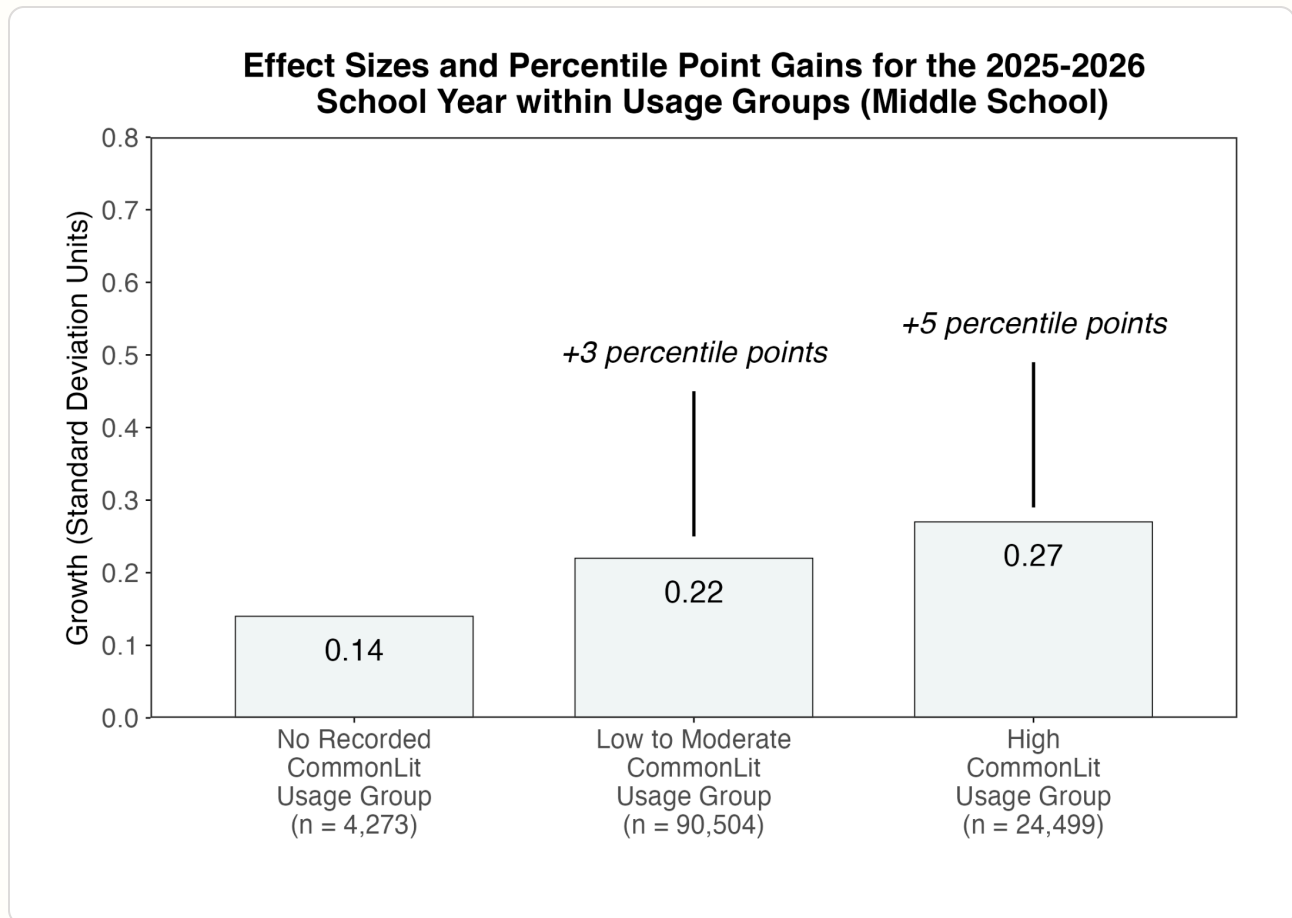


Figure 11. Effect sizes and percentile point gains for each usage group in the middle school grade band.

Within the high school grade band, there was an effect size difference of .36 standard deviations between students in the High CommonLit Usage group and the No Recorded CommonLit Usage group (see Figure 12). For a high school student who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 13 percentile points – from the 50th to the 63rd percentile. There was an effect size difference of .18 standard deviations between students in the Low to Moderate CommonLit Usage group and the No Recorded CommonLit Usage group. For a high school student who would score at the median without using CommonLit, using CommonLit in this way would raise their score by 7 percentile points – from the 50th to the 57th percentile. In sum, the reading growth observed among this specific sample of

high school students can be translated into 7-13 percentile point gains, when compared to high school students who would score at the median (50th percentile) without using CommonLit.

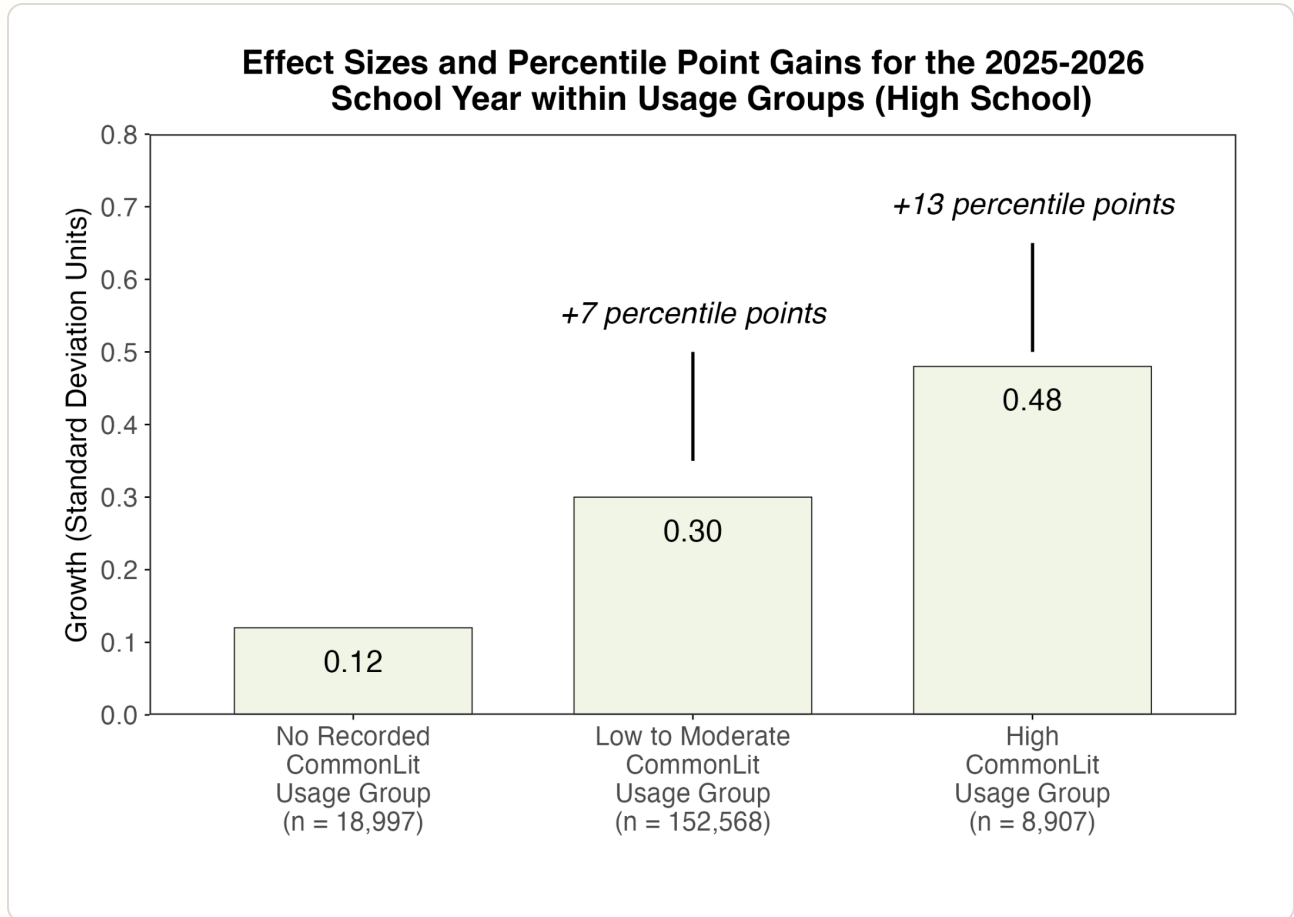


Figure 12. Effect sizes and percentile point gains for each usage group in the high school grade band.

Robustness Checks

Hierarchical models were conducted to address the nested structure of the data – i.e., students nested within classrooms nested within schools – and to test the robustness of the relationship between CommonLit usage and reading growth (see Appendix D for detailed reporting). Consistent with the ANCOVAs, these models showed that the High CommonLit Usage group saw significantly higher reading growth than the No Recorded CommonLit Usage group, Secondary grade band: $b = 0.17$, $SE = 0.016$, $t = 10.53$, $p < .001$; Middle School: $b = 0.23$, $SE = 0.034$, $t = 6.72$, $p < .001$; High School: $b = 0.11$, $SE = 0.023$, $t = 4.62$, $p < .001$. Additionally, the Low to Moderate CommonLit Usage group saw

significantly higher reading growth than the No Recorded CommonLit Usage group, Secondary grade band: $b = 0.05$, $SE = 0.013$, $t = 3.96$, $p < .001$; Middle School: $b = 0.10$, $SE = 0.031$, $t = 3.22$, $p = .001$; High School: $b = 0.03$, $SE = 0.014$, $t = 2.37$, $p = .018$. We also tested whether this relationship between CommonLit usage and student reading growth was robust across student groups. Specifically, we examined whether this relationship was moderated by schools' Title I eligibility. Results showed that the relationship between CommonLit usage and reading growth held among students at both Title I eligible schools and schools not eligible for Title I. We also tested whether this relationship was better described as non-linear. Results showed that the non-linear model fit the data significantly better than the linear model, and growth appeared to increase more gradually at the highest levels of usage. Taken together, the statistical analyses reported in this section robustly demonstrated that CommonLit usage was positively correlated with student reading growth.

Discussion

This study examined CommonLit usage data and Assessment Series data among 322,012 students in Grades 3-12 in the 2025-26 school year. For analyses, students in Grades 6-12 were placed in the "High CommonLit Usage group" if they engaged deeply with CommonLit 360 by accessing half or more of the Reading or Writing Lessons in a unit across at least three units of the curriculum. Students were placed in the "Low to Moderate CommonLit Usage group" if they accessed some CommonLit lessons but not at the depth of the first group. Finally, students were placed in the "No Recorded CommonLit Usage group" if they did not access any lessons on the platform. This study compared reading growth across these three groups of students.

First, descriptive patterns of performance gains were examined. Across all grade bands, students in the High CommonLit Usage group saw greater percentage point increases in the number of students who scored at the proficient level from the beginning to the end of the year, compared to students in the Low to Moderate CommonLit Usage group and the No Recorded CommonLit Usage group. Across all grade bands, a greater percentage of below proficient students in the High CommonLit Usage group scored at the proficient level on the Post-Assessment, compared to the other two usage groups.

Statistical analyses supported these descriptive patterns of performance gains. Specifically, in statistical models that controlled for baseline scores on the Pre-Assessment, grade level, and schools' Title I eligibility status, there was a significant positive relationship between CommonLit usage and reading growth across the secondary, middle school, and high school grade bands. Compared to their peers who would score at the median (50th percentile) without using CommonLit, engaging in CommonLit 360 deeply would raise scores by 7 percentile points (to the 57th percentile) for students in grades 6-12, by 5 percentile points (to the 55th percentile) for middle school students, and by 13 percentile points (to the 63rd percentile) for high school students.

Importantly, the relationship between CommonLit usage and reading growth was consistently positive and significant across several analytic methods. Specifically, the relationship held in models that controlled for classroom effects and school effects, suggesting that the gains observed were not due to unique factors within a particular classroom or school. Additionally, the relationship was positive and significant when examining usage as a continuous variable and as a categorical variable, suggesting that results were not reliant on researcher-defined usage groups. Moreover, the relationship held among students at Title I eligible schools and other schools, demonstrating robustness to school contexts and supporting the intention to develop instructional materials that are effective for all students. Additionally, this relationship was non-linear, with growth increasing more gradually at the highest levels of usage across all four grade bands. Findings for elementary school students were consistent with those reported for secondary, middle, and high school students and can be found in Appendix B. In sum, this study demonstrated that there was a significant positive relationship between CommonLit usage and reading growth, such that students saw greater reading gains the more deeply they engaged with CommonLit.

Limitations

Post-Hoc Quasi-Experimental Study Design

This study was a quasi-experimental study that defined treatment and comparison groups post-hoc based on observed CommonLit usage throughout the year. The lack of random assignment means that the groups may be different in ways beyond the level of

CommonLit usage. This study attempted to address potential selection bias by conducting baseline equivalence testing – i.e., comparing students' scores on the Pre-Assessment across groups (see Appendix E). Most group comparisons showed differences within the accepted guidelines from What Works Clearinghouse (WWC), and analyses involved statistical controls (for baseline scores on the Pre-Assessment, grade level, and Title I eligibility) to account for these differences. However, three comparisons exceeded the WWC threshold for baseline equivalence even with statistical adjustment: the No Recorded CommonLit Usage group versus the High CommonLit Usage group within the secondary grade band (Hedge's $g = 0.28$), middle school grade band (Hedge's $g = 0.31$) and within the high school grade band (Hedge's $g = 0.34$). As a result, findings regarding the relationship between High CommonLit Usage and reading growth for grades 6-12 students should be interpreted with some caution, as pre-existing differences between these two groups were large enough that statistical adjustment may not fully account for their influence on the observed growth outcomes. Beyond these three comparisons, students across groups may have differed on characteristics that were unmeasured in the study, and selection bias cannot be fully eliminated without random assignment. However, because random assignment of core instructional materials may not be feasible in practice, this study implemented a design that may provide promising evidence of the effectiveness of CommonLit while balancing demands on district and school leaders. Accordingly, claims from this study should be made with caution, as this design did not support strong causal inference.

