Tier II Efficacy Study

NORTH CAROLINA MIDDLE SCHOOLS

Learn By Doing, Inc.

https://www.albert.io

Tier II Efficacy Study

One Year of Albert Practice, Higher Middle School Reading Scores

Introduction

This is a Tier II efficacy study under the Every Student Succeeds Act (ESSA), designed to meet What Works Clearinghouse (WWC) Group Design Standards. It examines the impact of using Albert, an online supplemental practice platform, on 7th and 8th grade student performance on the 2025-2026 North Carolina End-of-Grade (EOG) Reading assessment. The study compares students at three middle schools that used Albert during the 2025-2026 school year with students at three middle schools in the same district that did not.

About the Company

Albert, operated by Learn By Doing, Inc., offers rigorous supplemental question banks across core and test prep subjects for grades 3-12. All question bank items are standards-aligned and paired with detailed, narrative-style explanations. Teachers and students use Albert flexibly: for independent practice, diagnostics, differentiated instruction, and summative assessment.

Study Setting

The study took place in a public school district in central North Carolina serving roughly 17,700 students. The district's enrollment is about 48% White, 25% Hispanic, 20% Black, and 7% multiracial.

The six middle schools in the study all serve grades 6-8, and both the Albert and comparison groups include suburban and rural schools. No middle school in the district used Albert in 2024-2025. In 2025-2026, three middle schools began using Albert through a year-long pilot, with most ELA practice in Albert's Adaptive ELA and NC EOG Reading content. Students at the three Albert schools answered more than 250,000 ELA practice questions during the year.

Study Design

This quasi-experimental study used student-level data provided by the district. The outcome is each student's 2025-2026 NC EOG Reading scale score. The pretest is the student's prior-year EOG Reading score (Grade 6 for current 7th graders and Grade 7 for current 8th graders), taken in Spring 2025, before any middle school in the district used Albert. Because EOG scale scores

differ by grade, pretest and outcome scores were standardized within grade.

Two schools in the district that enroll 7th and 8th graders, an alternative school and a virtual school, were left out of the study because they are not traditional middle schools and each had fewer than 75 tested 7th and 8th graders (92 students in total). Students were also excluded if they had no prior-year score (205 students) or an unclear English learner status code (137 students). The analytic sample included 2,138 students: 1,151 at Albert schools and 987 at comparison schools.

To build an equivalent comparison group, each Albert student was matched to one comparison student without replacement. Students were matched exactly on grade level and English learner status, then paired on a propensity score estimated from the following:

- Prior-year EOG Reading score
- Grade level
- English learner status
- Sex
- Race/ethnicity

Each match had to fall within 0.2 standard deviations of the logit of the propensity score. The matched sample included 695 Albert students and 695 comparison students. The other 456 Albert students were left out of the analysis, either because no comparison student was close enough or because their grade and English learner group had more Albert students than comparison students.

The following tables show the demographics and prior achievement of the matched sample.

Group	Students	White	Black	Hispanic	Multiracial	Other	Female	English Learner
No Albert	695	46.0%	17.8%	26.3%	8.6%	1.2%	49.6%	11.7%
Albert	695	46.0%	17.8%	26.3%	8.6%	1.2%	48.8%	11.7%

Other includes Asian, American Indian, and Pacific Islander students, who were too few to analyze separately.

Group	Prior-Year Score Mean	Prior-Year Score Standard Deviation
No Albert	548.4	9.5
Albert	548.6	9.4

After matching, the difference between groups on the prior-year score was Hedges' g = 0.02, below the 0.05 threshold at which the What Works Clearinghouse considers groups equivalent. The analysis also adjusted for prior-year score and demographics.

Outcomes

Students at Albert schools scored higher on the 2026 EOG Reading assessment than matched students at comparison schools. The effect is positive and statistically significant for 7th and 8th graders combined. Effects for each grade were also positive (0.09 standard deviations in 7th grade and 0.11 in 8th grade) but were not statistically significant on their own.

