

ESSA Evidence for Benchmark Advance and Benchmark Adelante

2021-2022 School Year Update

The last two school years, 2019–2020 and 2020–2021, have been difficult from the perspective of gathering program results data due to the cancelation of the 2020 state test assessments and the uneven 2020–2021 school year with in-person, remote, and hybrid instructional options. At Benchmark Education, we feel the coming school year will be different because we know more and have two years of experience to help us. Therefore, we are in the process of having an independent group study the efficacy of Benchmark Advance ©2022 in a large district in the Southeast Region of the U.S. We are expecting to have those results available in by fall of 2022.

Executive Summary

This report provides Every Student Succeeds Act (ESSA) evidence levels, as well as details and comments, on eight different studies where Benchmark Advance and Benchmark Adelante were used by students as their ELA core program. The studies occurred in six states, California, Colorado, Delaware, Michigan, Minnesota, and North Carolina. California is where Benchmark Advance and Benchmark Adelante are used by the largest number of districts due to the formal adoption process. Table 1 provides the title of the study, a short title used in the rest of the report, and the ESSA Evidence Level.

Table 1. Benchmark Advance and Benchmark Adelante Study Titles, Short Titles, and ESSA Evidence Level

| Benchmark Advance and Benchmark Adelante Study Title | Short Title | ESSA Evidence Level |
|---|-------------------------|---------------------|
| Smarter ELA/Literacy Results for Students in Indian River School District: Progress After Two Years of Benchmark Advance Instruction | Indian River Study 2018 | Moderate |
| Smarter ELA/Literacy Results for Students in Indian River School District and the State of Delaware After Initial Year of Benchmark Advance Instruction | Indian River Study | Moderate |
| Comparison of Dual Language Immersion and Monolingual Instruction Using Benchmark Advance and Benchmark Adelante | Chula Vista Study #1 | Moderate |
| Evidence Based Research in the Chula Vista (CA) Elementary School District on the Effectiveness of the Benchmark Advance and Benchmark Adelante Comprehensive Literacy Programs | MSA Study #1 | Promising |
| Percentage of Students in Met and Above Category More Than Doubled After Two Years of Benchmark Advance and Benchmark Adelante Instruction | Chula Vista Study #2 | Promising |
| Advancing to Proficiency: 2018 State Test Results for Students Using Benchmark Advance and Benchmark Adelante | Four State Study | Promising |
| California Benchmark Advance and Benchmark Adelante Districts CAASPP Results: After Initial Year of Instruction | 71 District Study | Promising |
| Comparison of Dual Language Immersion using Benchmark Adelante and English Only Instruction in Kindergarten and 1st Grade Classrooms in North Carolina | North Carolina Study #1 | Promising |

The number of participants ranges from below 100 participants (Chula Vista Study #1, Chula Vista Study #2, and MSA Study #1) to below 1,000 participants (Indian River Study and MSA Study #2) to greater than 500,000 participants (Four State Study).

Where statistical significance could be calculated, in three of the studies, it was in favor of the treatment groups. Where statistical significance could not be calculated, it was generally due to the use of state test aggregated results where an aggregated standard deviation was not provided or possible to obtain.

Where possible, the effect size was calculated using the Cohen's d formula¹. Effect sizes ranged from medium ($d = .55$ for 3rd to 4th grade and $d = .53$ for 4th to 5th grade in the Indian River Study) to large ($d = 1.08$ for 3rd to 5th grade in the Indian River Study 2018) to very large ($d = 1.22$ for English only and $d = 2.03$ for DLI in the Chula Vista Study #1) according to Cohen's convention for interpreting effect sizes. In John Hattie's² work on trying to determine what works best in education, he describes an effect size of: 0.5 as equivalent to a one-grade leap; 1.0 as equivalent to a two-grade leap; and above 0.4 as above average for the effects of all educational interventions he has studied. So, the effect sizes for these studies, where they could be calculated, are above average.

Introduction

This report provides a description of Every Student Succeeds Act (ESSA) evidence available, as of June 2019, for Benchmark Education's core elementary literacy and language programs, Benchmark Advance and Benchmark Adelante. These K-6 programs have been available for general use since the 2016–2017 school year and were used by one district during the 2015-2016 school year while participating in an early study. Bolding and highlighting was added to emphasize effect sizes, sample sizes, and when there was a control or comparison group(s).