Usage Group Criteria

The criteria used to define usage groups were specific to this study and differed from the criteria used in previous studies of CommonLit usage. Although the key finding of a significant positive relationship between CommonLit usage and reading gains has been consistent across studies, the threshold of what constitutes "high usage" has evolved over time. Therefore, a limitation is that it may not be possible to compare the outcomes of the high usage groups across studies. For example, the current study defined deep engagement with CommonLit 360 as accessing half or more of the Reading or Writing Lessons in a unit across three units, a change made from CommonLit's SY 2024-2025 study to better reflect shifts in instructional delivery in schools and to accommodate greater observed depth of engagement with Reading Lessons relative to Writing Lessons.

Although this study demonstrated that this high usage group showed the strongest outcomes, the high usage threshold is imperfect and has several limitations. For example, this threshold was agnostic to which units students and teachers accessed and in what order. Moreover, it did not reflect the most ideal implementation of the curriculum, which would be engaging deeply with all six units in order. Additionally, this threshold did not incorporate engagement in Culminating Tasks, as these are not yet instrumented into the platform. With new product developments, future studies may modify this threshold to strive for criteria that reflect more of the ideal implementation of CommonLit 360.

No Recorded Usage Comparison Group

This study made use of a convenient "no recorded usage" comparison group of students who did not access any CommonLit lessons on the platform but took the Assessment Series at the beginning and end of the year. It is important to note that this group was defined based on the absence of observable usage on the platform; as noted earlier, this does not fully capture usage that occurs outside of the platform, such as resource sharing (e.g., a teacher or coach downloading and sharing lessons with colleagues who do not use the platform), so some students in this group may have still received CommonLit 360 instruction through these unobserved channels. Although these students may also have presumably engaged in instructional materials from a different core ELA curriculum, their schools were in paid partnerships with CommonLit, suggesting the primary use case of the partnership was to access the assessment tools. A stronger comparison group would ideally include students who did not use CommonLit throughout the year at schools that were not partnered with CommonLit. This type of comparison would also require an outcome measure such as state ELA test scores, rather than CommonLit assessment scores. Moreover, it would be ideal to understand the alternative instructional materials used in the comparison group to understand whether the effects of CommonLit usage were compared to teacher-created materials, district-created materials, or materials from other ELA curriculum providers. Without this information, it can be difficult to interpret the comparisons between groups. Future research should incorporate more independent comparison groups and more clearly document the instructional alternatives experienced by the comparison students.

Ongoing Research

CommonLit is currently conducting a two-year, quasi-experimental study under the Education Innovation and Research grant program. This two-year study will examine the impact of CommonLit 360 on middle school students' state ELA test scores, using a comparative interrupted time series study design that addresses many limitations in the current study. For example, the outcome measure is independent from CommonLit. Moreover, the comparison group will be identified prior to the collection of state test data and matched on baseline ELA scores and available school-level demographics to address baseline equivalence. When possible, information about instructional alternatives will be collected to understand the contrast between treatment and comparison groups. This study also incorporates surveys and interviews to understand teachers' experiences with implementing the curriculum. Additionally, findings will be analyzed by external evaluators at the Consortium for Policy Research in Education (CPRE) at Teachers College, Columbia University. This two-year study represents the most rigorous evaluation of CommonLit to date, and is eligible to meet ESSA Tier 2 (Moderate Evidence) standards with positive findings.

References

- Brown, M. (2014). Investigating the Effects of CommonLit on Student Engagement in Middle School Reading Classrooms. <https://files.eric.ed.gov/fulltext/ED662965.pdf>.
- CommonLit (2023). CommonLit Assessments are Strongly Predictive of State Test Results in New York and Florida. <https://www.commonlit.org/blog/commonlit-assessments-are-strongly-predictive-of-state-test-results-in-new-york-and-florida/>.
- CommonLit (2024). Case Study: Lakewood Middle School's Success with CommonLit 360. <https://www.commonlit.org/blog/case-study-lakewood-middle-schools-success-with-commonlit-360/>.
- CommonLit (2025). Positive Relationship between CommonLit Usage and State ELA Scores in Maryland. <https://www.commonlit.org/blog/positive-relationship-between-commonlit-usage-and-state-ela-scores-in-maryland/>.
- 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. <https://m.weartv.com/resources/pdf/75e50529-b8f9-4b0f-9071-d8d7ef42f906-ERSReportFinal1.31.pdf>.
- Kuhfeld, M., Lewis, K., & Peltier, T. (2022). Reading achievement declines during the COVID-19 pandemic: evidence from 5 million US students in grades 3–8. Reading and Writing, 36(2), 245-261. <https://link.springer.com/content/pdf/10.1007/s11145-022-10345-8.pdf>.
- Kuhfeld, M. & Lewis K. (2025). 5 years after COVID-19 hit: Test data converge on math gains, stalled reading recovery. <https://www.brookings.edu/articles/5-years-after-covid-19-hit-test-data-converge-on-math-gains-stalled-reading-recovery/>.
- NAEP (2024). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2022 Reading Assessment (Grades 4 and 8).
- von Hippel, P. T. (2024). Multiply by 37 (or divide by 0.027): A surprisingly accurate rule of thumb for converting effect sizes from standard deviations to percentile points. Educational Evaluation and Policy Analysis, 01623737241239677. <https://doi.org/10.3102/01623737241239677>.

Appendices

Appendix A: Inclusion Criteria for the Student Sample

Before any inclusion criteria were applied, the dataset contained 2,066,519 assessments taken by 1,114,361 students. These assessments included CommonLit's Pre-Assessment, Mid-Year Assessment, Post-Assessment, and Free Reading Assessment. First, Free Reading Assessments were filtered out, as this assessment is not a part of CommonLit's Assessment Series. Next, we filtered out assessments with missing scaled scores, assessments not taken at grade level, assessments with more than 20% of answers missing, and assessments taken after the first attempt. Next, we filtered out students who did not take both the Pre-Assessment and Post-Assessment. Afterward, we filtered out students who did not take the Pre-Assessment before the Mid-Year Assessment and those who did not take the Pre-Assessment before the Post-Assessment. We then filtered out those who did not take the Pre-Assessment at the beginning of the year (7/21/25 – 11/1/25) and the Post-Assessment at the end of the year (3/1/26 – 6/30/26). Finally, we filtered out students who were not enrolled in schools within the United States. Figure A1 provides specific details about the percent of assessments or students filtered out at each stage, along with counts for each assessment type and unique students at each stage. The final sample included 322,012 students who took both the Pre-Assessment and Post-Assessment. Of this sample, 91% (291,745 students) also took the Mid-Year Assessment.

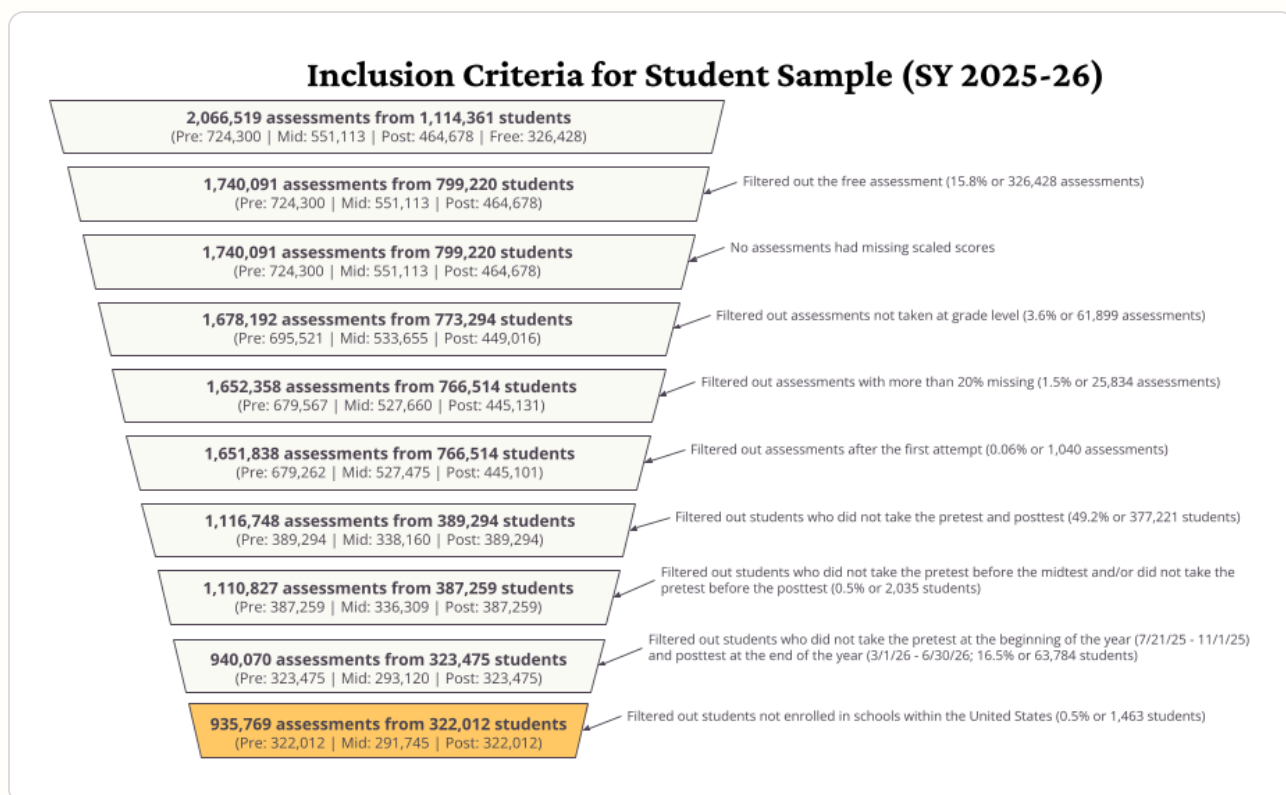


Figure A1. Funnel of inclusion criteria for the student sample (school year 2025-26).

Appendix B: Results for the Elementary Grade Band (Grades 3-5)

How Did Students and Teachers Use CommonLit?

Elementary school students submitted an average of 6 digital lessons and their teachers downloaded an average of 5 lessons, for an average of 11 lessons. Within this average, 8 lessons were from the CommonLit Text Library and 3 were CommonLit Target Lessons. To compare outcomes between elementary school students, we created usage groups. Students were placed in the “High CommonLit Usage Group” if they accessed 36 or more lessons. Students were placed in the “Low to Moderate CommonLit Usage Group” if they accessed between 1-35 lessons. Finally, students were placed in the “No Recorded CommonLit Usage Group” if they did not access any lessons through the platform.

How Did Students Perform on the CommonLit Assessment Series?

We examined descriptive statistics of students’ growth on the CommonLit Assessment Series within each of the usage groups (for additional descriptives, see Appendix C). Among elementary school students in the High CommonLit Usage group, 27% of students

scored at the proficient level on the Pre-Assessment and 48% of students scored at the proficient level on the Post-Assessment, representing a 21 percentage point increase (see Figure A2). Among elementary school students in the Low to Moderate CommonLit Usage group, 30% of students were proficient on the Pre-Assessment and 48% of students were proficient on the Post-Assessment, representing an 18 percentage point increase. Finally, among the elementary school students in the No Recorded Usage group, 32% of students were proficient on the Pre-Assessment and 49% were proficient on the Post-Assessment, representing a 17 percentage point increase.

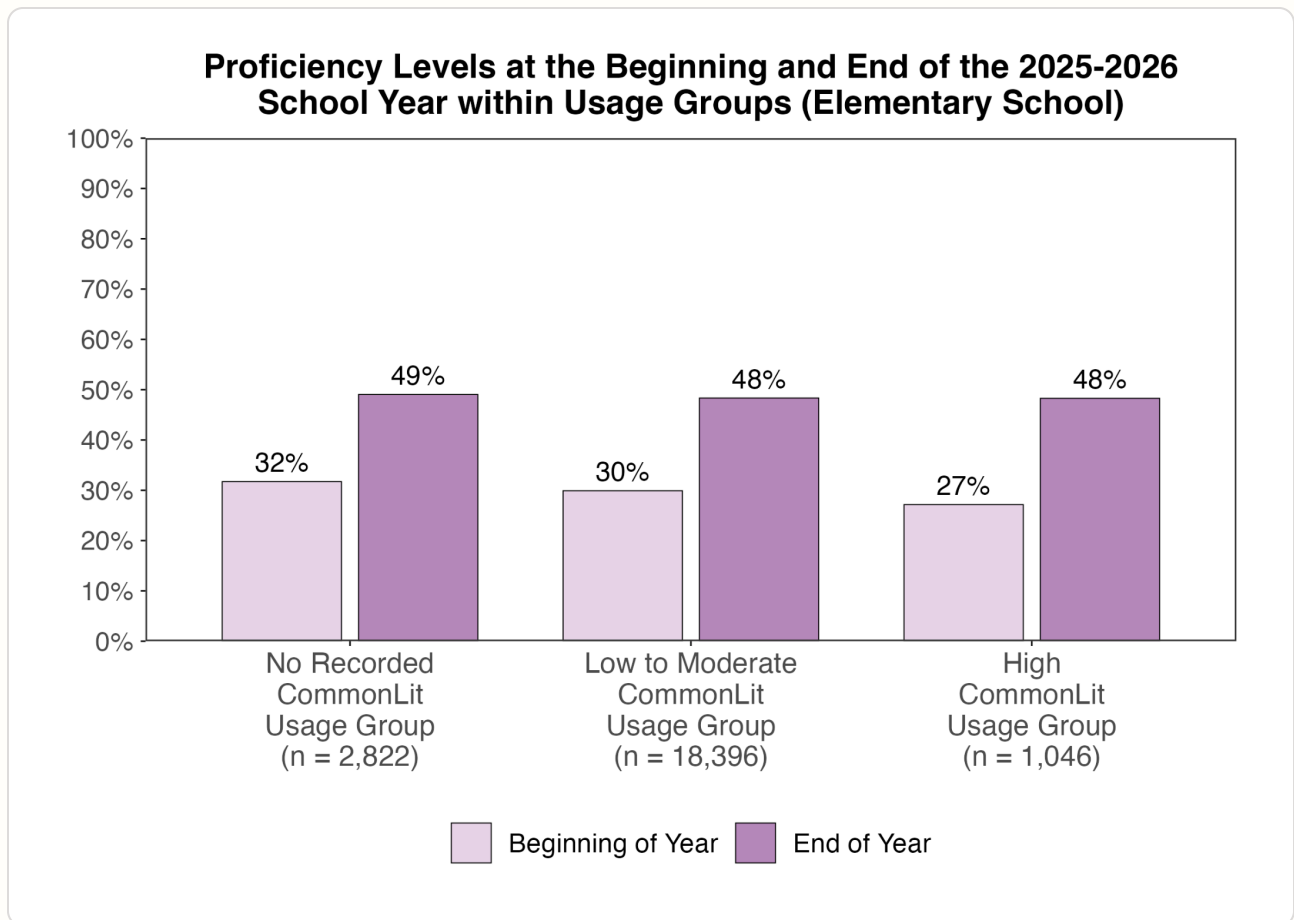


Figure A2. Percent of elementary school students scoring at the proficient level at the beginning and end of the year within each usage group.

