



COMMONLIT

DIGITAL: EVIDENCE OF EFFECTIVENESS

PREPARED BY AD HOC ANALYTICS, LLC FOR COMMONLIT, INC.



UPDATED MAY 2022



INTRODUCTION

CommonLit is a nonprofit organization dedicated to ensuring that all students, particularly low-income students, graduate high school with the reading, writing, critical thinking, communication, and collaboration skills necessary to succeed in college and career. To achieve these goals, CommonLit operates a free online reading program at www.commonlit.org for 3rd-12th grade classrooms. On the website, teachers can select lessons and reading passages and assign them to students.

In 2016, CommonLit was awarded a grant from the Department of Education under the Innovative Approaches to Literacy program to develop the online tool and evaluate its effectiveness. This report was prepared by Ad Hoc Analytics, LLC to summarize the key findings of the evaluation efforts between October 2016 and June 2020.





METHODS

This report is based on evaluation conducted by Ad Hoc Analytics between October 2016 and April 2021. To evaluate CommonLit's effectiveness, the research team investigated the following questions:

- Do students who use CommonLit more frequently perform better on CommonLit's grade-level reading assignments?
- Do students who use CommonLit perform better on state standardized English Language Arts (ELA) assignments?
- Do students in low-income (Title I) schools experience greater gains on CommonLit's assignments?
- Do students in Special Education classes experience greater gains on CommonLit's assignments?
- Do students who use CommonLit's Guided Reading Mode feature experience greater gains on CommonLit assignments?
- Do students in Mexico experience gains from using CommonLit's Spanish library?

To answer these research questions, Ad Hoc Analytics studied numerous data sets involving millions of rows of data. This report summarizes five salient findings that emerged through their evaluation efforts.





KEY FINDING #1

STUDENTS EXPERIENCE GREATER GAINS WITH INCREASED COMMONLIT USAGE

Ad Hoc Analytics examined student performance on CommonLit reading assignments. For the 157,632 students that completed at least four reading assignments¹ during the 2017-2018 school year, evaluators found that the growth in student performance tended to increase as they completed more CommonLit assignments. These results were statistically significant ($p < 0.001$) for up to 25 assignments (see Appendix A). Exhibit 1 illustrates these findings.

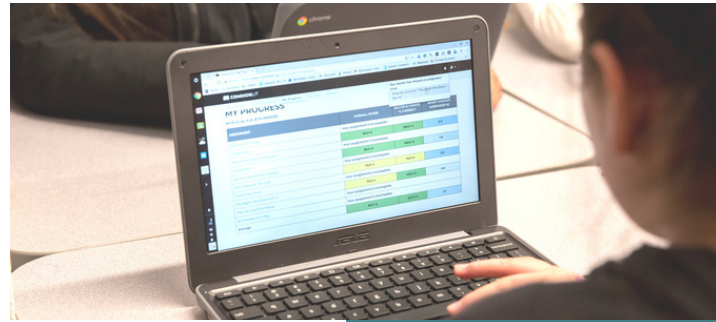
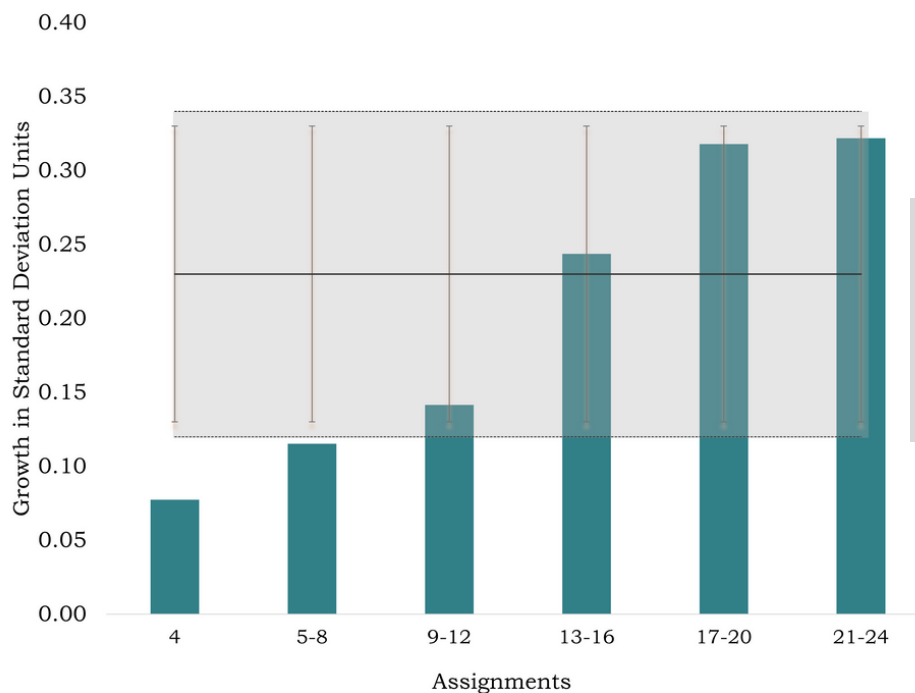


EXHIBIT 1: SCHOOL YEAR GAINS IN READING PERFORMANCE BY NUMBER OF COMPLETED ASSIGNMENTS (2017-2018)



Expected reading gains for 1 year of learning in grades 6-8 (Source: Bloom, et al, 2008).²

¹ The average score of the first two assignments was the baseline measure. The average score of the final two assignments was the follow-up measure. This methodology required a minimum of four completed assignments for each student. Those with fewer completed readings were excluded from analysis.

² According to Bloom et al. (2008, Table 1, p. 10), the average expected reading gain for 1 year of learning in grades 6-7 English Language Arts is 0.23 (+/-0.11) standard deviation units. Students who completed 9 or more CommonLit Library lessons demonstrated reading gains within this range.



