Empirical Education

Newsela

Impact on California Schools

August 1, 2017

This study of Newsela implementation in California schools during the 2015-2016 school year finds evidence of a positive effect of Newsela on reading. This report identifies the effects of active Newsela usage on achievement compared to similar California schools not using Newsela.

Impact of Newsela

Active use of Newsela results in improved student outcomes on the reading portion of the California state test compared to a matched sample of non-users. The effect of Newsela on schools is, on average across grades 4-8, a 1.34 percentage point increase in the proportion of students performing "near standard" or "above standard" in the Reading area of the California Smarter Balanced ELA Assessment. We have high confidence of this result ($p < .05$). Figure 1 shows the overall result.

Schools that were active users of Newsela performed better on the state reading assessment than schools that were not Newsela users.

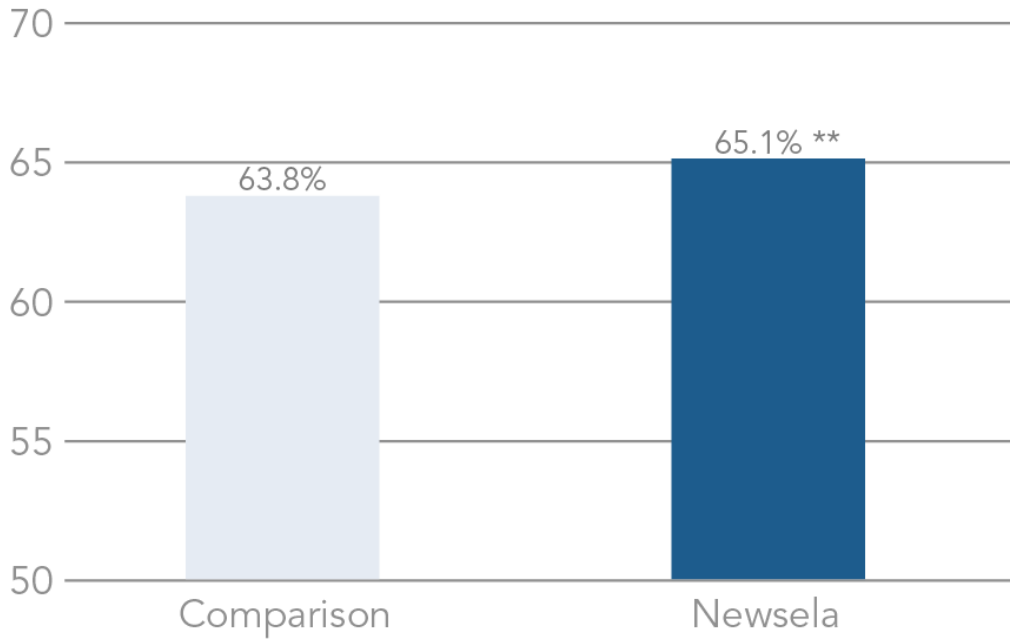

FIGURE 1. EFFECT OF NEWSELA.COMPARISON OF COMPARISON AND NEWSELA SCHOOLS IN PROPORTION OF STUDENTS PERFORMING NEAR OR ABOVE STANDARD, PERCENTAGE POINTS.

Note. Asterisks indicate significant differences between comparison and product (* $p < .2$, ** $p < .05$). Reported percentages are adjusted for group differences at baseline.

While this appears to be a modest result, it is roughly equivalent to an additional 24 days of reading instruction (using an accepted conversion method). Also, it is a conservative estimate in that it is an average result across schools where not all students were active in the program.

Since California reports achievement separately for grades within schools, the impact of Newsela was estimated separately for each grade level included in the analysis. The impact is positive across all grades as displayed in Figure 2. The magnitude of the effect varies across grade levels and is statistically significant only in grades 4, 5 and 7.