Out of the 762 elementary school students in the High CommonLit Usage group who scored below proficient on the Pre-Assessment, 33% (251 students) were proficient on the Post-Assessment (see Figure A3). Out of the 12,894 elementary school students in the Low to Moderate CommonLit Usage group who scored below proficient on the Pre-

Assessment, 30% (3,838) were proficient on the Post-Assessment. Finally, out of the 1,927 elementary school students in the No Recorded Usage Group who scored below proficient on the Pre-Assessment, 29% (557) were proficient on the Post-Assessment.

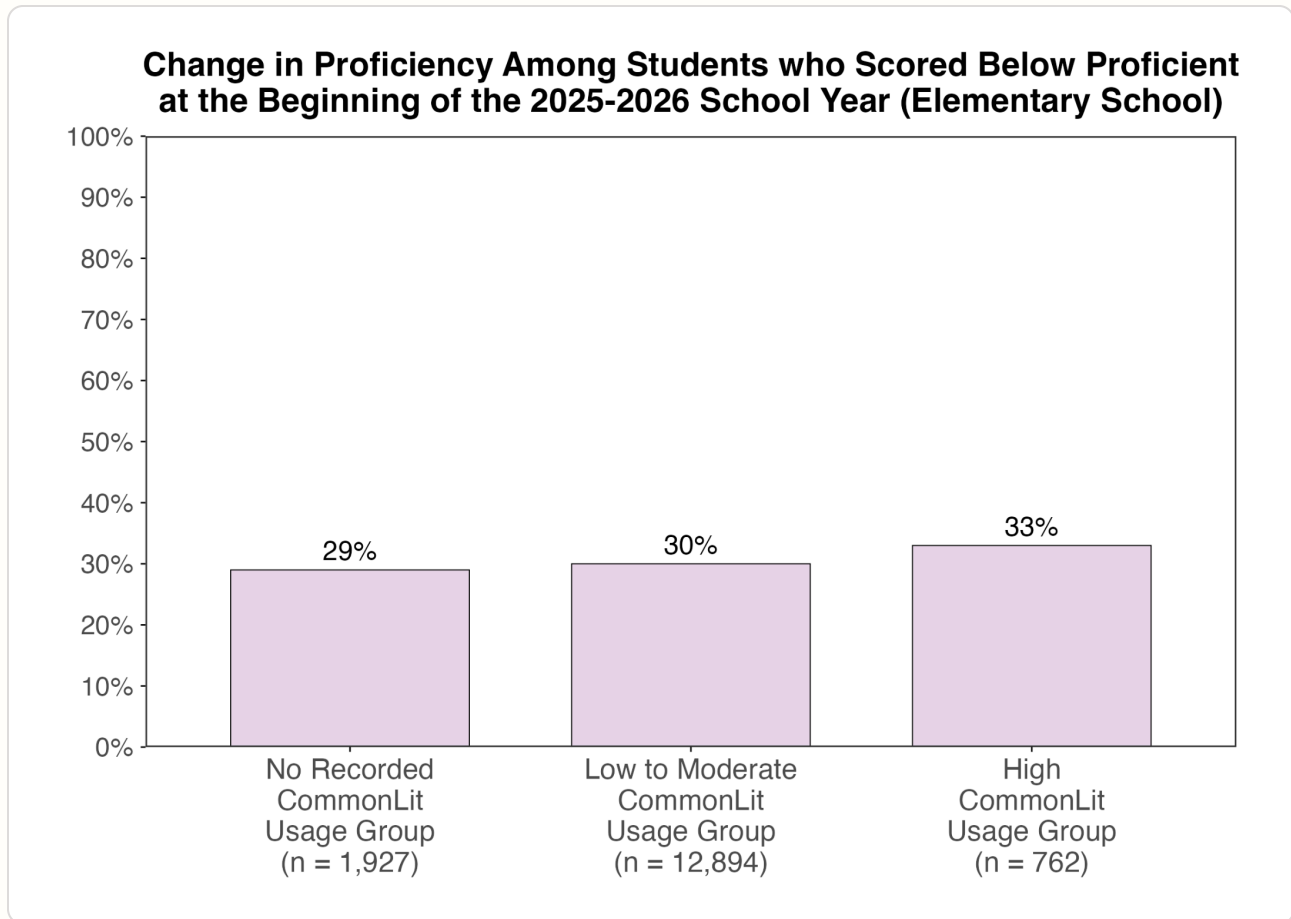


Figure A3. Percent of below proficient elementary school students who scored at the proficient level on the Post-Assessment within each of the usage groups.

Was CommonLit Usage Correlated with Student Reading Growth?

Next, we examined whether students' reading growth was significantly different across the three usage groups through conducting a one-way analysis of covariance (ANCOVA). This model predicted students' scaled score growth from the categorical CommonLit usage group variable (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage), controlling for students' baseline reading comprehension growth on the CommonLit Pre-Assessment, students' grade level, and schools' Title I eligibility. The outcome variable (reading growth) and students' baseline

scores on the CommonLit Pre-Assessment were z-scored by grade level prior to analysis, in order to place students on a common scale of relative performance across grades.

Indeed, results revealed that students' reading growth significantly differed across usage groups, $F(2, 22257) = 12.53, p < .001$. Based on follow-up analyses, differences between each of the usage groups were statistically significant. Elementary school students in the High CommonLit Usage group had higher reading growth than those in both the Low to Moderate CommonLit Usage group, $t(22257) = 4.71, p < .001$, and the No Recorded CommonLit Usage group, $t(22257) = 6.22, p < .001$. Elementary school students in the Low to Moderate CommonLit usage group also had higher reading growth than those in the No Recorded CommonLit Usage group, $t(22257) = 3.83, p < .001$. Therefore, CommonLit usage was positively correlated with elementary school students' reading growth. Figure A4 shows the model-adjusted standardized reading growth averages for each of the groups among elementary school students.

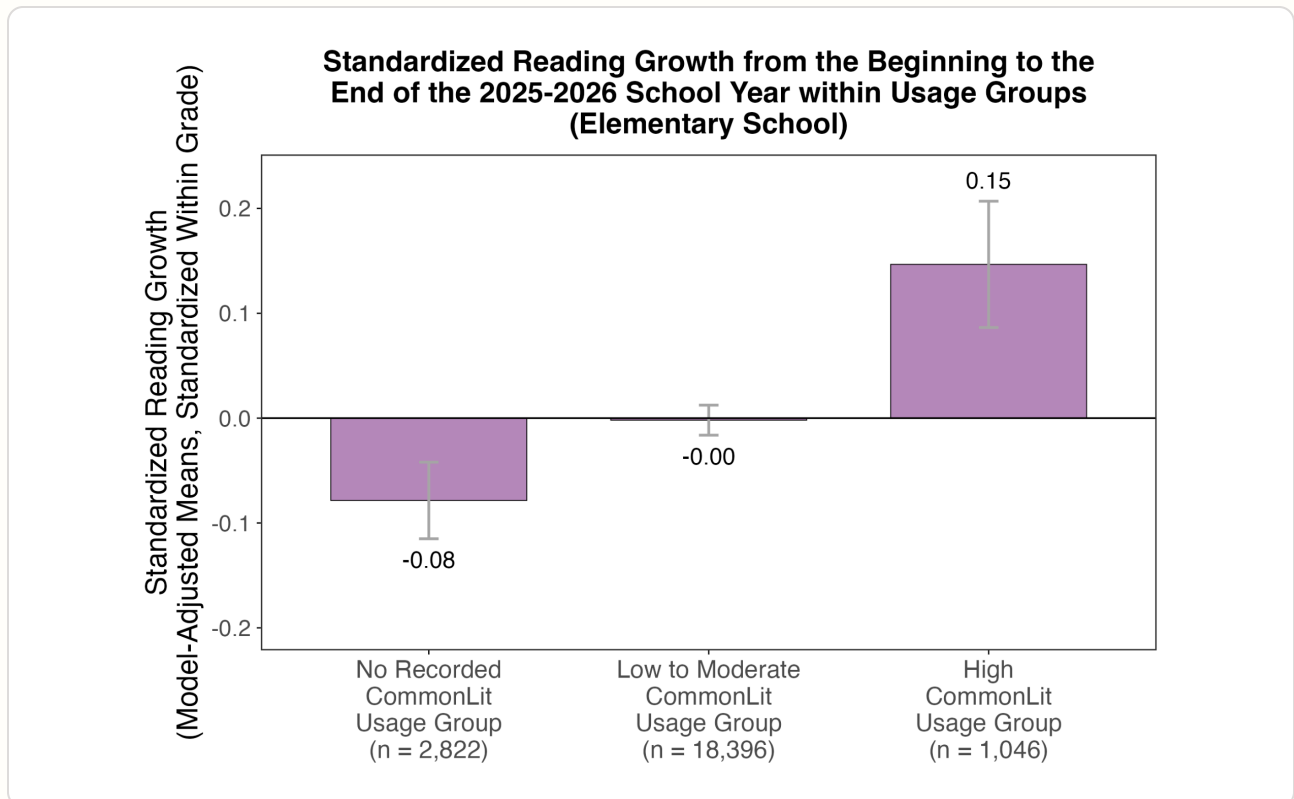


Figure A4. Model-adjusted standardized reading growth averages for each of the three usage groups (No Recorded CommonLit Usage, Low to Moderate CommonLit Usage, and High CommonLit Usage) among elementary school students. Differences between groups were statistically significant.

These findings can be translated into percentile point gains. Among elementary school students, we calculated the following weighted effect sizes: $d = .43$ for the No Recorded CommonLit Usage group, $d = .46$ for the Low to Moderate CommonLit Usage group, and $d = .54$ for the High CommonLit Usage group. There was an effect size difference of .11 standard deviations between students in the High CommonLit Usage group and the No Recorded Usage group (see Figure A5). For an elementary school student who would score at the median without using CommonLit, using CommonLit to a higher degree would raise their score by 4 percentile points – from the 50th to the 54th percentile. Additionally, there was an effect size difference of .03 standard deviations between students in the Low to Moderate CommonLit Usage group and the No Usage group. For an elementary school student who would score at the median without using CommonLit, using CommonLit to a moderate degree would raise their score by 1 percentile point – from the 50th to the 51st percentile. Thus, the reading growth observed among this particular sample of elementary school students can be translated into 1-4 percentile point gains, when compared to those who would score at the median (50th percentile) without using CommonLit.

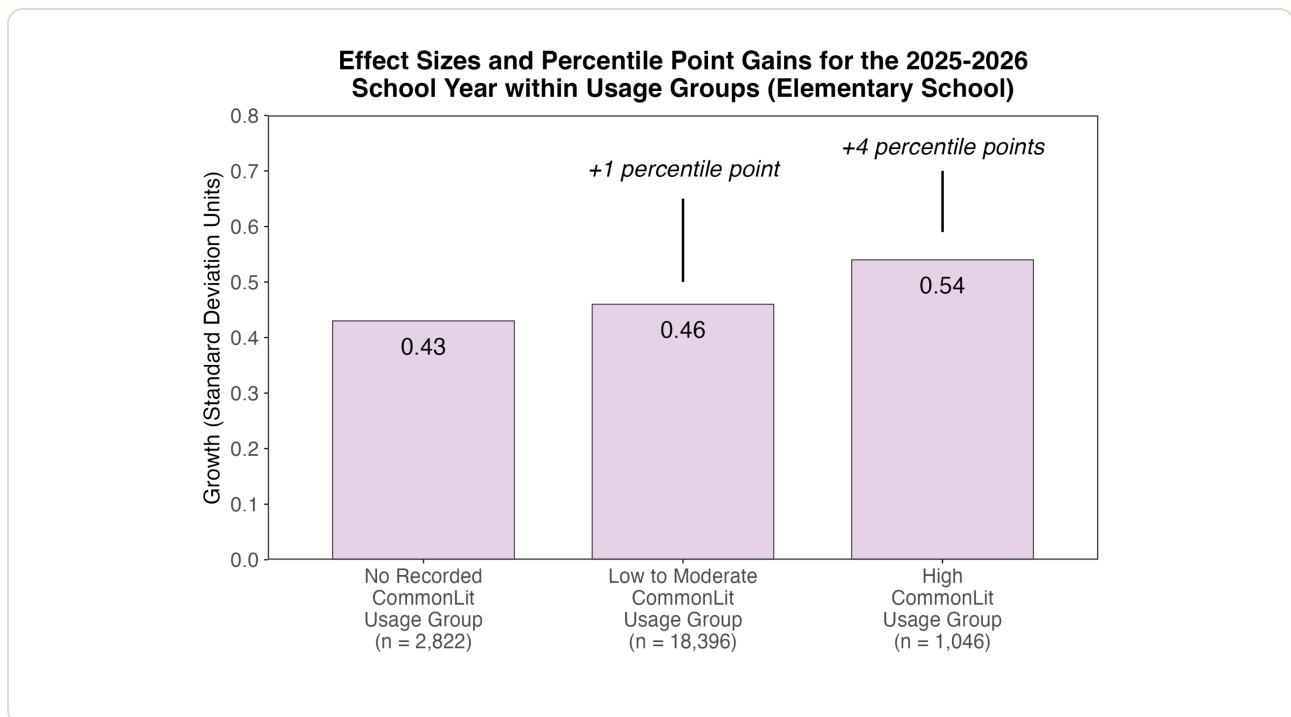


Figure A5. Effect sizes and percentile point gains for each usage group in the elementary school grade band.