KEY FINDING #2

INCREASED COMMONLIT USAGE IS ASSOCIATED WITH MORE STUDENTS AT PROFICIENT OR MASTERY LEVEL ON A STATE ASSESSMENT

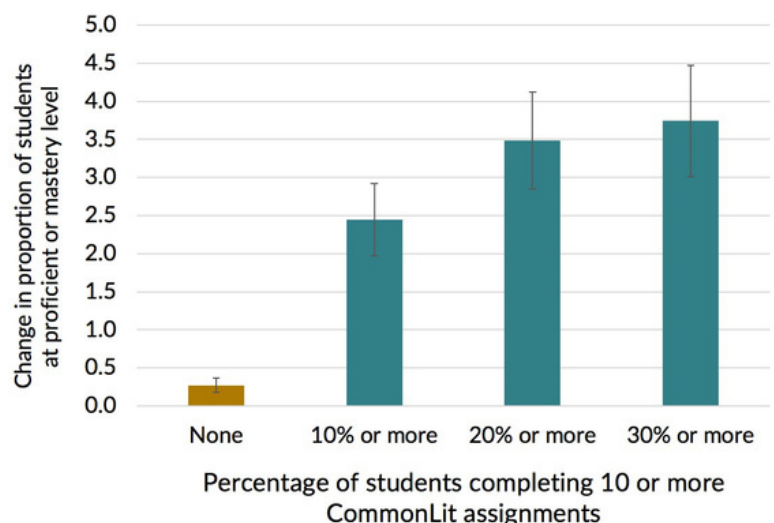
An important next step was to analyze whether CommonLit performance gains measured using CommonLit assignments were meaningful indicators of student performance on standardized state English Language Arts assessments. To measure whether students who used CommonLit did better on state assessments than students who did not, Ad Hoc Analytics studied publicly available data on student performance on the Florida Standards Assessment (FSA). This data is reported by school and grade, so they defined each school and grade as a cohort. Then, they matched this data with CommonLit's school and grade level usage data in Florida.

SUMMARY OF FINDING

Ad Hoc Analytics found that the relationship between CommonLit usage and performance on the FSA varied based on how many CommonLit assignments students completed. Specifically, cohorts in which students used CommonLit at higher levels saw greater improvements on their state assessments compared to cohorts where CommonLit usage was nonexistent or minimal.

Exhibit 2 shows that in schools where a higher proportion of students completed at least 10 CommonLit assignments, a higher proportion of students scored at level 4 or 5 on the FSA compared to the year before. This finding was statistically significant ($p < 0.05$).

EXHIBIT 2: PERCENTAGE OF STUDENTS AT PROFICIENT OR MASTERY LEVEL (FSA) AS A FUNCTION OF COMMONLIT USAGE





DATA SUMMARY

The FSA data included end-of-year English Language Arts (ELA) assessments for the 2016-2017 and 2017-2018 school years. The 2016-2017 data represented the baseline measure of student ELA performance *before* exposure to CommonLit. The 2017-2018 data represented ELA performance *after one year of exposure* to CommonLit for participating cohorts. Ad Hoc Analytics measured CommonLit usage as the percentage of students in each cohort who completed CommonLit assignments. The final dataset contained 7,360 cohorts, of which 1,260 included CommonLit students. These 1,260 cohorts were composed of 150,164 students who completed 695,377 CommonLit assignments.

MODEL

Ad Hoc Analytics used regression analysis to assess the effect of CommonLit usage on the percentage of students in each cohort with high achievement on the FSA. High achievement was defined as scoring at a level 4 (“proficient”) or 5 (“mastery”) on the FSA’s 5-point scale. The model used the following control variables: baseline ELA performance, grade level, and percentages of students in different demographic groups (females, under-represented minorities, English Language Learners, or economically disadvantaged). The final model was found to be a strong predictor of state assessment outcomes, explaining 49 percent of the variance in the percentage of students with high achievement within cohorts (See Appendix B).



KEY FINDING #3

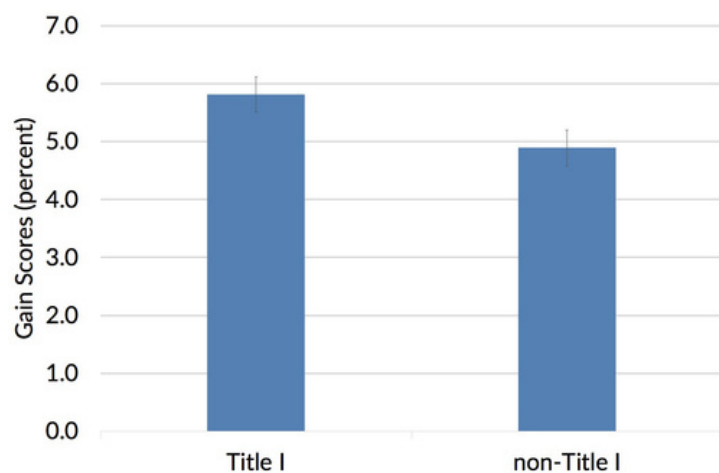
STUDENTS FROM TITLE I SCHOOLS HAVE GREATER GAINS ON COMMONLIT

Earlier research found that at least 59 percent of students using CommonLit come from Title I designated schools or small rural schools (as defined in the federal Small Rural School Achievement program).³ Given that CommonLit reaches mostly students in high-need schools, an important next step was to study how it benefits these students.

To determine whether CommonLit might have a special impact on low-income students, Ad Hoc Analytics studied the relationship between Title I status (a school-level characteristic) and student-level growth in performance on CommonLit assignments over the course of the school year. Specifically, evaluators asked whether students from Title I schools had greater gains in CommonLit performance compared to students from non-Title I schools.