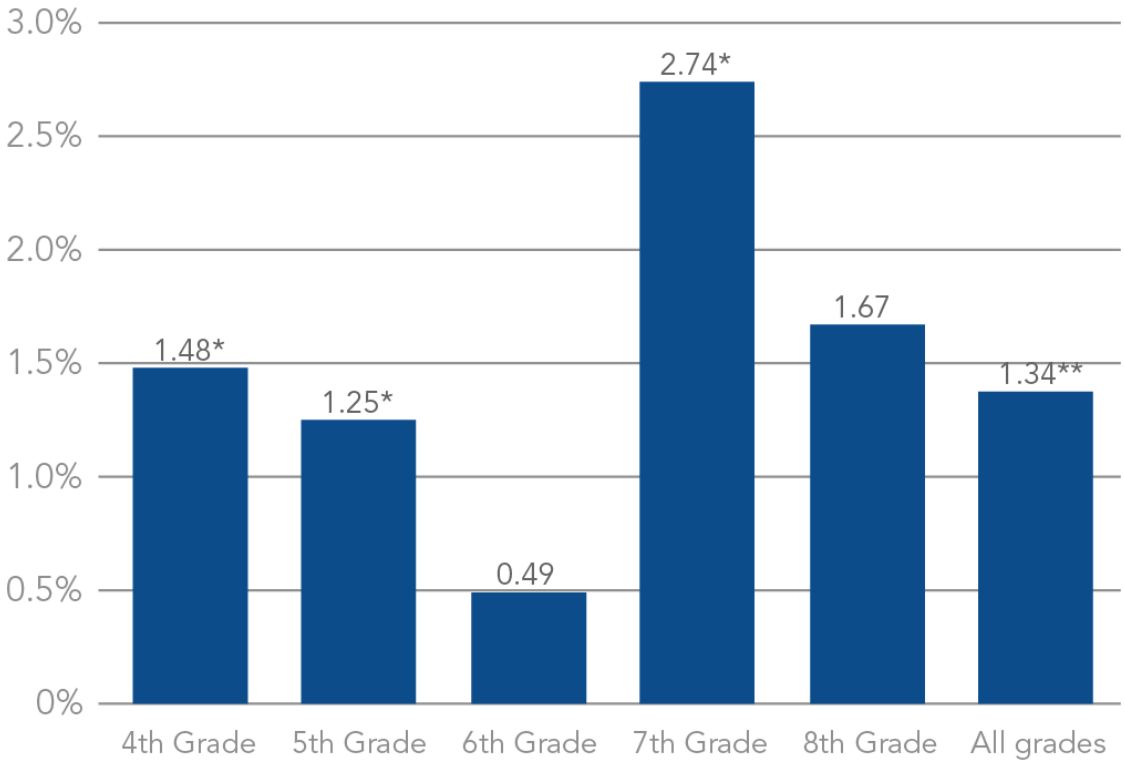

FIGURE 2. EFFECT OF NEWSELA.INCREASE IN PROPORTION OF STUDENTS PERFORMING NEAR OR ABOVE STANDARD, PERCENTAGE POINTS.

Note. Asterisks indicate significant differences between comparison and product (* $p < .2$, ** $p < .05$). Reported percentages are adjusted for group differences at baseline.

Differential Impact

We also tested the impact of Newsela on schools that differed in the percentage of the following subgroups of students: economically disadvantaged, English learners, major ethnic groups (Black, White, and Hispanic). We did not find a difference in the impact of Newsela on schools regardless of the percentage of students with these characteristics.

Newsela had a positive impact on schools regardless of their student subgroup composition.

Correlation between Newsela Use and School Outcomes

There is a significant association between average product usage, measured as the number of Newsela quizzes taken per student, and the student outcome, measured as the percentage of students performing near or above standard on the state reading assessment. This association was found in all grades, and we have high confidence of this result ($p < .05$). No statistically significant differences across grades were identified.

Use of Newsela is positively correlated with reading test performance.

Composition of the Sample

During the 2015-2016 school year, Newsela was implemented in more than eight thousand schools in California. For the purposes of this study, the pool of Newsela schools was limited to those where at least some students took at least one quiz during the year. As shown in technical details, a smaller sample was used for the comparison study.

Students in Newsela schools were similar to the general California population

In the full sample, there were 2,525 such schools, and Newsela users comprised 35% of the student population of these schools. Due to the availability of achievement data, only grades four through eight were included.

Figure 3 shows the distribution of demographic variables in Newsela schools in California. Each bar represents the range of a variable (10th through 90th percentile) with a white line representing the median.