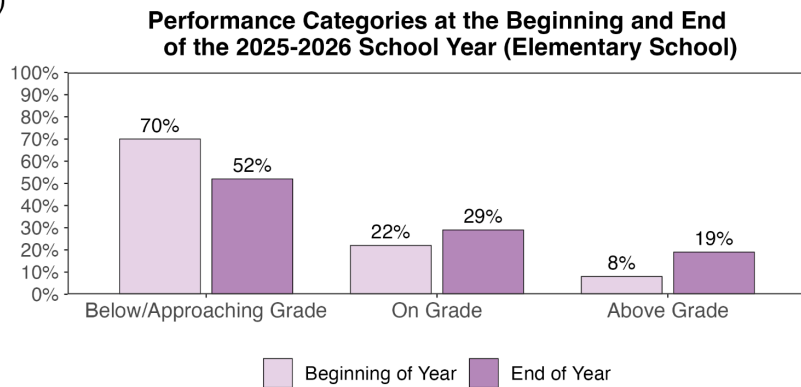
We also conducted hierarchical models to account for the nested structure of the data. Moreover, we tested whether the relationship between CommonLit usage and reading growth was dependent on schools' Title I eligibility. These results are reported in Appendix D, alongside similar models conducted for the middle school and high school grade bands. Results were consistent with the ANCOVA, providing support for the robustness of the relationship between CommonLit usage and reading growth.

Appendix C: Additional Descriptives

Performance Categories (Elementary School)

On the CommonLit Pre-Assessment, 70% of elementary school students scored "below or approaching grade level," 22% scored "on grade level," and 8% scored "above grade level." On the CommonLit Post-Assessment, 52% scored "below or approaching grade level," 29% scored "on grade level," and 19% scored "above grade level" (see Figure A6).

(i)



(ii)

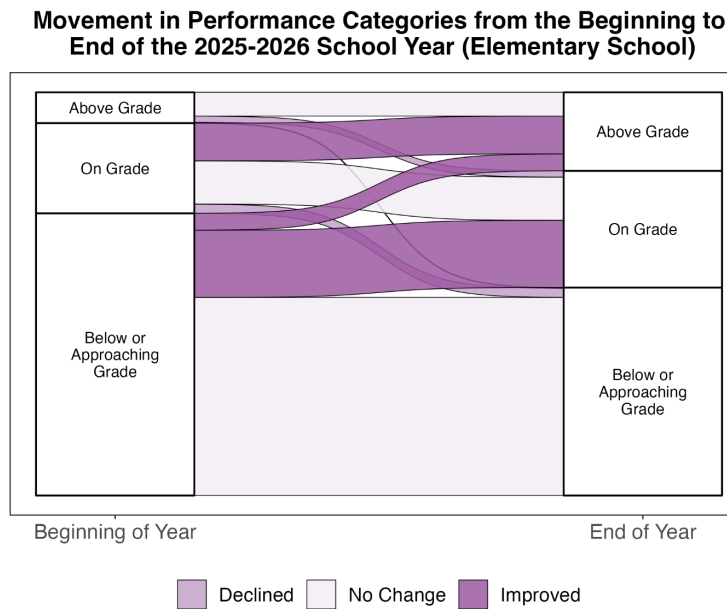


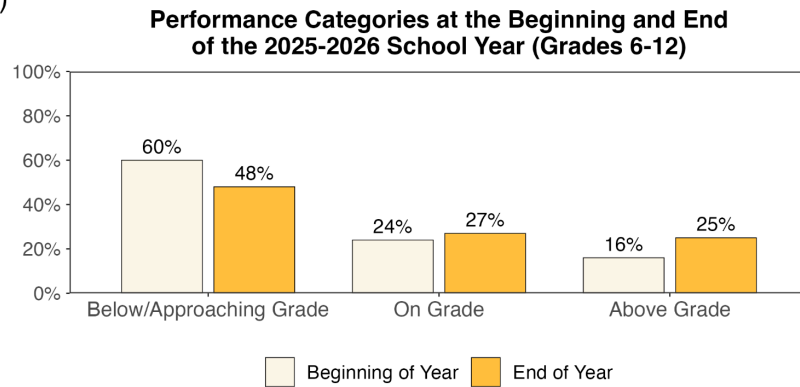
Figure A6. Percent of elementary school students in each performance category at the beginning and the end of the school year, based on scores from the Assessment Series. (i) Overall distribution by performance category at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Performance Categories (Grades 6-12)

On the CommonLit Pre-Assessment, 60% of students in Grades 6-12 scored “below or approaching grade level,” 24% scored “on grade level,” and 16% scored “above grade level.” On the CommonLit Post-Assessment, 48% of students in Grades 6-12 scored “below

or approaching grade level," 27% scored "on grade level," and 25% scored "above grade level" (see Figure A7).

(i)



(ii)

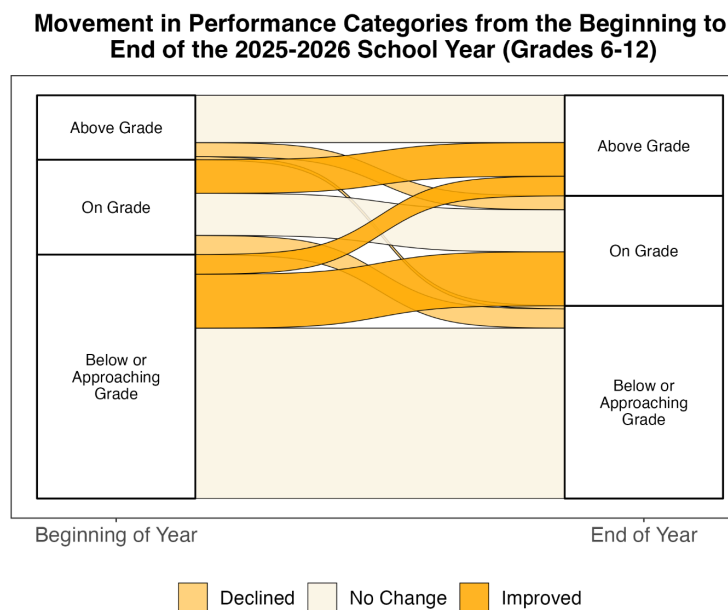


Figure A7. Percent of students in grades 6-12 (middle and high school) in each performance category at the beginning and the end of the school year, based on scores from the Assessment Series. (i) Overall distribution by performance category at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Performance Categories (Middle School)

On the CommonLit Pre-Assessment, 58% of middle school students scored “below or approaching grade level,” 25% scored “on grade level,” and 17% scored “above grade level.” On the CommonLit Post-Assessment, 48% of middle school students scored “below or approaching grade level,” 27% scored “on grade level,” and 25% scored “above grade level” (see Figure A8).

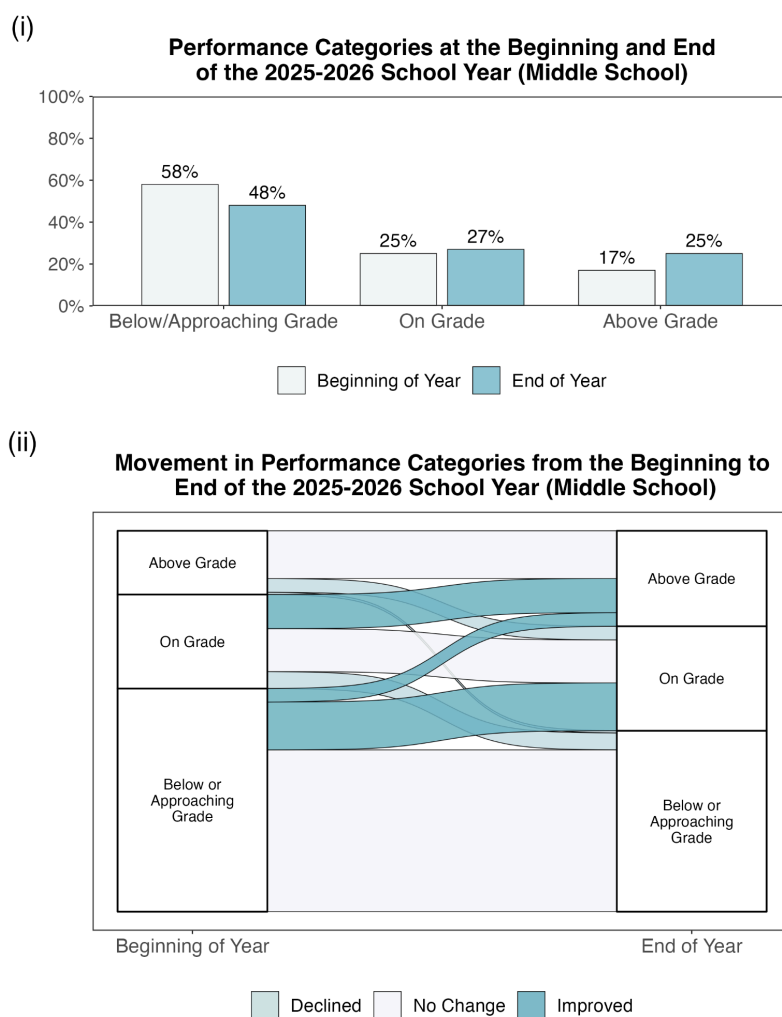


Figure A8. Percent of middle school students in each performance category at the beginning and the end of the school year, based on scores from the Assessment Series. (i) Overall distribution by performance category at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Performance Categories (High School)

On the CommonLit Pre-Assessment, 62% of high school students scored “below or approaching grade level,” 23% scored “on grade level,” and 15% scored “above grade level.” On the CommonLit Post-Assessment, 48% of high school students scored “below or approaching grade level,” 27% scored “on grade level,” and 25% scored “above grade level” (see Figure A9).

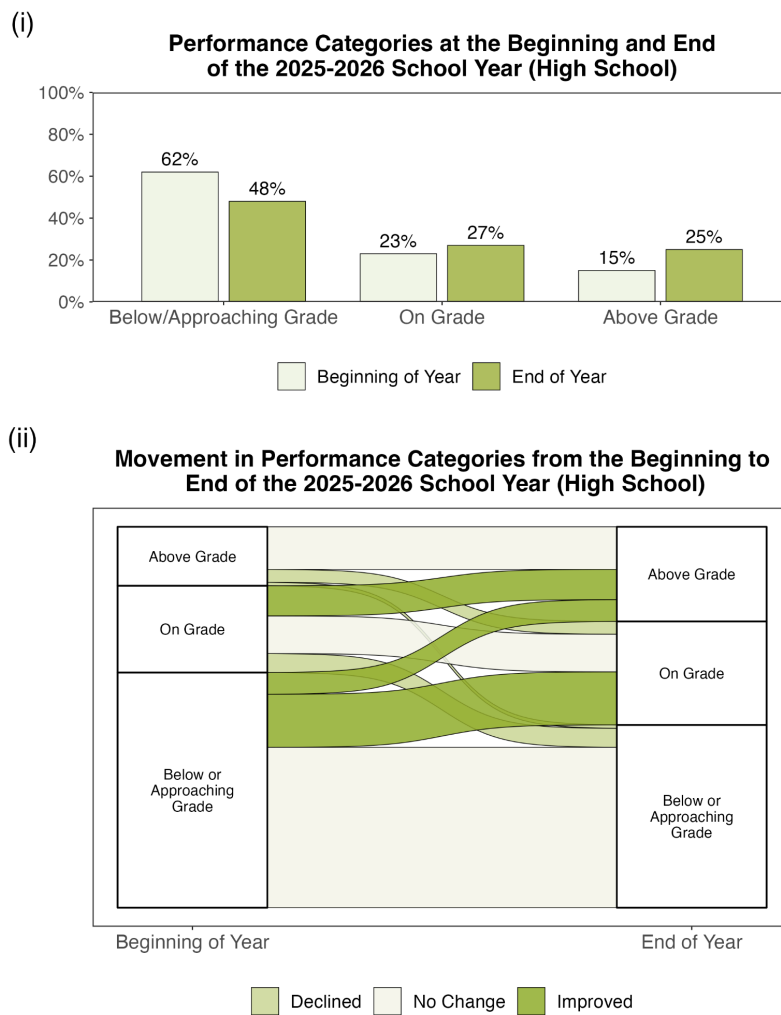


Figure A9. Percent of high school students in each performance category at the beginning and the end of the school year, based on scores from the Assessment Series. (i) Overall distribution by performance category at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Proficiency (Elementary School)

Next, we examined the number of students who scored at the proficient level (“on grade level” or “above grade level”). On the CommonLit Pre-Assessment, 30% of elementary school students (6,681 students) scored at the proficient level (“on grade level” or “above grade level”; see Figure A10). At the end of the year, 48% of elementary school students (10,781 students) scored at the proficient level on the CommonLit Post-Assessment. That is, an additional 4,100 students scored at the proficient level from the beginning to the end of the school year, representing a 61% increase in proficiency.

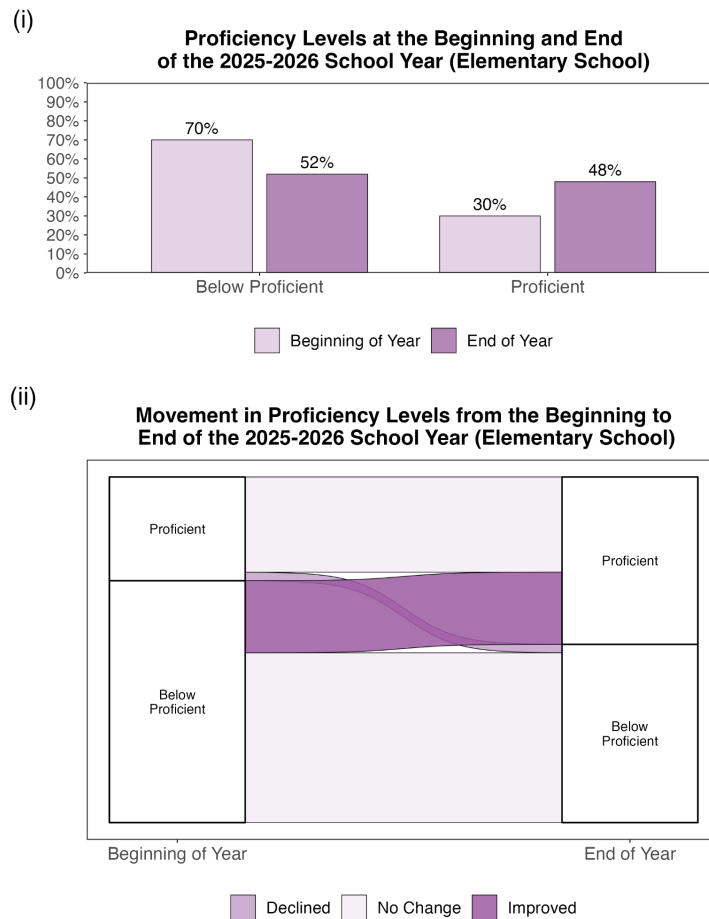


Figure A10. Percent of elementary school students scoring at the below proficient and proficient levels from the beginning to the end of the year. (i) Overall distribution by proficiency level at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Proficiency (Grades 6-12)

On the CommonLit Pre-Assessment, 40% of students in Grades 6-12 (118,449 students) scored at the proficient level. On the CommonLit Post-Assessment, 52% of students in Grades 6-12 (156,494 students) scored at the proficient level (see Figure A11). In other words, an additional 38,045 students in Grades 6-12 scored at the proficient level from the beginning to the end of the school year, which was a 32% increase in the number of students in Grades 6-12 who scored at the proficient level from the beginning to the end of the year.