Moderate ESSA Evidence

Currently, three of the Benchmark Advance and Benchmark Adelante studies are quasi-experimental, using existing groups, and demonstrating baseline equivalence for those groups, qualify for moderate ESSA evidence. In the first study, **Indian River Study 2018**, the salient information includes:

- The **treatment group** included students in the 3rd grade before Benchmark Advance was used as the core ELA program during 2015–2016 school year, and in the 5th grade, after two years of instruction with Benchmark Advance, during the 2017–2018 school year. There were **854 students in the 3rd grade** at the pre-test and **876 students in the 5th grade** at post-test.
- The **control group** included students in the same grades who attended schools in the other 12 school districts in Delaware who received ELA instruction with other materials and had no exposure to Benchmark Advance materials. There were **6,387 students in the 3rd** grade control group at the pre-test and **6,464 students in the 5th** grade control group at the post-test.

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

² <https://visible-learning.org/hattie-ranking-influences-effect-sizes-learning-achievement/>

- Using Delaware’s Smarter ELA/Literacy Assessment as the outcome measure and using confidence intervals to evaluate the difference between the groups, the treatment and control groups at the pre-test, 2016 administration, had confidence intervals that overlapped.
- At post-test, 2018 administration, there was no overlap of the confidence intervals, indicating a **statistical significance difference** between the Indian River students and students in the other 12 school districts in the 3rd to the 5th grade group after the two years of Benchmark Advance instruction, **in favor of the treatment group**.
- The size of the difference between the Indian River students and the other 12 school districts was calculated using the Cohen’s d effect size³. For the all student treatment group, the effect size was **$d = 1.08$ for the 3rd to 5th grade students**. For the all student control group, the effect size was $d = .89$ for the 3rd to 5th grade students. For the nine subgroups, six of the Indian River subgroups were similar at pre-test with overlapping confidence intervals and all Indian River subgroups had larger effect sizes than the other 12 school district comparison group.

In the second study, **Indian River Study**, the salient information includes:

- The **treatment groups** included students in the 3rd or 4th grade during 2015–2016 school year, before new instruction, and in the 4th or 5th grade during the 2016–2017 school year, when the Benchmark Advance instruction occurred. There were **854 students in the 3rd to 4th grade** at the pre-test and **853 students in the 4th to 5th grade** at post-test.
- The **control groups** included students in the same grades, 3rd to 4th and 4th to 5th, who attended schools in the other 14 school districts in Delaware and received ELA instruction with other materials. There were **8,443 students in the 3rd to 4th** grade control group at the pre-test and **8,358 students in the 4th to 5th grade** control group at the post-test.
- Using Delaware’s Smarter ELA/Literacy Assessment as the outcome measure and using confidence intervals to evaluate the difference between the groups, the treatment and control groups at the pre-test, 2016 administration, had confidence intervals that overlapped.
- At post-test, 2017 administration, there was no overlap of the confidence intervals, indicating a **statistical significance difference** between the Indian River students and students in the other 14 school districts in both the 3rd to 4th and the 4th to 5th grade groups after the initial year of Benchmark Advance instruction, **in favor of the treatment groups**.
- The size of the difference between the Indian River students and the other 14 school districts was calculated using the Cohen’s d effect size. For the treatment groups effect sizes were **$d = 0.55$ for the 3rd to 4th grade students and $d = 0.53$ for the 4th to 5th grade students**. For the control groups effect sizes were $d = 0.45$ for the 3rd to 4th grade students and $d = 0.40$ for the 4th to 5th grade students.

³ Cohen’s d is defined as the difference between two means divided by the standard deviation of the pooled groups or of the control group alone. For these studies, the pooled standard deviation was used. According to Cohen (1988), an effect size of 0.2 is small, 0.5 is medium, and 0.8 is large. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second edition). New York, NY: Psychology Press, Taylor & Francis Group.

In the third study, **Chula Vista Study #1**, the salient information includes:

- This study was an extension of the original Chula Vista study (MSA Study #1), with additional information from the district to identify students in the different types of instruction and includes two years of instruction using Benchmark Advance and/or Benchmark Adelante.
- **Three groups** were considered in this study, students receiving **Dual Language Immersion (DLI) instruction (n = 42)**, student receiving **English only instruction (n = 31)**, and all students in the **Chula Vista Elementary School District**.
- There was **no statistically significant difference between the DLI and English only groups at the pre-test**, 2015 administration of the California Assessment of Student Performance and Progress (CAASPP). Aggregated CAASPP results for the Chula Vista Elementary School District made statistical comparison impossible.
- There was a **statistically significant difference** between these groups of students at the post-test with an **effect size of $d = 1.22$ for the English only group and $d = 2.03$ for the DLI group**.