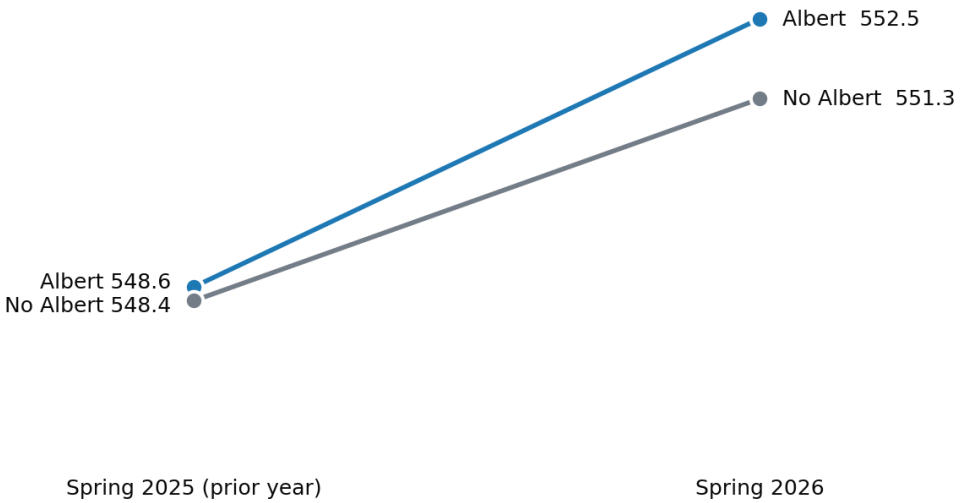
Effect (SD)	95% CI	p	Improvement Index	Students
+0.10	0.03 to 0.18	0.01	+4.1	1,390

Effects are regression-adjusted differences in standardized EOG Reading scores, converted to Hedges' g.

An effect of 0.10 standard deviations is about one scale-score point on the EOG Reading assessment. Relative to comparison students, who sit at the 50th percentile by definition, the average Albert student scored near the 54th percentile, a gain of 4.1 percentile points.

Same starting point, bigger reading gains with Albert

Average NC EOG Reading scale scores for matched 7th and 8th graders.
The groups started 0.2 points apart. One year later, Albert students were 1.2 points ahead.



Group	2026 EOG Reading Mean	2026 EOG Reading Standard Deviation	Gain from Prior Year Mean	Gain from Prior Year Standard Deviation
No Albert	551.3	10.0	2.9	6.0
Albert	552.5	9.6	3.8	6.0

Study Findings

The analysis found a statistically significant positive effect of attending an Albert school on EOG Reading scores. In a regression that adjusts for prior-year score, grade, sex, race/ethnicity, and English learner status, students at Albert schools scored 0.10 standard deviations higher than matched comparison students ($p = 0.01$).

Because students were grouped in six schools and Albert use was decided at the school level, standard errors were clustered by school. The result held up in other checks: leaving out any one school gave effects from 0.06 to 0.13, all statistically significant, and the same regression run on all 2,138 students, without matching, gave a similar effect of 0.09 ($p = 0.02$).

These results show that in a district where no middle school used Albert the year before, students at schools that adopted Albert made larger reading gains over one school year than similar students at schools that did not.

ESSA Tier II Criteria

This study meets the requirements for Tier II (Moderate Evidence) under the Every Student Succeeds Act (ESSA), based on the following criteria:

- **Well-Designed Quasi-Experiment:** Uses a matched comparison-group design intended to meet What Works Clearinghouse Group Design Standards, with propensity score matching to control for selection bias.
- **Demonstrated Baseline Equivalence:** The matched groups were equivalent on prior-year achievement (Hedges' g = 0.02), within What Works Clearinghouse thresholds.
- **Statistically Significant Positive Effect:** Shows a positive, statistically significant effect on an objective, state-standardized reading assessment, with significance tests that account for school clustering.
- **Large, Multi-Site Sample:** Includes 1,390 students across six middle schools, meeting the U.S. Department of Education's ESSA guidance of at least 350 students across more than one site.

Middle Schools in the Study

Contact us for more information, including a list of schools included in the analysis.

Appendix

Study Notes

This is a retroactive study: the district shared existing de-identified student records after the school year ended, and Albert had no role in assigning schools to use Albert.

Analysis

Propensity scores were estimated with logistic regression, and matching and outcome models were run in Python using the statsmodels package. Matching used nearest neighbor within exact grade and English learner groups, in random order, with a caliper of 0.2 standard deviations of the logit of the propensity score. The outcome model was an ordinary least squares regression of the standardized 2026 score on attending an Albert school, standardized prior-year score, grade, sex, race/ethnicity, and English learner status. Standard errors were clustered by school. Because clustered standard errors can understate uncertainty when there are only a few clusters, we also ran a wild cluster bootstrap with 9,999 replications and Webb six-point weights, clustering on school. The effect remained positive and statistically significant ($p = 0.025$).