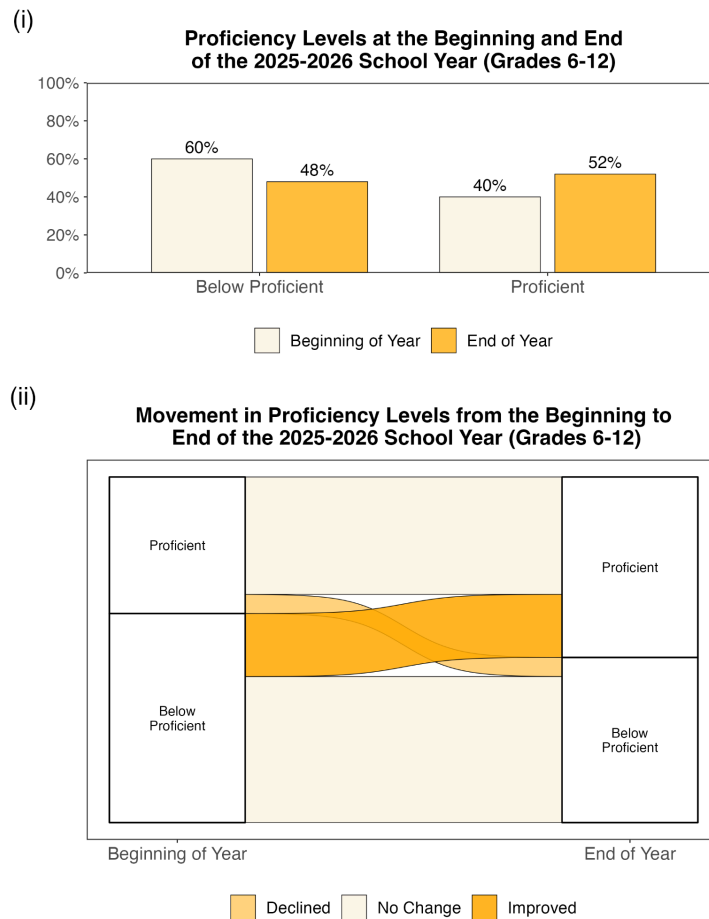


Figure A11. Percent of students in grades 6-12 (middle and high school) scoring at the below proficient and proficient levels from the beginning to the end of the year. (i) Overall distribution by proficiency level at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Proficiency (Middle School)

On the CommonLit Pre-Assessment, 41% of middle school students (49,395 students) scored at the proficient level. On the CommonLit Post-Assessment, 52% of middle school students (62,615 students) scored at the proficient level (see Figure A12). In other words, an additional 13,220 middle school students scored at the proficient level from the beginning to the end of the school year, which was a 27% increase in the number of middle school students who scored at the proficient level from the beginning to the end of the year.

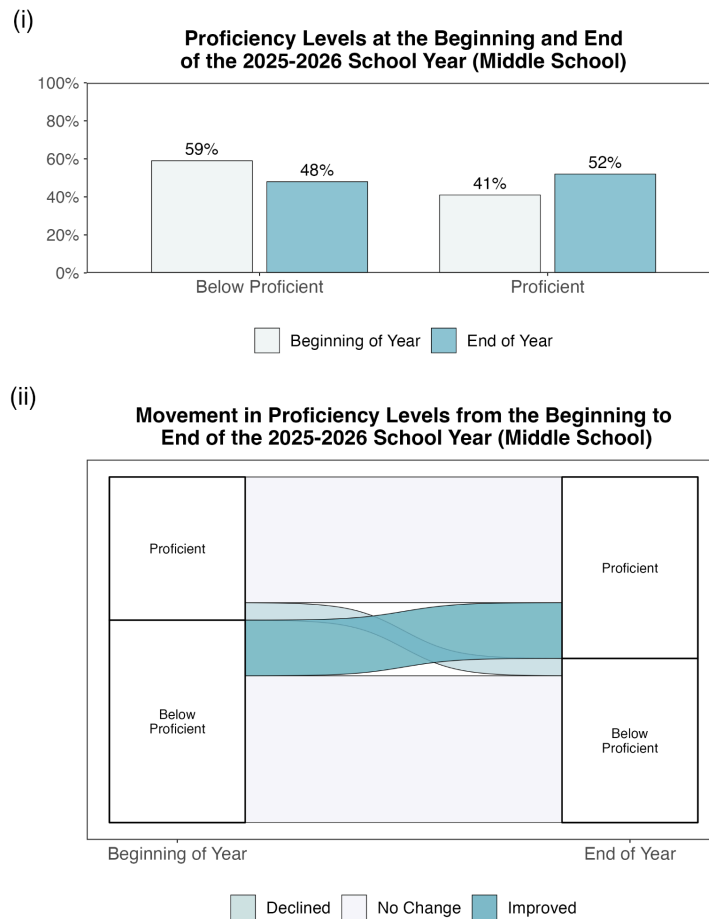


Figure A12. Percent of middle school students scoring at the below proficient and proficient levels from the beginning to the end of the year. (i) Overall distribution by proficiency level at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Proficiency (High School)

On the CommonLit Pre-Assessment, 38% of high school students (69,054 students) scored at the proficient level. On the CommonLit Post-Assessment, 52% of high school students (93,879 students) scored at the proficient level (see Figure A13). In other words, an additional 24,825 high school students scored at the proficient level from the beginning to the end of the school year, which translated into a 36% increase in the number of high school students who scored at the proficient level from the beginning to the end of the year.

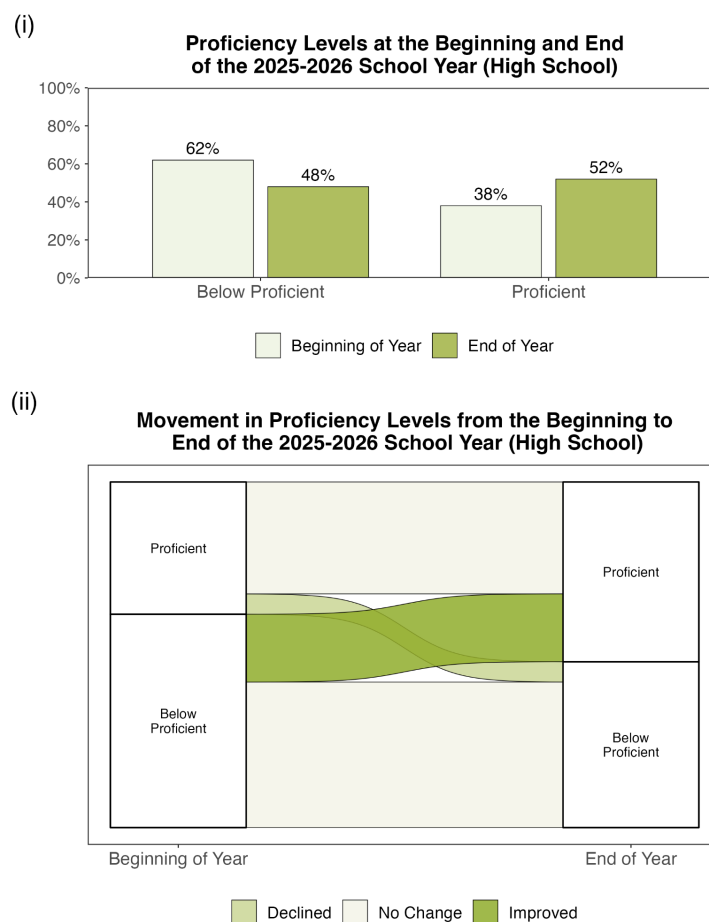


Figure A13. Percent of high school students scoring at the below proficient and proficient levels from the beginning to the end of the year. (i) Overall distribution by proficiency level at each time point. (ii) Individual student movement between categories, colored by whether performance declined, held steady, or improved.

Percent Increase in Proficiency Within Usage Groups (Elementary School)

Among elementary school students in the High CommonLit Usage group, there was a 78% increase in the number of students who scored at the proficient level on the Pre-Assessment to the Post-Assessment (284 students to 505 students; see Figure A14).

Among elementary school students in the Low to Moderate CommonLit Usage Group, there was a 62% increase (5,502 students to 8,892 students). Finally, among elementary school students in the No Recorded Usage group, there was a 55% increase (895 students to 1,384 students).

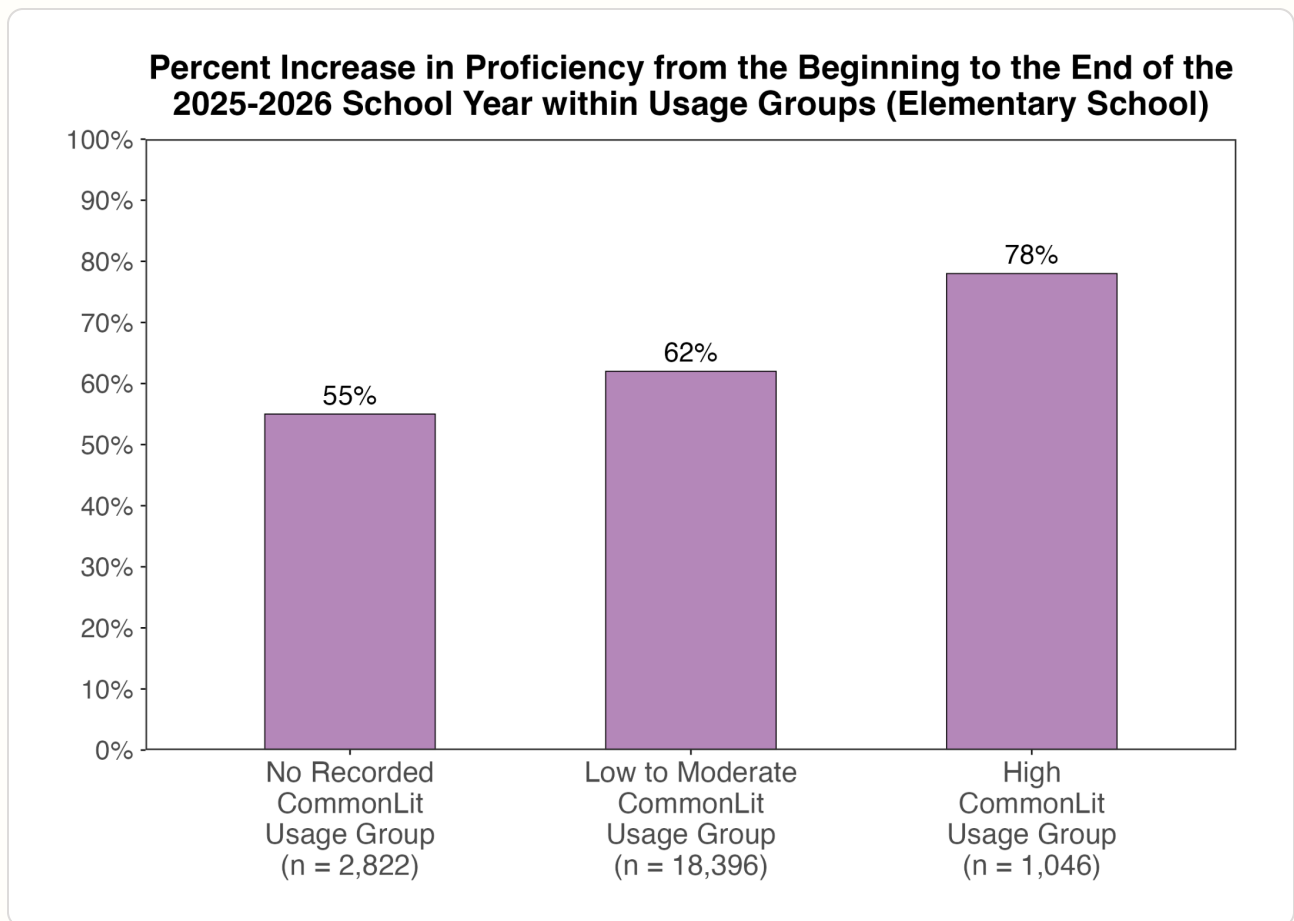


Figure A14. Percent increase in the number of proficient elementary school students from the beginning to the end of the year within each of the usage groups.

Percent Increase in Proficiency Within Usage Groups (Grades 6-12)

We also examined the percent increase in the raw number of students who scored at the proficient level from the beginning to the end of the year within each of the usage groups.

To calculate this, we took the difference between the number of students who scored at the proficient level at the beginning and end of the year, and divided it by the number of students who scored at the proficient level at the beginning of the year. Among students in Grades 6-12 in the High CommonLit Usage group, there was a 41% increase in the number of students who scored at the proficient level on the Pre-Assessment to the Post-Assessment (12,729 students to 17,977 students; see Figure A15). Among students in Grades 6-12 in the Low to Moderate CommonLit Usage Group, there was a 33% increase (94,701 students to 125,902 students). Finally, among students in Grades 6-12 in the No Recorded CommonLit Usage group, there was a 14% increase (11,019 students to 12,615 students).