Using CommonLit assignment results from the 2017-2018 school year, Ad Hoc Analytics found that students from Title I schools experienced greater gains in reading performance than students from non-Title I schools, as shown in Exhibit 3 ($p = 0.025$, Cohen's $d = 0.042$). Although the effect size was relatively small,⁴ researchers have argued that outcomes such as this one can be meaningful, particularly when they impact many students.⁵

EXHIBIT 3: GAIN SCORES BY TITLE I STATUS



³ Title I is a federal program that provides funding to schools where at least 40% of students come from low-income families. The Small Rural School Achievement program is a federal program for small schools located in areas with low population density.

⁴ Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

⁵ Glass, G.V., McGaw, B. and Smith, M.L. (1981). *Meta-Analysis in Social Research*. London: Sage.



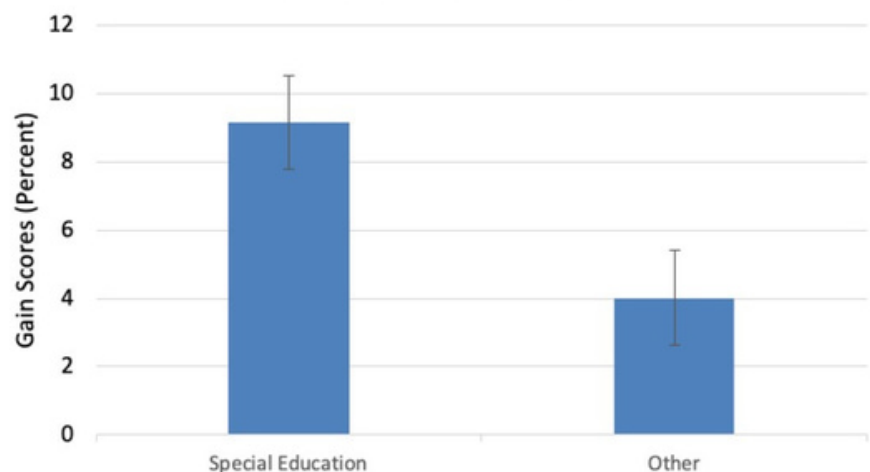
KEY FINDING #4

STUDENTS IN SPECIAL EDUCATION CLASSES HAVE GREATER GAINS ON COMMONLIT

Ad Hoc Analytics continued its analysis to determine whether CommonLit has a special positive effect for students in Special Education classes. To do so, they used a similar method they had used when analyzing CommonLit's impact in Title I schools, comparing students in classes labeled as Special Education to those in classes without this designation. Specifically, they looked at students' scores on CommonLit assignments from the beginning of the 2018-2019 school year to their CommonLit scores at the end of the school year. More than 30,000 students in Special Education classes used CommonLit in 2018-2019, and this analysis focused on those that completed at least two on grade-level assignments at the beginning and end of the school year.

First, they found that more students in Special Education classes on CommonLit experienced gains in their performance. 63 percent (63.0%) of these students increased their scores over the academic year, compared to 57 percent (57.3%) of General Education students.

EXHIBIT 4: GAIN SCORES (PERCENT) OF SPECIAL EDUCATION AND GENERAL EDUCATION STUDENTS WHOSE SCORES IMPROVED IN SY18-19



Second, students in Special Education classes experienced greater gains over the course of the academic school year compared to their peers ($p < 0.01$, Cohen's $d = 0.242$). Students in Special Education classes improved their CommonLit assignment scores by an average of 9.2 percentage points while their peers in General Education classes improved their assignment scores by 4.0 percentage points. CommonLit's free website includes tools like text-to-speech read aloud, enhanced text options, translation into 26 languages, and text chunking (Guided Reading Mode).



KEY FINDING #5

STUDENTS WHO USE GUIDED READING MODE HAVE GREATER GAINS ON COMMONLIT

Ad Hoc Analytics sought to measure the impact of Guided Reading Mode (GRM) on students' reading comprehension. GRM "chunks" the text for students, only displaying a small paragraph at a time. The feature asks students to answer a comprehension question, and a correct answer "unlocks" the next portion of the text. Evaluators looked at whether students who use GRM more frequently made greater gains in performance on CommonLit assignments.

Because GRM can be enabled or disabled for individual assignments, they calculated a Guided Reading Mode Index (GRMI), defined as the proportion of assignments submitted with GRM enabled. They used a multiple regression analysis to measure the contribution of the GRMI on gains in students' performance on CommonLit assignments. Employing a multiple regression model allowed them to control for other variables that can contribute to gains in students' performance. These variables included schools' Title I status and whether classes were designated as Special Education, both of which were shown to impact student gains (see Key Findings #3 and #4).

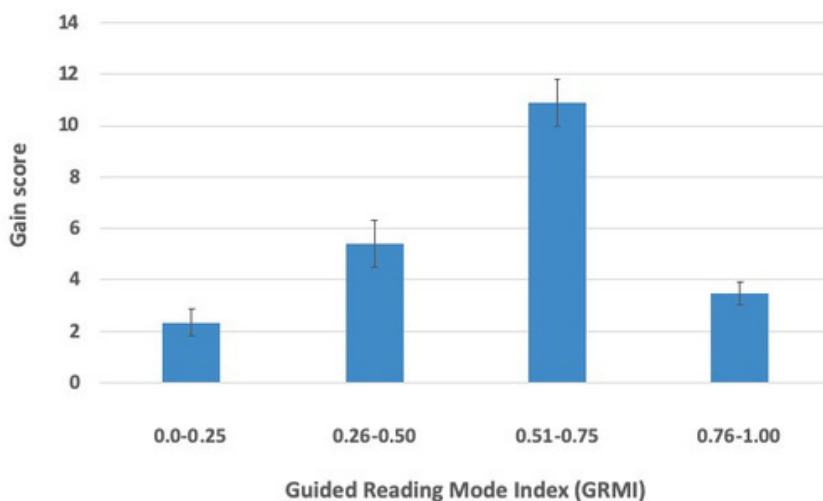


EXHIBIT 5: GAIN SCORES FOR STUDENTS BY GRMI LEVEL

Ad Hoc Analytics found that when a higher proportion of a students' assignments had GRM enabled, they experienced greater performance gains. Exhibit 5 shows that as the proportion of assignments with GRM enabled increased, so did students' performance gains, and that the greatest gains were seen when 51% to 75% of a student's assignments had GRM enabled.