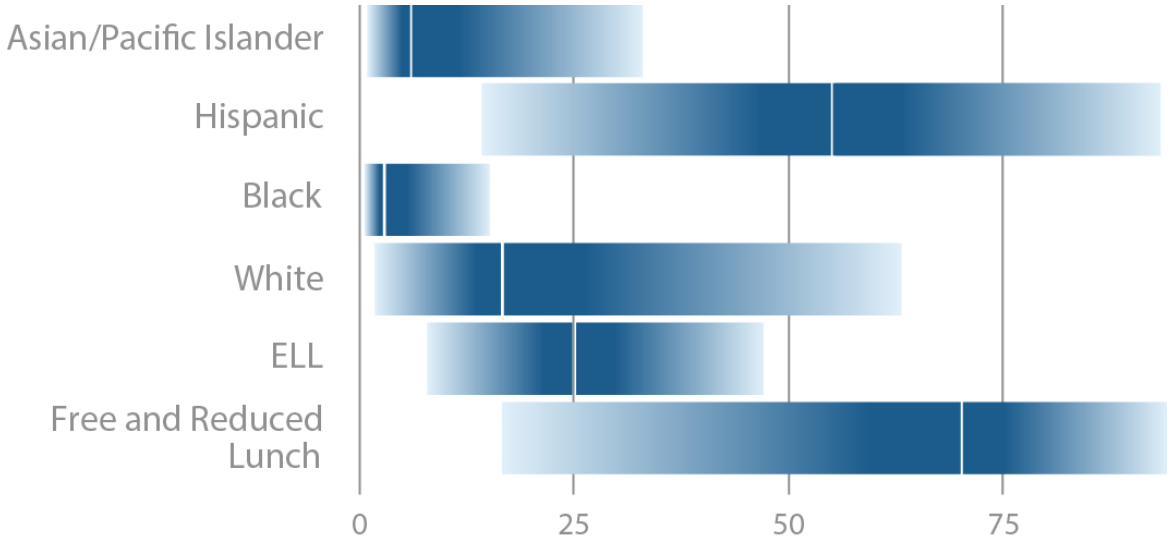

FIGURE 3. CHARACTERISTICS OF NEWSELA SCHOOLS IN CALIFORNIA

Conclusion

In this comparison study, schools that were active users of Newsela performed better on state reading assessments than schools that were not Newsela users. These results held true across all demographic subgroups.

Moderate evidence of positive impact

The study design used in this study meets the requirements for the level of Moderate evidence as defined in the Every Student Succeeds Act.

The positive average impact on schools with active Newsela users compared to non-users in conjunction with the positive correlation between product usage and outcomes among all users allows for a conclusion that the study provides Moderate evidence of impact. This leads to the recommendation of additional studies in other states and in districts using student-level data. It is important to note that analysis at the aggregate level does not allow inferences about impact on individual students. It shows that classes (grades) in schools that adopted Newsela and used it actively had better outcomes but it does not imply that requiring an individual student to take a large number of quizzes will lead to a higher test score. In order to better understand Newsela's impact on students, an assessment of student-level achievement and usage data is required.

Technical Details

Methodology

This report examines the effect of Newsela usage on reading performance in California schools during the 2015/16 school year, as measured by the Reading area of the Smarter Balanced ELA assessment. Reading area test results for each student are assigned to one of three achievement levels: Below Standard, Near Standard, or Above Standard. Test results are published at the class level, i.e. percentages of students in each of the three achievement levels among all students at a given grade level in a given school. For the purposes of evaluation, Newsela usage was aggregated at the class level in participating schools to match the available achievement data.

The analytical sample was established as follows:

- Student-level Newsela usage records were aggregated at the class level and classes with at least one quiz taken were selected
- Schools with most active use were selected (see Analysis section below for specific usage criteria)
- Newsela school data were linked to NCES demographic data
- Results from the 2016 Smarter Balanced test were collected for all schools in California and linked to Newsela school data.

The impact of Newsela was assessed using matched comparison group design whereby each Newsela class was matched to up to four classes of the same grade level in non-user schools with similar demographic characteristics and prior year test performance. Correlational analysis was performed using data from all Newsela classes with any greater-than-zero number of quizzes taken.

Data

The following table shows the baseline equivalence of the Newsela and comparison schools.