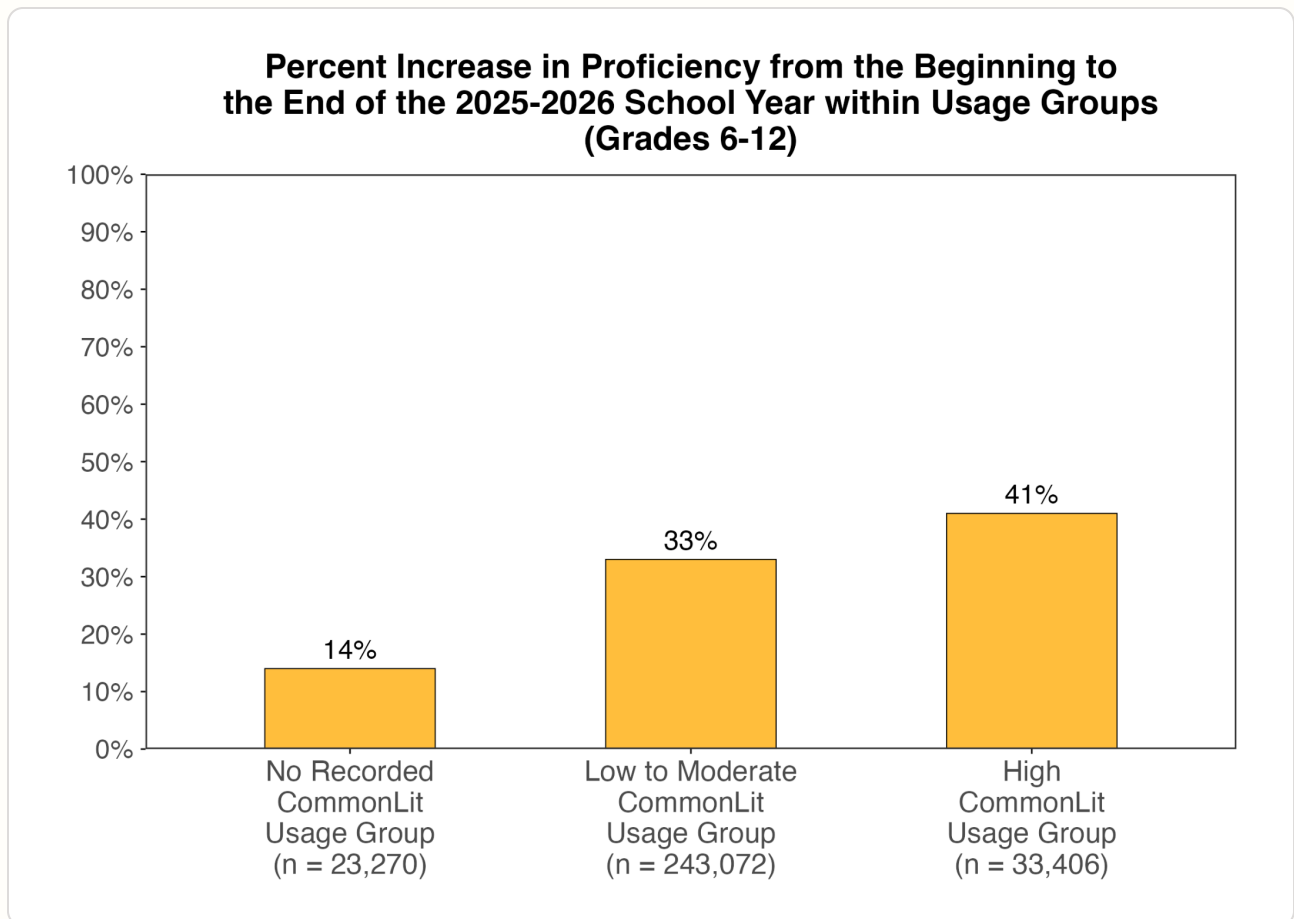


Figure A15. Percent increase in the number of proficient students in grades 6-12 (middle and high school) from the beginning to the end of the year within each of the usage groups.

Percent Increase in Proficiency Within Usage Groups (Middle School)

Among middle school students in the High CommonLit Usage group, there was a 33% increase in the number of students who scored at the proficient level on the Pre-Assessment to the Post-Assessment (9,757 students to 12,972 students; see Figure A16). Among middle school students in the Low to Moderate CommonLit Usage Group, there was a 26% increase (37,427 students to 47,133 students). Finally, among middle school students in the No Recorded CommonLit Usage group, there was a 14% increase (2,211 students to 2,510 students).

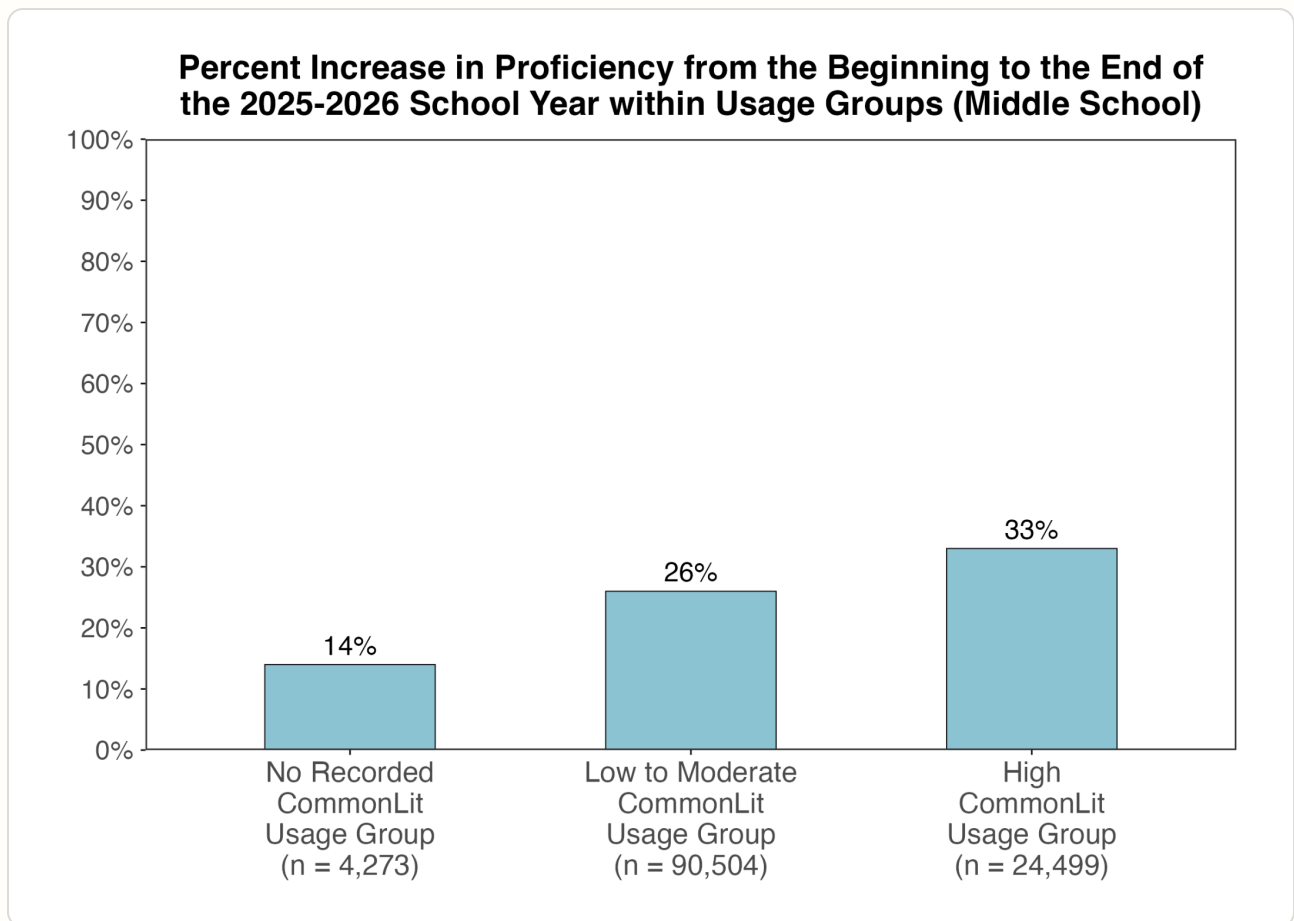


Figure A16. Percent increase in the number of proficient middle school students from the beginning to the end of the year within each of the usage groups.

Percent Increase in Proficiency Within Usage Groups (High School)

Among high school students in the High CommonLit Usage group, there was a 68% increase in the number of students who scored at the proficient level on the Pre-

Assessment to the Post-Assessment (2,972 students to 5,005 students; see Figure A17). Among high school students in the Low to Moderate CommonLit Usage Group, there was a 38% increase (57,274 students to 78,769 students). Finally, among high school students in the No Recorded CommonLit Usage group, there was a 15% increase (8,808 students to 10,105 students).

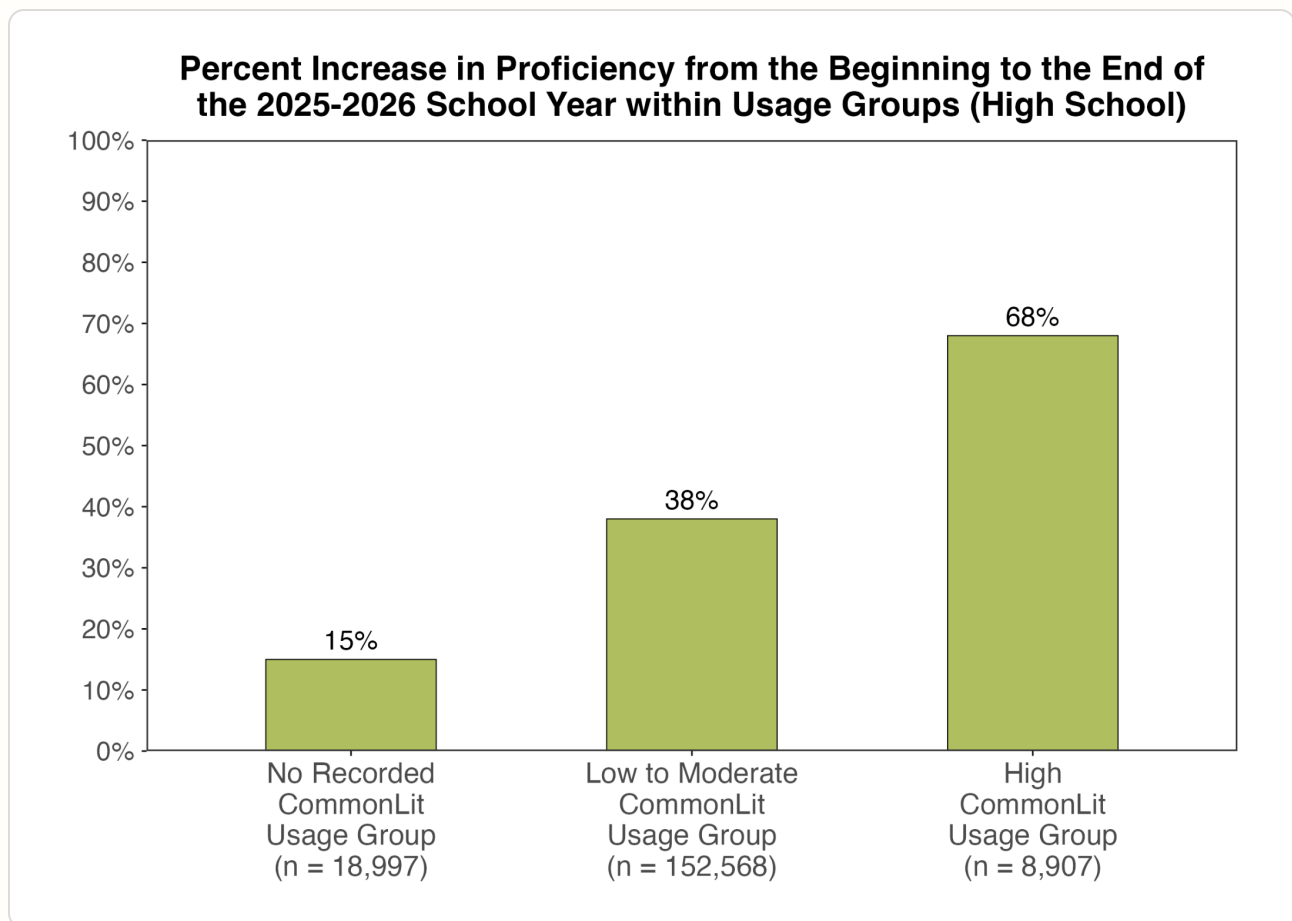


Figure A17. Percent increase in the number of proficient high school students from the beginning to the end of the year within each of the usage groups.

Appendix D: Hierarchical Models

Hierarchical models were conducted to address the nested structure of the data – i.e., students nested within classrooms nested within schools – and to test the robustness of the relationship between CommonLit usage and reading growth. For each grade band, four hierarchical models were conducted. First, we ran a model with the categorical CommonLit usage group variable (no recorded usage, low to moderate usage, high usage)

to test whether results from the ANCOVA held when accounting for the nested structure of the data. Second, we ran a model with a continuous CommonLit usage variable to test whether results held when not using a researcher-created usage group variable. Third, we ran a model that tests for moderation by Title I eligibility to examine whether the relationship between CommonLit usage and reading growth held for students at Title I eligible and other schools. Fourth, we ran a non-linear model using a natural spline transformation of the continuous usage variable to test whether the relationship between usage and growth was better characterized as non-linear (e.g., with diminishing or accelerating returns at different levels of usage) rather than strictly linear. Across all hierarchical models, the following covariates were included: scores from the CommonLit Pre-Assessment (standardized to a mean of 0 and standard deviation of 1), Title I eligibility (1 = Title I eligible), grade level, classroom-level CommonLit usage, and classroom-level scores from the CommonLit Pre-Assessment. Additionally, teachers and schools were modeled as random intercepts. Analyses were conducted separately for four grade bands: Elementary (grades 3-5), a combined Secondary grade band (grades 6-12), Middle School (grades 6-8), and High School (grades 9-12).