KEY FINDING #6

STUDENTS IN MEXICO HAD SIGNIFICANT GAINS ON COMMONLIT

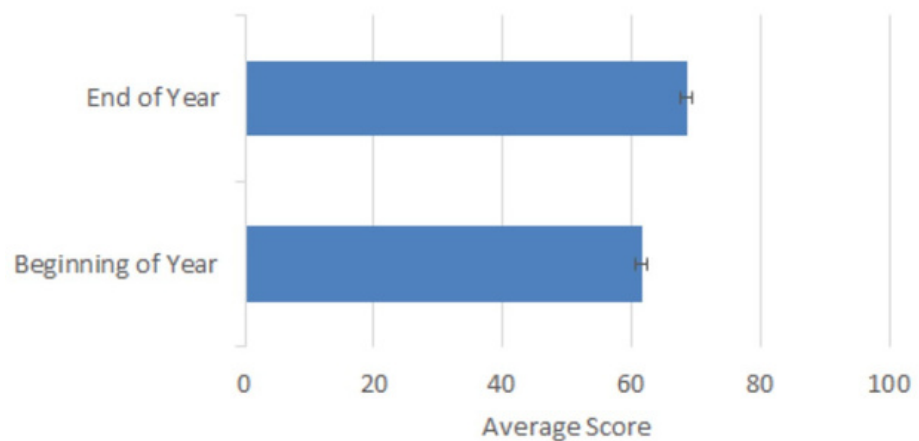
CommonLit's Spanish library offers localized content for students in Latin America. CommonLit has formed partnerships with local governments, NGOs, and networks of schools to deliver training and curriculum to teachers. To assess the impact of these efforts, Ad Hoc Analytics examined growth in performance on CommonLit assignments over the 2019-2020 school year for students in Mexico. This analysis was designed to assess performance gains for Mexican students using CommonLit in Spanish.

The average gain in performance for the 641 students in this analysis was found to be 7.0 points. A paired t-test found this comparison to be statistically significant, $t(640) = 7.137, p < 0.001$, Cohen's $d = 0.282$.

¹Although the effect size was relatively small,¹ researchers have argued that outcomes such as this one are meaningful, particularly when they impact many students.²

The 641 Mexican students completed a total of 10,942 CommonLit assignments in Spanish, averaging 17.1 assignments each.

EXHIBIT 6: STUDENT PERFORMANCE SCORES SY19-20 (MEXICO ONLY)



¹ Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

² Glass, G.V., McGaw, B. and Smith, M.L. (1981). *Meta-Analysis in Social Research*. London: Sage.



APPENDIX A:

METHODOLOGY FOR KEY FINDING #1

Ad Hoc Analytics examined student performance growth on CommonLit assignments during the 2017-2018 school year. First, they measured the average difference between students' early CommonLit assignment scores compared to their later scores and determined whether these differences were statistically significant. Second, the research team examined the factors associated with students' growth using multilevel modeling.

DATA

The data included results from students who had at least four graded CommonLit assignments that were on their grade level. By limiting the sample in this way, evaluators could compare the mean score of the first two assignments with the mean score of the most recent two assignments for each student.¹ Further, they only included students who had completed their first two assignments at the beginning of the school year and their most recent two assignments at the end of the school year. The use of two assignments at each end of the year allowed evaluators to establish a degree of reliability in their measures without overly restricting the available sample for analysis. The final sample included assignment scores from 157,632 students who had at least four graded assignments.



¹ The sample for these analyses eliminated any assignments that were assigned, not started; in progress; or submitted, since these assignments had not yet received their final grade. In addition, we eliminated any assignments that were completed by students who were above or below their grade level to control for the relative rigor or ease of the assignment. Finally, evaluators excluded any student who did not complete at least 4 assignments.



ANALYSIS

Ad Hoc Analytics first studied whether students' performance on CommonLit assignments improved significantly from the beginning to the end of the school year. They then used multilevel modeling to determine which factors at the student, classroom, and school levels contributed most to students' improvement.

Evaluators first assessed whether the data for each number of completed assignments followed a normal distribution. When the distributions were normal, they used a dependent samples t-test. When they were non-normal, they used the Wilcoxon Signed-Rank Test, which is appropriate for non-normal distributions. When normal distributions were detected, evaluators assessed the mean difference in performance between the earliest and most recent assignments. When non-normal distributions were detected, they assessed the median difference in performance. Finally, they did not assess the statistical significance of test differences for samples smaller than 10 students.

For the entire sample of students, performance on most recent assignments ($Mdn = 66.3$) was significantly higher than earliest assignments ($Mdn = 63.4$; $z = -42.74$ $p = 0.001$). The effect size of this relationship is small ($r = 0.11$). Considering the brevity of student time spent on these assignments, a small effect size is still very meaningful. It is also noteworthy that as the number of assignments completed by students grew, students' scores increased by a greater margin. Table A.1 shows the gains in student scores by the number of CommonLit assignments they completed.