TABLE 1. BASELINE EQUIVALENCE OF THE NEWSELA AND COMPARISON SAMPLES

Characteristic	Newsela	Comparison	Pooled standard deviation	Difference as proportion of standard deviation
Average school size	570	595	316.01	-0.08
% students near or above standard in prior year (used as the pretest in the analysis.)	65.1	65.2	17.58	0
% economically disadvantaged	58.1	58.1	31.63	0
% Black	5.3	5.7	9.37	-0.05
% Hispanic	48.3	47.6	30.4	0.02
% English language learners	24.1	22.6	13.9	0.11
Community affluence indicator	3.8	3.8	1.21	0.06

Note. The Affluence Indicator uses an algorithm developed by MDR to rank the socioeconomic status of an institution. Data points ranging from specific variables to census data were incorporated into the formula.

ANALYSIS

The estimate of Newsela's impact was assessed by comparing proportions of students near or above standard in the most active Newsela classes and comparison classes, using a linear regression adjustment for differences in student demographics, school characteristics, and pretest scores. Two alternative criteria were used to select most active classes:

1) 24 or more quizzes per tested student
2) Top quintile of classes by the number of quizzes taken (more than 10 quizzes per student).

The first criterion is suggested by Newsela's own formative research but results in too small a sample (144 Newsela classes). Use of the top quintile is suggested by prior *Evidence as a Service* studies and results in a larger sample size (394 Newsela classes).

RESULTS

Results for the two samples described above are reported in Table 2 and Table 3, respectively. The estimates of the average impact are nearly identical in both samples, although the statistical significance is higher in the second sample due to the larger sample size. Estimates of grade-level effects are based on smaller numbers of observations and exhibit greater variability. Results reported in the body of the report were obtained from the larger second sample – Newsela classes in the top quintile by usage.

TABLE 2. EFFECT OF NEWSELA ON THE PROPORTION OF STUDENTS PERFORMING NEAR OR ABOVE STANDARD - MOST ACTIVE USERS

Grade	Effect	Standard error	*p value*	Comparison mean	Standard deviation of the mean	Effect size
All grades	1.35	0.88	0.13	63.8	17.4	0.08
4	0.55	2.00	0.78	62.6	18.0	-
5	2.27	1.64	0.17	63.5	17.3	0.13
6	1.45	1.65	0.38	62.7	16.9	-
7	2.98	2.76	0.28	65.4	16.7	-
8	-1.12	2.50	0.66	67.7	17.1	-

Note. Number of quizzes per student equals or exceeds 24 in the 144 Newsela classes. Comparison group is comprised of 560 matched non-user classes. Asterisks indicate significant differences between comparison and product (* $p < .2$, ** $p < .05$). Effect size not calculated for estimates that are not statistically significant.

TABLE 3. EFFECT OF NEWSELA ON THE PROPORTION OF STUDENTS PERFORMING NEAR OR ABOVE STANDARD - TOP QUINTILE USERS

Grade	Effect	Standard error	*p value*	Comparison mean	Standard deviation of the mean	Effect size
All grades	1.34	0.51	0.01	63.8	17.4	0.08
4	1.48	1.13	0.19	62.6	18.0	0.08
5	1.25	0.93	0.18	63.5	17.3	0.07
6	0.49	1.01	0.63	62.7	16.9	
7	2.74	1.43	0.06	65.4	16.7	0.16
8	1.67	1.47	0.26	67.7	17.1	

Note. A total of 394 Newsela classes included – top quintile by the number of quizzes per student (number of quizzes per student equals or exceeds 10). Comparison group is comprised of 1,538 matched non-user classes. Asterisks indicate significant differences between comparison and product (* $p < .2$, ** $p < .05$). Effect size not calculated for estimates that are not statistically significant.

Analysis of the correlation between Newsela use and student outcomes was performed using linear regression of the same outcome variable on student demographics, school characteristics, and pretest scores using all schools with non-zero Newsela usage. The appropriateness of a linear model (as opposed to a non-linear) is indicated by the results of a test performed using a generalized additive model. The usage indicator was obtained by dividing total articles accessed by total school enrollment.

In both impact and correlational analyses, average and grade-level effects were estimated. However, no statistically significant grade-level effects were established in the correlational analyses.

Note that these results only establish an association between class-level average usage and aggregate student outcomes. They cannot be used to predict the effect of an increase in Newsela usage on individual students or infer an optimal level of use.