Across all grade bands, results across all hierarchical models showed that higher CommonLit usage significantly predicted higher reading growth, with no evidence of moderation by Title I eligibility (see Tables A1-A12). For models that included the categorical CommonLit usage variable, results showed that the High CommonLit Usage group saw significantly higher reading growth than the No Recorded Usage group, Elementary: $b = 0.23$, $SE = 0.06$, $t = 3.96$, $p < .001$; Secondary grade band: $b = 0.17$, $SE = 0.02$, $t = 10.53$, $p < .001$; Middle: $b = 0.23$, $SE = 0.03$, $t = 6.72$, $p < .001$; High: $b = 0.11$, $SE = 0.02$, $t = 4.62$, $p < .001$. Additionally, the Low to Moderate CommonLit Usage group saw significantly higher reading growth than the No Recorded Usage group, Elementary: $b = 0.09$, $SE = 0.03$, $t = 2.68$, $p = .008$; Secondary grade band: $b = 0.05$, $SE = 0.01$, $t = 3.96$, $p < .001$; Middle: $b = 0.10$, $SE = 0.03$, $t = 3.22$, $p = .001$; High: $b = 0.03$, $SE = 0.01$, $t = 2.37$, $p = .018$.

For models that included the continuous CommonLit usage variable, results showed that CommonLit usage significantly predicted reading growth, Elementary: $b = 0.02$, $SE = 0.003$, $t = 7.51$, $p < .001$; Secondary: $b = 0.01$, $SE = 0.001$, $t = 25.44$, $p < .001$; Middle: $b = 0.02$, $SE = 0.001$, $t = 24.31$, $p < .001$; High: $b = 0.01$, $SE = 0.001$, $t = 9.98$, $p < .001$. Finally, for models that tested for moderation, results showed that the relationship between CommonLit

usage and reading growth was not significantly moderated by Title I eligibility, Elementary: $b = -0.002$, $SE = 0.002$, $t = -0.84$, $p = .401$; Secondary: $b = -.0005$, $SE = 0.0004$, $t = -1.44$, $p = .151$; Middle: $b = -0.001$, $SE = 0.001$, $t = -1.57$, $p = .118$; High: $b = -0.0002$, $SE = 0.001$, $t = -0.34$, $p = .735$.

Given that the continuous usage variable showed a significant, positive linear relationship with reading growth, we also tested whether this relationship was better described as non-linear. This is important because a strictly linear model assumes that each additional lesson accessed is associated with the same increment in reading growth, regardless of a student's overall usage level; a non-linear model allows this benefit to vary and tests whether the returns accelerate, plateau, or diminish at higher levels of usage. To test this, we fitted a spline-based model that allowed the relationship between usage and growth to flex across three degrees of freedom, and compared its fit to the linear continuous model using a likelihood ratio test. Across all grade bands, the non-linear model fit the data significantly better than the linear model, Elementary: $\chi^2(2) = 16.65$, $p < .001$; Secondary: $\chi^2(2) = 64.30$, $p < .001$; Middle: $\chi^2(2) = 23.53$, $p < .001$; High: $\chi^2(2) = 40.88$, $p < .001$, indicating that the relationship between CommonLit usage and reading growth is not purely linear.

To characterize this non-linear relationship in practical terms, we generated predicted reading growth across the full range of CommonLit usage in each grade band using the spline models, and compared these predictions to a standard linear model (see Figure A18). Across all four grade bands, predicted growth increased with usage, but the rate of improvement was not constant: growth increased most steeply at lower levels of usage and more gradually at the highest levels of usage. This pattern held across all four grade bands (Elementary, Secondary, Middle School, and High School), though the size of the pattern varied somewhat by grade band (see Tables A4, A8, A12, and A16 for benchmark predictions at the 10th, 25th, 50th, 75th, and 90th percentiles of usage within the Elementary, Secondary, Middle, and High grade bands).

In sum, there was a positive relationship between CommonLit usage and reading growth across models and grade bands, and this relationship held for students at Title I eligible schools and other schools. Additionally, this relationship was non-linear, with growth increasing more gradually at the highest levels of usage across all four grade bands.

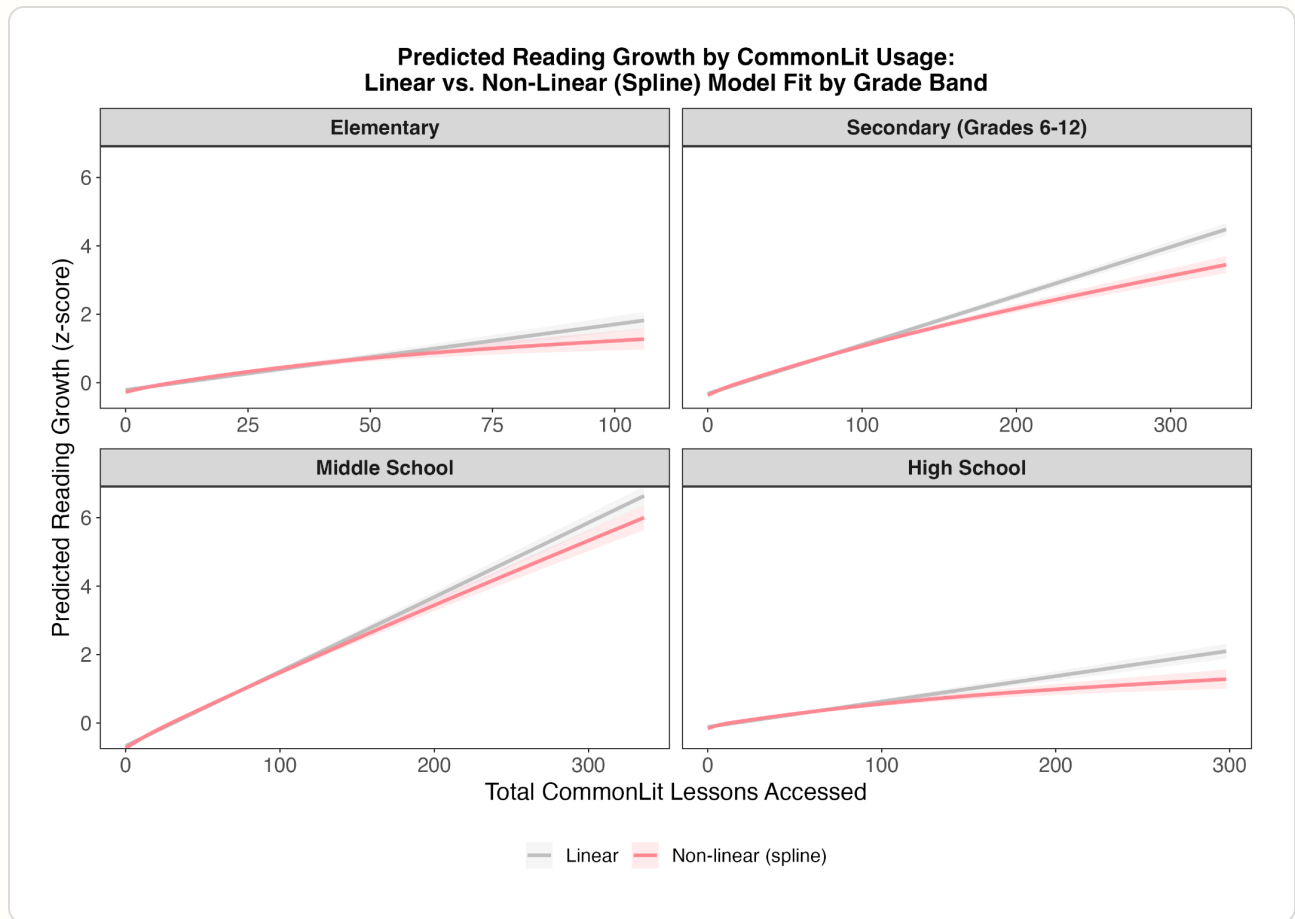


Figure A18. Predicted reading growth by total CommonLit lessons accessed, comparing linear and non-linear (spline, $df = 3$) model fits across four grade bands. Shaded bands represent ± 1 standard error. Across all four grade bands, the spline model shows growth increasing more gradually at higher usage levels, diverging below the linear model's straight-line trend — a pattern consistent with diminishing returns to additional CommonLit usage. Note that the x-axis range differs across grade bands, reflecting differences in the underlying distribution of total lessons accessed within each group.

Elementary School

Table A1. Hierarchical linear model predicting reading growth from CommonLit usage groups among elementary school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.07	0.27
School (Intercept)	0.03	0.16
Residual	0.87	0.93

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.04	0.04	431.54	-0.83	.408
Low to Moderate CommonLit Usage Group	0.09	0.03	1421.55	2.68	.008
High CommonLit Usage Group	0.23	0.06	1985.38	3.96	< .001
Pre-Assessment Scores	-0.22	0.01	21278.85	-31.10	< .001
Title I Eligibility	-0.14	0.04	132.14	-3.58	< .001
Grade 3	0.03	0.03	905.90	1.15	.252
Grade 4	0.04	0.03	905.47	1.59	.112
Classroom Effect: Avg. Pre-Assessment Scores	0.14	0.03	914.36	4.61	< .001

Table A2. Hierarchical linear model predicting reading growth from CommonLit usage (continuous) among elementary school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.07	0.27
School (Intercept)	0.02	0.16
Residual	0.86	0.93

Predictor	Estimate	SE	df	t-value	p-value
Intercept	0.05	0.03	168.04	1.47	.144
Total CommonLit Lessons	0.02	0.003	21262.12	7.51	< .001
Pre-Assessment Scores	-0.23	0.01	21259.66	-31.69	< .001
Title I Eligibility	-0.15	0.04	130.76	-3.83	< .001
Grade 3	0.04	0.03	911.23	1.40	.162
Grade 4	0.04	0.03	906.82	1.52	.130
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.003	16893.35	-5.28	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.15	0.03	900.12	4.91	< .001

Table A3. Hierarchical linear model testing for moderation by Title I eligibility among elementary school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.07	0.27
School (Intercept)	0.02	0.16
Residual	0.86	0.93

Predictor	Estimate	SE	df	t-value	p-value
Intercept	0.05	0.03	169.58	1.51	.133
Total CommonLit Lessons	0.02	0.003	8550.43	6.69	< .001
Pre-Assessment Scores	-0.23	0.01	21265.64	-31.70	< .001
Title I Eligibility	-0.15	0.04	127.54	-3.75	< .001
Grade 3	0.04	0.03	910.10	1.36	.173
Grade 4	0.04	0.03	906.35	1.48	.139
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.003	14818.35	-5.35	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.16	0.03	903.84	4.97	< .001
CommonLit Lessons x Title I Eligibility	-0.002	0.002	1067.98	-0.84	.401

Table A4. Non-linear (spline) hierarchical model predicting reading growth from CommonLit usage among elementary school students, with predicted growth at usage benchmarks.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.07	0.27
School (Intercept)	0.02	0.15
Residual	0.86	0.93

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.22	0.05	804.71	-4.68	< .001
Total CommonLit Lessons (Spline Term 1)	0.94	0.12	11037.82	7.81	< .001
Total CommonLit Lessons (Spline Term 2)	1.42	0.22	19013.11	6.47	< .001
Total CommonLit Lessons (Spline Term 3)	1.48	0.32	12390.28	4.63	< .001
Pre-Assessment Scores	-0.23	0.01	21284.55	-31.80	< .001
Title I Eligibility	-0.16	0.04	128.72	-4.04	< .001
Grade 3	0.05	0.03	906.01	1.82	.069
Grade 4	0.04	0.03	897.48	1.53	.125
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.003	16283.76	-5.38	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.16	0.03	871.40	5.06	< .001

Percentile	Total Lessons Accessed	Predicted Growth (z-score)
10%	0	-0.27
25%	2	-0.20
50%	7	-0.06
75%	15	0.12
90%	28	0.38

Grades 6-12

Table A5. Hierarchical linear model predicting reading growth from CommonLit usage groups among students in grades 6-12.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.05	0.22
Residual	0.78	0.88

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.02	0.02	2692.17	-1.04	.299
Low to Moderate CommonLit Usage Group	0.05	0.01	18819.94	3.96	< .001
High CommonLit Usage Group	0.17	0.02	23543.46	10.53	< .001
Pre-Assessment Scores	-0.40	0.002	294844.39	-210.62	< .001
Title I Eligibility	0.001	0.02	849.41	0.05	.963
Grade 6	-0.11	0.02	3871.28	-6.32	< .001
Grade 7	-0.08	0.02	4269.49	-5.16	< .001
Grade 8	-0.09	0.02	4586.42	-5.82	< .001
Grade 10	0.002	0.01	38962.68	0.24	.809
Grade 11	0.01	0.01	41054.23	1.15	.249
Grade 12	-0.02	0.01	54699.79	-1.39	.164
Classroom Effect: Avg. Pre-Assessment Scores	0.26	0.01	6740.28	29.10	< .001

Table A6. Hierarchical linear model predicting reading growth from CommonLit usage (continuous) among students in grades 6-12.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.04	0.21
Residual	0.78	0.88

Predictor	Estimate	SE	df	t-value	p-value
Intercept	0.02	0.01	1143.78	1.59	.111
Total CommonLit Lessons	0.01	0.0006	296259.22	25.44	< .001
Pre-Assessment Scores	-0.40	0.002	294709.19	-211.67	< .001
Title I Eligibility	0.001	0.02	853.08	0.07	.942
Grade 6	-0.09	0.02	3933.33	-5.26	< .001
Grade 7	-0.06	0.02	4321.60	-4.02	< .001
Grade 8	-0.07	0.02	4606.86	-4.68	< .001
Grade 10	0.01	0.01	38944.10	0.76	.445
Grade 11	0.02	0.01	40878.10	2.23	.026
Grade 12	0.01	0.01	54075.89	0.66	.512
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	165116.10	-20.84	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.27	0.01	6753.37	29.58	< .001

Table A7. Hierarchical linear model testing for moderation by Title I eligibility among students in grades 6-12.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.04	0.21
Residual	0.78	0.88

Predictor	Estimate	SE	df	t-value	p-value
Intercept	0.02	0.01	1134.61	1.52	.130
Total CommonLit Lessons	0.01	0.001	152450.79	24.15	< .001
Pre-Assessment Scores	-0.40	0.002	294724.59	-211.68	< .001
Title I Eligibility	0.004	0.02	848.53	0.24	.808
Grade 6	-0.09	0.02	3940.23	-5.30	< .001
Grade 7	-0.06	0.02	4308.81	-4.01	< .001
Grade 8	-0.07	0.02	4596.40	-4.68	< .001
Grade 10	0.01	0.01	38958.79	0.77	.439
Grade 11	0.02	0.01	40891.79	2.23	.026
Grade 12	0.01	0.01	54078.48	0.67	.506
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	164306.85	-20.86	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.27	0.01	6751.21	29.58	< .001
CommonLit Lessons x Title I Eligibility	-0.001	0.0004	5007.75	-1.44	.151