TABLE A.1:
SCHOOL YEAR
GAINS IN
READING
PERFORMANCE
BY NUMBER OF
COMPLETED
COMMONLIT
ASSIGNMENTS
(2017-2018)

Number of Assignments	Median Difference	Z-test Statistic	Degrees of Freedom/ N	Statistical Significance
4 assignments	1.8	-19.385	62,582	***
5-8 assignments	2.9	-31.621	75,326	***
9-12 assignments	4.2	-17.031	14,492	***
13-16 assignments	5.6	-14.86	3,724	***
17-20 assignments	8.5	-8.92	790	***
21-24 assignments	13	-6.511	412	***
25+ assignments	2.5	-1.502	306	
All assignments	2.81	-42.735	157,632	***

* $p < .05$; ** $p < .01$; *** $p < .001$



MEASURES

Next, evaluators used multilevel modeling to assess the factors that contributed to student improvement on assignments. Because students are nested within classrooms, which are nested within schools, it was necessary to control for this nested data structure with multilevel modeling. The multilevel model examined the student, classroom, and school level predictors of success on CommonLit assignments. These predictors are presented in Table A.2.

TABLE A.2:
STUDENT, CLASS, AND SCHOOL LEVEL
EFFECTS ON COMMONLIT SCORES

Student-Level	Coefficient (B)	Standard Error	Statistical Significance
Pre-Test Scores	.36	.002	***
CommonLit Participation (2 yrs)	.78	.51	
CommonLit Participation (2 semesters)	1.10	.22	***
High Usage (15+ Courses)	2.29	.43	***
Grade 3	-5.68	1.07	***
Grade 4	-2.19	.79	**
Grade 5	-1.76	.54	***
Grade 6	-.39	.38	
Grade 7	-1.03	.30	***
Grade 9	1.41	.36	***
Grade 10	-1.06	.42	*
Grade 11	.68	.53	
Grade 12	-5.35	.76	***
Intercept	61.60	.21	***
Classroom-Level			
Small-Medium Usage Classroom (5-19 students)	1.19	.59	*
Medium Usage Classrooms (20-34 students)	2.79	.66	***
Medium-High Usage Classrooms (35-49 students)	3.83	.75	***
High Usage Classrooms (50+ students)	4.47	.56	***
School-Level			
Title 1	-3.14	0.42	***
Variance Explained		Total Variance Explained at Each Level	Proportion of Variance Explained by Variables
Student-Level (e)		68.5%	13.7%
Classroom-Level (r)		25.5%	36.3%
School-Level (u)		5.9%	61.5%



Student Level Predictors

To control for students' innate academic abilities, evaluators included a "pre-test score" variable, measured by students' average on their first two CommonLit assignments. Students who scored higher on their first two assignments were more likely to earn higher scores on their end of year assignments. In addition, students who submitted more CommonLit assignments also tended to improve final performance on their most recent two assignments. Specifically, students who completed fifteen or more assignments increased their performance by an average of almost 2.3 points. Next, students who participated in CommonLit for two semesters during the 2017-18 school year were significantly more likely to improve their reading scores compared to students who participated in CommonLit for only one semester by an average of 1.1 points.

While there was a statistically significant difference between scores for students in 8th grade (used as the baseline grade) and students in 4th, 5th, 7th, 9th, 10th, and 12th grades, their scores only differed by a couple of percentage points. Students in 3rd grade saw lower scores than students in 8th grade, which is likely due to the fact that there were fewer 3rd grade students using CommonLit and fewer options for 3rd grade assignments on the site. There were no significant differences between 6th and 11th grade students' final assignment scores and those of 8th grade students. Taken together, these findings indicate that it was important to control for students' grade levels in the model, but grade levels do not account for major differences in students' final scores.

Classroom- and School-Level Predictors

Evaluators found that students whose teachers assigned multiple students CommonLit assignments tended to improve their final CommonLit scores. At the class level, as the number of students completing CommonLit assignments increased, students' scores grew. Specifically, students whose teacher assigned CommonLit to 5-19 students increased their



scores by almost 1.2 points, followed by an increase of 2.8 points (20-34 students), 3.8 points (35-49 students), and 4.5 points (more than 50 students). This finding suggests that teachers who used CommonLit more widely were able to engender more performance growth in their students than teachers who used CommonLit with fewer students. At the school-level, students who attended schools designated as Title I had significantly lower final scores than students who attended schools with wealthier students.



FINDINGS

On average, students who completed at least four CommonLit assignments increased their performance by 2.5 points during the 2017-2018 school year. As students completed more assignments, they tended to improve their scores by wider margins, as shown in Exhibit 1.

The multilevel model predicted students' end of year performance on CommonLit assignments. 69 percent of the variation that predicts student-level performance on CommonLit assignments was explained by student-level characteristics. An additional 26 percent of variability in the outcome was at the classroom level and 6 percent at the school level. While our student-level measures account for only 14 percent of the variation in CommonLit scores, 36 and 62 percent of classroom and school level variation are accounted for by our measures at each of those levels.

These analyses confirm that students' performance on CommonLit assignments improved over the 2017-2018 school year. Further, as the number of CommonLit assignments students completed increased, so too did the amount of growth they showed. Multilevel modeling showed that variables at the student, classroom, and school levels contributed to students' end of year scores on CommonLit assignments. Importantly, students with high CommonLit usage and students whose teachers used CommonLit more widely showed greater growth than students for whom CommonLit was only a small part of their curriculum.



APPENDIX B:

METHODOLOGY FOR KEY FINDING #2

Ad Hoc Analytics also assessed whether participation in CommonLit increased student performance at the grade level on a standards-aligned state assessment.

DATA

This analysis drew information from the following two data sources:

Florida Assessment Data: Evaluators compiled two years of English Language Arts (ELA) end of year assessment scores for Florida public schools, aggregated at the grade level ($n = 11,229$). For this analysis, 2016-2017 data represents the baseline measure of student ELA performance before exposure to CommonLit. The 2017-2018 data represents ELA performance after one year of exposure to CommonLit for participating grades within schools.

To analyze specific cohorts of students, 2016-2017 assessment data were matched to 2017-2018 assessment data for the next grade within the same school (for instance, 3rd grade data from 2016-2017 was matched to 4th grade data from 2017-2018). As a result of this alignment, a number of rows of data were lost from analysis ($n = 1,158$) due to missing baseline data (e.g., no 2nd grade FSA data) or missing follow up data (e.g., no 6th grade FSA data for schools without 6th grade).



CommonLit Data: Ad Hoc Analytics used data from graded CommonLit assignments completed by students in Florida public schools submitted during the 2017-2018 school year. Because this analysis was conducted at the grade level, the final dataset only included students whose grade could be identified. Ultimately, the dataset represented 150,164 students who completed 695,377 CommonLit assignments. Evaluators then aggregated students to their grade level and school to identify those cohorts with students who had taken CommonLit assignments.

The final dataset included 1,260 cohorts with at least one CommonLit student and 6,100 cohorts with no CommonLit students, bringing the total to 7,360 cohorts.

MEASURES

Ad Hoc Analytics built a model designed to predict the proportion of students within a cohort who had reached the proficient or mastery levels (Level 4 or Level 5) on the FSA during the 2017-18 school year (Florida Dept. of Education, 2019). To control for the academic ability of students within classes, evaluators used the average score of the FSA (by grade) during the 2016-17 school year. They called this variable "assessment pre."

CommonLit Usage: In total, Ad Hoc Analytics created five variables that measure CommonLit usage by grade:

- 1) *CommonLit usage (any):* proportion of students who completed at least one CommonLit assignment during the 2017-2018 school year
- 2) *Low CommonLit usage:* proportion of students who completed only one CommonLit assignment during the 2017-2018 school year²
- 3) *High CommonLit usage:* proportion of students who completed ten or more CommonLit assignments during the 2017-2018 school year
- 4) *Two years CommonLit usage:* proportion of students who also had a CommonLit assignment from the 2016-2017 school year
- 5) *Two to four teachers CommonLit usage:* proportion of students who were given CommonLit assignments by two or more teachers



² They performed quartile analyses using the CommonLit user data aggregated to the student-level. Out of the 150,164 students, evaluators found that the lowest quartile included all students with only one assignment (n=44,539 students), while the highest quartile included those students with six or more CommonLit assignments (n=41,336 students).



Control Variables: Ad Hoc Analytics included several control variables to help further explain the cohorts' academic growth:

- 1) *Female students percentage:* proportion of female students³
- 2) *ELL percentage:* proportion of English Language Learners⁴
- 3) *Under-represented minority percentage:* proportion of students identifying as African American, Hispanic, or two or more races
- 4) *Percent economically disadvantaged:* proportion of students from economically disadvantaged families⁵
- 5) *Grade 06 - Grade 09 indicator:* variables indicating grade level

ANALYSIS

Ad Hoc Analytics performed a regression analysis to assess the relative contribution of each of the defined variables in predicting the proportion of students who reached level 4 or 5 on the FSA for all cohorts. They employed standard multiple regression procedures and checks (Field, 2005; Tabachnick & Fidell, 2007).

FINDINGS

The regression analysis did predict the proportion of students who reached mastery on the FSA $F(14,2721) = 185.543, p < 0.001$. The effect size (adjusted R^2) for this analysis was 0.487, meaning that a little less than 49 percent of the variance in the percentage of students in each cohort scoring a 4 or 5 on the FSA can be explained by the predictor variables selected. The coefficients table generated for this analysis (see Table B.1) indicates that the statistically significant predictors for assessment post were "assessment pre," "CommonLit usage (low)," "percent students with at least 10 completed assignments," "percent economically disadvantaged," "percent female students," "grade 06 indicator," "grade 07 indicator," "grade 08 indicator," and "grade 09 indicator." All other predictor variables were non-significant.

³ Previous research has found that female students tend to earn higher scores on ELA standardized tests than their male peers (Reilly et al., 2019).

⁴ In Florida, ELLs are classified based on their language skills and whether they were raised speaking a native language other than English or were raised in a linguistically isolated family or were raised in an American Indian or Alaskan Native family. Retrieved from: <http://www.fldoe.org/accountability/data-sys/edu-info-accountabilityservices/pk-12-public-school-data-pubs-reports/students.stml>

⁵ Florida Department of Education DOE Information Database Requirements Volume 1: Automated Student Information System Automated Student Data Elements as cited from: www.fldoe.org/core/



TABLE B.1:
REGRESSION COEFFICIENTS
FOR PREDICTING
PERCENTAGE OF STUDENTS
WITH HIGH ACHIEVEMENT ON
2017-2018 FSA ELA

	<i>B</i>	<i>SE B</i>	β
Step 1			
Constant	-103.960	4.112	
Assessment Pre	.409	.013	.514 *
Step 2			
Constant	-61.306	5.584	
Assessment Pre	.314	.016	.395 *
Percent of Students with:			
Any CommonLit Usage	-0.017	0.019	-.025
CommonLit Usage (Low)	-.112	.040	-.061 *
At Least 10 Completed Readings	.129	.052	.046 *
Two to Four Teachers	-.132	.097	-.021
Two Years CommonLit Usage	-.071	.111	-.009
Economically Disadvantaged Percentage	-.276	.011	-.440 *
English Language Learners (ELL) Percentage	-.019	.021	-.016
Female Students Percentage	.164	.035	.065 *
Minority Students Percentage	.004	.011	.007
Grade 06 Indicator	-3.571	.956	-.052 *
Grade 07 Indicator	-5.738	.513	-.167 *
Grade 08 Indicator	-5.457	.567	-.153 *
Grade 09 Indicator	-9.249	2.850	-.045 *

Note $R^2 = .264$ for Step 1; $\Delta R^2 = .223$ for Step 2 ($p s < .001$). * $p < .05$.