Table A8. Non-linear (spline) hierarchical model predicting reading growth from CommonLit usage among secondary school students, with predicted growth at usage benchmarks.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.24
School (Intercept)	0.04	0.21
Residual	0.78	0.88

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.33	0.02	4340.26	-16.20	< .001
Total CommonLit Lessons (Spline Term 1)	1.80	0.08	109162.16	23.81	< .001
Total CommonLit Lessons (Spline Term 2)	3.03	0.15	103021.74	19.79	< .001
Total CommonLit Lessons (Spline Term 3)	3.73	0.25	33090.88	14.81	< .001
Pre-Assessment Scores	-0.40	0.002	294719.66	-211.65	< .001
Title I Eligibility	-0.001	0.02	842.43	-0.11	.912
Grade 6	-0.10	0.02	3940.65	-6.07	< .001
Grade 7	-0.07	0.02	4298.40	-4.62	< .001
Grade 8	-0.08	0.02	4566.22	-5.23	< .001
Grade 10	0.01	0.01	38568.56	0.99	.323
Grade 11	0.02	0.01	40534.90	2.54	.011
Grade 12	0.01	0.01	53401.42	1.18	.240
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	154140.72	-19.92	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.27	0.01	6759.27	29.96	< .001

Percentile	Total Lessons Accessed	Predicted Growth (z-score)
10%	1	-0.34
25%	4	-0.28
50%	12	-0.13

Percentile	Total Lessons Accessed	Predicted Growth (z-score)
75%	31	0.15
90%	58	0.53

Middle School

Table A9. Hierarchical linear model predicting reading growth from CommonLit usage groups among middle school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.23
School (Intercept)	0.04	0.20
Residual	0.84	0.92

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.12	0.04	3052.49	-3.32	< .001
Low to Moderate CommonLit Usage Group	0.10	0.03	4715.98	3.22	.001
High CommonLit Usage Group	0.23	0.03	5303.20	6.72	< .001
Pre-Assessment Scores	-0.33	0.003	117466.15	-106.75	< .001
Title I Eligibility	-0.03	0.02	383.92	-1.42	.156
Grade 6	-0.02	0.01	3590.58	-1.22	.225
Grade 7	0.01	0.01	4941.36	0.86	.389
Classroom Effect: Avg. Pre- Assessment Scores	0.28	0.01	2751.93	19.03	< .001

Table A10. Hierarchical linear model predicting reading growth from CommonLit usage (continuous) among middle school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.05	0.23
School (Intercept)	0.04	0.20
Residual	0.83	0.91

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.002	0.02	510.69	-0.11	.917
Total CommonLit Lessons	0.02	0.001	117275.21	24.31	< .001
Pre-Assessment Scores	-0.33	0.003	117295.91	-108.73	< .001
Title I Eligibility	-0.03	0.02	390.02	-1.17	.243
Grade 6	-0.01	0.01	3600.59	-1.03	.302
Grade 7	0.01	0.01	4940.12	0.75	.455
Classroom Effect: Avg. CommonLit Lessons	-0.02	0.001	84838.07	-21.85	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.29	0.01	2756.87	19.72	< .001

Table A11. Hierarchical linear model testing for moderation by Title I eligibility among middle school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.05	0.23
School (Intercept)	0.04	0.20
Residual	0.83	0.91

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.004	0.02	510.12	-0.22	.828
Total CommonLit Lessons	0.02	0.001	78016.98	23.48	< .001
Pre-Assessment Scores	-0.33	0.003	117306.25	-108.74	< .001
Title I Eligibility	-0.02	0.02	387.94	-0.97	.332
Grade 6	-0.02	0.01	3609.55	-1.09	.277
Grade 7	0.01	0.01	4957.41	0.77	.444
Classroom Effect: Avg. CommonLit Lessons	-0.02	0.001	83679.62	-21.90	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.29	0.01	2752.11	19.72	< .001
CommonLit Lessons x Title I Eligibility	-0.001	0.001	1976.01	-1.57	.118

Table A12. Non-linear (spline) hierarchical model predicting reading growth from CommonLit usage among middle school students, with predicted growth at usage benchmarks.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.05	0.23
School (Intercept)	0.04	0.20
Residual	0.83	0.91

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.71	0.04	5120.53	-19.59	< .001
Total CommonLit Lessons (Spline Term 1)	2.68	0.12	78345.20	22.91	< .001
Total CommonLit Lessons (Spline Term 2)	5.24	0.25	53585.22	21.07	< .001
Total CommonLit Lessons (Spline Term 3)	6.54	0.37	19255.77	17.57	< .001
Pre-Assessment Scores	-0.33	0.003	117307.12	-108.75	< .001
Title I Eligibility	-0.03	0.02	392.31	-1.35	.179
Grade 6	-0.02	0.01	3616.64	-1.37	.170
Grade 7	0.01	0.01	4918.60	0.65	.517
Classroom Effect: Avg. CommonLit Lessons	-0.02	0.001	76626.44	-21.42	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.29	0.01	2761.99	19.68	< .001

Percentile	Total Lessons Accessed	Predicted Growth (z-score)
10%	3	-0.65
25%	8	-0.51
50%	22	-0.17
75%	45	0.33
90%	68	0.81

High School

Table A13. Hierarchical linear model predicting reading growth from CommonLit usage groups among high school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.05	0.22
Residual	0.74	0.86

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.03	0.02	1233.34	-1.53	.126
Low to Moderate CommonLit Usage Group	0.03	0.01	13563.66	2.37	.018
High CommonLit Usage Group	0.11	0.02	17874.01	4.62	< .001
Pre-Assessment Scores	-0.45	0.002	177337.67	-186.93	< .001
Title I Eligibility	0.03	0.02	497.57	1.47	.142
Grade 10	0.001	0.01	26396.90	0.17	.862
Grade 11	0.01	0.01	27763.58	0.99	.322
Grade 12	-0.02	0.01	36865.36	-1.57	.117
Classroom Effect: Avg. Pre-Assessment Scores	0.26	0.01	4052.37	23.01	< .001

Table A14. Hierarchical linear model predicting reading growth from CommonLit usage (continuous) among high school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.04	0.21
Residual	0.74	0.86

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.01	0.02	563.43	-0.57	.572
Total CommonLit Lessons	0.01	0.001	178558.38	9.98	< .001
Pre-Assessment Scores	-0.45	0.002	177335.94	-187.10	< .001
Title I Eligibility	0.03	0.02	492.60	1.25	.213
Grade 10	0.003	0.01	26390.86	0.38	.701
Grade 11	0.02	0.01	27532.23	1.58	.114
Grade 12	-0.01	0.01	36214.76	-0.41	.684
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	93077.89	-6.79	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.26	0.01	4067.99	23.30	< .001

Table A15. Hierarchical linear model testing for moderation by Title I eligibility among high school students.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.05	0.21
Residual	0.74	0.86

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.01	0.02	559.84	-0.58	.560
Total CommonLit Lessons	0.01	0.001	84844.84	9.32	< .001
Pre-Assessment Scores	-0.45	0.002	177340.81	-187.10	< .001
Title I Eligibility	0.03	0.02	489.98	1.27	.202
Grade 10	0.003	0.008	26384.08	0.39	.701
Grade 11	0.02	0.01	27528.02	1.58	.114
Grade 12	-0.01	0.01	36192.44	-0.41	.684
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	92896.41	-6.79	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.26	0.01	4067.66	23.30	< .001
CommonLit Lessons x Title I Eligibility	-0.0001	0.001	3169.24	-0.34	.735

Table A16. Non-linear (spline) hierarchical model predicting reading growth from CommonLit usage among high school students, with predicted growth at usage benchmarks.

Random Effect	Variance	Std. Dev.
Teacher (Intercept)	0.06	0.25
School (Intercept)	0.04	0.21
Residual	0.74	0.86

Predictor	Estimate	SE	df	t-value	p-value
Intercept	-0.18	0.02	1836.81	-7.98	< .001
Total CommonLit Lessons (Spline Term 1)	0.84	0.09	56303.25	9.45	< .001
Total CommonLit Lessons (Spline Term 2)	1.27	0.17	71543.48	7.29	< .001
Total CommonLit Lessons (Spline Term 3)	1.39	0.29	21329.25	4.84	< .001
Pre-Assessment Scores	-0.45	0.002	177331.67	-187.07	< .001
Title I Eligibility	0.03	0.02	480.14	1.21	.228
Grade 10	0.01	0.01	26218.34	0.57	.568
Grade 11	0.02	0.01	27401.30	1.86	.063
Grade 12	0.001	0.01	35936.97	0.08	.938
Classroom Effect: Avg. CommonLit Lessons	-0.01	0.001	84434.52	-5.84	< .001
Classroom Effect: Avg. Pre-Assessment Scores	0.27	0.01	4069.14	23.77	< .001

Percentile	Total Lessons Accessed	Predicted Growth (z-score)
10%	0	-0.16
25%	3	-0.11
50%	8	-0.04
75%	19	0.05
90%	44	0.23

Appendix E: Baseline Equivalence Tests

[What Works Clearinghouse \(WWC\)](#) states that the baseline difference between groups in Hedge's g must be less than or equal to 0.05 for comparisons to be considered equivalent without statistical adjustments. If Hedge's g is between 0.05 and 0.25, statistical adjustments for baseline differences are required to satisfy baseline equivalence. To test for baseline equivalence, analyses of variance (ANOVAs) were conducted on the standardized Pre-Assessment scores by usage groups and Hedge's g was calculated for each comparison. Table A17 provides the descriptive statistics for students' standardized scores on the Pre-Assessment within each grade band and usage group.

Table A17. Average standardized scores on the Pre-Assessment by grade band and usage group.

Grade Band	Usage Group	<i>N</i>	<i>M</i>	<i>SD</i>
Elementary	No recorded usage	2,822	0.09	0.99
	Low to moderate usage	18,396	-0.01	1.00
	High usage	1,046	-0.09	0.99
Secondary	No recorded usage	23,270	0.21	1.04
	Low to moderate usage	243,072	-0.01	1.00
	High usage	33,406	-0.07	0.96
Middle	No recorded usage	4,273	0.25	1.04
	Low to moderate usage	90,504	0.00	1.00
	High usage	24,499	-0.05	0.97
High	No recorded usage	18,997	0.20	1.04
	Low to moderate usage	152,568	-0.02	0.99
	High usage	8,907	-0.14	0.95

For the elementary school grade band, baseline scores on the Pre-Assessment differed significantly across usage groups, $F(2, 22261) = 17.98$, $p < .001$. Follow-up analyses showed that the no usage group had significantly higher baseline scores ($M = 0.09$, $SD = 0.99$) than the low to moderate usage group ($M = -0.01$, $SD = 1.00$), $p < .001$, and the high usage group ($M = -0.09$, $SD = 0.99$), $p < .001$; the low to moderate usage group also scored significantly higher than the high usage group, $p = .019$. Baseline scores also differed significantly across usage groups for the secondary grade band, $F(2, 299745) = 627.4$, $p < .001$, the

middle school grade band, $F(2, 119273) = 158.3$, $p < .001$, and the high school grade band, $F(2, 180469) = 509.8$, $p < .001$. Within the secondary grade band, follow-up analyses showed that the no usage group had significantly higher baseline scores ($M = 0.21$, $SD = 1.04$) than the low to moderate usage group ($M = -0.01$, $SD = 1.00$), $p < .001$, and the high usage group ($M = -0.07$, $SD = 0.96$), $p < .001$; the low to moderate usage group also scored significantly higher than the high usage group, $p < .001$. Within the middle school grade band, the no usage group had significantly higher baseline scores ($M = 0.25$, $SD = 1.04$) than the low to moderate usage group ($M = 0.00$, $SD = 1.00$), $p < .001$, and the high usage group ($M = -0.05$, $SD = 0.97$), $p < .001$; the low to moderate usage group also scored significantly higher than the high usage group, $p < .001$. Within the high school grade band, baseline scores were significantly different across the no usage group ($M = 0.20$, $SD = 1.04$), low to moderate usage group ($M = -0.02$, $SD = 0.99$), and high usage group ($M = -0.14$, $SD = 0.95$), all p 's $< .001$. We calculated Hedge's g for each group comparison for each grade band (see Table A18).

Table A18. Baseline equivalence comparisons by grade band.

Grade Band	Comparison	Hedge's g
Elementary	No recorded usage vs. low to moderate usage	0.10
	No recorded usage vs. high usage	0.18
	Low to moderate usage vs. high usage	0.08
Secondary	No recorded usage vs. low to moderate usage	0.22
	No recorded usage vs. high usage	0.28
	Low to moderate usage vs. high usage	0.06
Middle	No recorded usage vs. low to moderate usage	0.25
	No recorded usage vs. high usage	0.31
	Low to moderate usage vs. high usage	0.05
High	No recorded usage vs. low to moderate usage	0.22
	No recorded usage vs. high usage	0.34
	Low to moderate usage vs. high usage	0.12

Because Hedge's g fell between 0.05 and 0.25 for several group comparisons across grade bands, we included students' Pre-Assessment scores in analyses to statistically account

for these differences. Three comparisons, however, exceeded the WWC threshold of 0.25: the no usage vs. high usage comparison within the secondary grade band ($g = 0.28$), the no usage vs. high usage comparison within the middle school grade band ($g = 0.31$) and the no usage vs. high usage comparison within the high school grade band ($g = 0.34$). Per WWC standards, baseline equivalence cannot be established for these three comparisons even with statistical adjustment for Pre-Assessment scores. As a result, findings regarding the relationship between high CommonLit usage and reading growth for secondary, middle, and high school students should be interpreted with some caution, as pre-existing differences between the no usage and high usage groups in these grade bands were large enough that statistical adjustment may not fully account for their influence on the observed growth outcomes. All other group comparisons across grade bands fell within the 0.05-0.25 range and were adequately addressed through the inclusion of Pre-Assessment scores as a covariate in subsequent analyses.