Although low CommonLit usage was associated with lower percentages of students with FSA scores at Level 4 or Level 5, when students had high CommonLit usage, higher proportions of them reached Level 4 or Level 5. This finding suggests that when students used CommonLit occasionally (i.e., one assignment), they had poorer performance on their state assessment scores compared to other students. However, when students used CommonLit at high levels (i.e., 10 or more reading assignments), they improved their state assessment scores compared to other students. Students who used CommonLit only minimally did not improve, while students who used CommonLit more frequently saw improvements in their FSA scores in English Language Arts.

Beyond CommonLit usage measures, the significant predictors included "percent economically disadvantaged," "percent female students," and grade indicators. For grades with larger percentages of economically disadvantaged students, the average state assessment score was lower. For grades with larger percentages of female students, the average state assessment score was higher. Both match the reality of findings in education research. Grades 6, 7, 8, and 9 were associated with lower percentages of students that attained a 4 or a 5 on the state assessment at follow up.

One caveat to the current report is the issue of representativeness of students in the highest grade at any school. Because of the need to align data across years (e.g., 6th grade assessment in 2016-2017 with 7th grade assessment in 2017-2018), 6th grade assessment data for 2016-2017 were only retained for analysis if that school also included 7th graders. For those schools without 7th graders, the 6th grade data for 2016-2017 was dropped from the analysis. As a result, the 6th grade assessment data only represents those 6th grade classes in schools where there is also a 7th grade class (e.g. from K-8 schools).



APPENDIX C:

METHODOLOGY FOR KEY FINDING #3

Ad Hoc Analytics also assessed the relationship between student growth in performance on CommonLit assignments and their schools' Title I status. Specifically, they sought to answer whether students from Title I schools had greater gains on CommonLit performance compared to students from non-Title I public schools.

DATA

The dataset for this analysis consisted of graded CommonLit assignments from the 2017-2018 school year. The dataset was restricted in several ways. First, only public school students were included in the analysis, which ensured that the non-Title I schools group did not include students from private or parochial schools. Second, only students with at least two assignments completed before October 1, 2017 and at least two assignments completed after March 30, 2018 were included in the dataset. Restricting the data in this way operationalized the time span of a full academic year between baseline and end of year time points for the analysis. Using four assignments completed at any point during the academic year would have resulted in much shorter time spans between the time points of interest. Further, ensuring students had completed at least two assignments allowed evaluators to get a more robust measure of students' performance at each end of the school year.





MEASURES

Students' beginning of year performance scores were calculated using the mean of their scores on their first two graded assignments from the 2017-2018 school year. Similarly, students' average score on their final two assessments completed in the 2017-2018 school year measured their end of year performance. Gain scores represented the difference between the beginning of year and end of year scores for each student. The Title I status variable came from the National Center for Education Statistics (NCES) Common Core of Data.⁶



ANALYSIS

An independent samples t-test (Tabachnick & Fidell, 2007) was used to determine whether the gain scores for students in Title I schools were significantly different from those of students in non-Title I schools.

FINDINGS

Students from Title I schools experienced greater gains in reading performance, over the course of the school year, compared to those from non-Title I schools, $t(11275) = 2.239$, $p = 0.025$. The effect size for this analysis was $d = 0.042$. While the reported effect size for the current analysis is relatively small (Cohen, 1988), even a small effect when multiplied over ten thousand students has considerable impact for a low-cost, scalable, digital platform. Exhibit 3 shows average gain scores achieved by students over the course of the academic year based on their schools' Title I status. The error bars for each column indicate the standard error of measurement found for each measure.

⁶ <https://nces.ed.gov/ccd/pubschuniv.asp>

APPENDIX D:

METHODOLOGY FOR KEY FINDING #4

Next, researchers studied whether CommonLit had a special effect for students in Special Education. Specifically, they asked whether students in Special Education classrooms experienced greater gains in their scores on CommonLit assignments than students in other classrooms.

DATA

The dataset for this analysis consisted of graded CommonLit assignments from the 2018-2019 school year. The dataset was restricted in similar ways to the Title I analysis (Key Finding #3). The final dataset included a subset of public school students with at least two assignments completed before October 1, 2018 and at least two assignments completed after March 30, 2019.



MEASURES

Students' beginning of year performance scores were calculated as the mean of their scores on graded assignments completed before October 1, 2018. Similarly, students' average score on their assessments completed after March 31, 2019 measured their end of year performance. Gain scores represented the difference between the beginning and end of year scores for each student. When a teacher creates a class on CommonLit, he or she can select its subject(s). Evaluators used classes marked as "Special Education" to determine whether or not a student was in a Special Education classroom. The final dataset included 5,974 students.

ANALYSIS

An independent samples t-test (Tabachnick & Fidell, 2007) was used to determine whether the gain scores for Special Education students were significantly different from that of their peers.



FINDINGS

63 percent (63.0%) of these students increased their scores over the academic year, compared to 57 percent (57.3%) of General Education students. Further, students in Special Education classes experienced greater gains over the course of the academic school year compared to their peers, $t(5972) = 3.731, p < 0.001$. The effect size for this analysis was $d = 0.242$, which is considered a small effect (Cohen, 1988). Students in Special Education classes improved their CommonLit assignment scores by an average of 9.2 percentage points while their peers in General Education classes improved their assignment scores by 4.0 percentage points.



APPENDIX E:

METHODOLOGY FOR KEY FINDING #5

Ad Hoc Analytics studied whether student growth on CommonLit assignments from the 2018-2019 school year varied based on how frequently they used Guided Reading Mode (GRM). They used regression analysis to determine what role GRM plays in students' performance growth while controlling for other factors.

DATA

As with Key Findings #3 and #4, Ad Hoc Analytics looked at data from public school students who had completed at least four on-grade assignments across the school year (two before October 1 and two after March 31). Structuring the data in this way allowed them to focus their analysis on a full school year time frame and look at performance changes over a common time period. The final dataset included assignments from 5,983 students.

**MEASURES**

As in the analysis for Key Finding #4, students' beginning of year performance scores were calculated as the mean of their scores on graded assignments completed before October 1, 2018. Similarly, students' average score on their assignments completed after March 31, 2019 measured their end of year performance. Gain scores represented the difference between the beginning and end of year performance scores for each student.

Because GRM can be enabled or disabled for individual assignments, evaluators defined a *Guided Reading Mode Index (GRMI)* to be the proportion of CommonLit assignments the student completed with GRM enabled. For instance, if a student had 10 on-grade assignments, 7 of which had GRM enabled, the GRMI would be 0.7. By calculating the proportion of assignments completed with GRM, evaluators could include the measure at the student level. This index was the primary variable of interest in predicting students' gain scores.



Ad Hoc Analytics also included several control variables to ensure the effect they saw could be attributed to GRM usage and not other factors. These variables included:

- 1) *Beginning of Year Score*: a measure of students' prior reading ability as defined above
- 2) *Title I Schoolwide Status*: an indicator of whether a student's school receives schoolwide Title I funding
- 3) *Special Education Class*: an indicator defined at the class level of whether the student was part of a class designated as a Special Education class
- 4) *English Language Learner (ELL) Class*: an indicator defined at the class level of whether the student was part of a class designated as an ELL class
- 5) *GRMI-ELL Interaction*: the result of multiplying the GRMI by the ELL Status indicator, which allowed evaluators to examine how these two variables work together

ANALYSIS

Ad Hoc Analytics used a multiple regression analysis to predict student gain scores, measuring the relative contribution of the five control variables and the GRMI. They employed standard multiple regression procedures and checks (Field, 2005; Tabachnick & Fidell, 2007).

FINDINGS

Ad Hoc Analytics found that GRM did have a positive impact on students' performance. More specifically, as students' proportion of assignments completed using GRM increased, so did their performance gains.

The multiple regression model proved to be statistically significant in predicting students' gain scores, $F(6,5578) = 326.042, p < .001$. The adjusted R^2 for this analysis was 0.259, meaning that just under 26 percent of the variance in gain scores can be explained by the predictor variables selected.



Table E.1 shows the regression coefficients produced by the model. Of the control variables, only the ELL Class Indicator was not statistically significant ($p = 0.097$). The GRMI-ELL Interaction variable, however, was found to have a negative correlation with students' gain scores, meaning that for students in ELL classes, having a higher proportion of assignments with GRM enabled actually led to smaller gains in reading performance. All of the other control variables were found to be positive predictors of students' gain scores, which is consistent with previous findings.

TABLE E.1:
REGRESSION
COEFFICIENTS FOR
PREDICTING
PERFORMANCE
GAIN SCORES OF
STUDENTS (2018-
2019 SCHOOL YEAR)

	<i>B</i>	<i>SE B</i>	β
Constant	38.341	1.038	
Title I - school wide status	1.133	.556	.024 *
Beginning of year reading performance	-.554	.013	-.504 *
Special Education class indicator	5.501	1.666	.050 *
English Language Learners (ELL) Class indicator	4.243	2.554	.048
Guided Reading Mode Index (GRMI)	1.817	.691	.031 *
GRMI-ELL interaction	-14.942	3.517	-.133 *

Note: adjusted $R^2 = .259$ ($p < .001$). * $p < .05$. This analysis included 5,983 students.



APPENDIX E:

METHODOLOGY FOR KEY FINDING #6

Ad Hoc Analytics examined growth in performance on CommonLit assignments over the 2019-2020 school year for students in Mexico. This analysis was designed to assess performance gains for Mexican students using CommonLit in Spanish.

**DATA**

The dataset for this analysis consisted of CommonLit assignments in Spanish completed by Mexican students during the 2019-2020 school year and graded by their teachers. The dataset was restricted to those students with at least two assignments completed before March 1, 2020 and at least two assignments completed after March 30, 2020. This restriction resulted in an average time span between first and final readings to be 216.0 ($SD = 50.1$) days.

MEASURES

Students' beginning of year performance scores were calculated using the mean of their scores on their first two graded assignments from the 2019-2020 school year. Similarly, students' average score on their final two assessments completed in the 2019-2020 school year measured their end of year performance. Gain scores were defined as the difference between the beginning of year and end of year scores for each student.

ANALYSIS

A paired samples t-test (Tabachnick & Fidell, 2007) was used to determine whether the difference between beginning and end of year performance measures at the student level were significantly different.



FINDINGS

Students in Mexico experienced significant gains in reading performance over the course of the school year, $t(1033) = 6.253$, $p < .001$, Cohen's $d = 0.194$ (small effect size).

When looking at all assignments completed by these 1,034 students ($n = 17,969$), the average number per student was 22.7 assignments, ranging between 4 and 45 assignments for each student. Finally, it was determined that nearly 3 in 4 of these assignments (72.2 percent) were in Spanish.

While the reported effect size for the current analysis is relatively small (Cohen, 1988), even a small effect when multiplied over a thousand students has considerable impact for a low-cost, scalable, digital platform. Exhibit 6 shows average beginning and end of year performance measures achieved by students over the course of the academic year in Mexico. The error bars for each column indicate the standard error of measurement found for each measure.



APPENDIX F:

REFERENCES

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Florida Department of Education (2019). 2018-19 Florida standards assessments English language arts and mathematics fact sheet. Retrieved from <http://www.fldoe.org/core/fileparse.php/5663/urlt/FSAELA-MathFS1819.pdf>

Reilly, D., Neumann, D., & Andrews, G. (2019). Gender Differences in Reading and Writing Achievement: From the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445-458.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.



CommonLit, Inc.
660 Pennsylvania Avenue SE
Suite 302
Washington, DC 20003

info@commonlit.org
www.CommonLit.org