Promising ESSA Evidence

Currently, there are five Benchmark Advance and Benchmark Adelante studies that fit into the promising ESSA evidence. Two studies are related to the Chula Vista Elementary School District school participating in the initial Benchmark Advance and Benchmark Adelante study, designed and conducted by Main Street Academix (MSA), two studies are related to the many districts in California, Colorado, Michigan, and Minnesota who adopted Benchmark Advance and Benchmark Adelante since the 2016–2017 school year, and the last study, North Carolina Study #1, is related to the comparison of Spanish instruction using Benchmark Adelante to English instruction in kindergarten and 1st grade classrooms in North Carolina.

For the first Chula Vista report, **MSA Study #1**, the salient information includes:

- The participants in this study are from **School A Elementary School in Chula Vista** and are compared to the growth of students in the **other elementary schools in Chula Vista** and **students across the state of California**. During the 2015–2016 school year there were **98 students in the 4th grade and 93 students in the 5th grade** that received Benchmark Advance and/or Benchmark Adelante instruction in School A.
- Using the CAASPP as the outcome measure, from pre-test (2015 CAASPP administration) to post-test (2016 CAASPP administration) the 4th grade students went from 28 percent of students in the Met or Exceeded performance categories on the CAASPP to 51 percent of students in Met or Exceeded, **a 23 percentage point increase** after receiving Benchmark Advance and/or Benchmark Adelante instruction. The **comparison groups had an 11 percentage point increase for students in other Chula Vista elementary schools and a 6 percentage point increase for students across the state of California**.
- For the 5th grade students, School A students went from 31 percent to 70 percent in the Met or Exceeded performance categories, **a 39 percentage point increase**. The 5th grade comparison groups had an **11 percentage point increase for student in other Chula Vista elementary schools and a 9 percentage point increase for students across the state of California**.

The second Chula Vista report, **Chula Vista Study #2**, provides results for the second academic year of instruction with Benchmark Advance and/or Benchmark Adelante, during the 2016–2017 school year.

- Students included in this report were in 3rd grade prior to Benchmark Advance and/or Benchmark Adelante instruction in the 2014–2015 school year and in the 5th grade after two years of Benchmark Advance and/or Benchmark Adelante instruction in the 2016–2017 school year.
- These **students started at 28 percent (n = 95)** in the Met or Exceeded performance categories on the 2015 CAASPP administration, **grew to 51 percent (n = 98)** in the Met or Exceeded after 1 year of instruction, and **reached 62 percent (n = 85)** in the Met or Exceeded performance categories by the 2017 CAASPP administration.

In the **Four State Study**, the salient information includes:

- By the 2017–2018 school year, **over 180 school districts in four states** used Benchmark Advance and/or Benchmark Adelante as the core English/Spanish Language Arts curriculum in kindergarten through either the 5th or 6th grade across all or almost all elementary schools in the districts.
- The districts using the Benchmark Education Company programs, identified as **BEC Districts**, **represent over 500,000 students** receiving ELA/SLA instruction using one or both programs. The districts who were not using Benchmark Education Company programs, identified as Non-BEC Districts, included over 1.6 million student receiving instruction using other materials.
- The state tests in California, Colorado, Michigan, and Minnesota do not report an aggregated standard deviation so calculating statistical significance and effect sizes are not possible. This study looks at the movement of students into the proficient category on the state tests from the four states. The growth of the percentage of **students in the proficient categories increased by 3.53 percentage points for BEC Districts** and **1.39 percentage points for Non-BEC Districts** from the 2017 to 2018 administrations of the state tests.
- The subgroups included in this study, including the BEC Districts and Non-BEC Districts growth in the percentage of student in the proficient categories, include students who are: Identified as **English Learners (BEC Districts up 2.15, Non-BEC Districts up 0.87)**; **Economically disadvantaged (BEC Districts up 4.19, Non-BEC Districts up 2.34)**; **Students with Disabilities (BEC Districts up 2.51, Non-BEC Districts up 1.63)**; **Female (BEC Districts up 3.47, Non-BEC Districts up 1.20)**; **Male (BEC Districts up 3.58, Non-BEC Districts up 1.68)**; **American Indian/Alaska Native (BEC Districts up 3.58, Non-BEC Districts up 1.45)**; **Asian (BEC Districts up 2.11, Non-BEC Districts up 1.39)**; **Black/African American (BEC Districts up 3.13, Non-BEC Districts up 0.54)**; **Hispanic/Latino (BEC Districts up 3.76, Non-BEC Districts up 3.07)**; **Two or more races (BEC Districts up 1.85, Non-BEC Districts up 1.69)**; and **White (BEC Districts up 1.89, Non-BEC Districts up 0.33)**;

In the **71 District Study**, the salient information includes:

- During the initial year Benchmark Advance and Benchmark Adelante were available for adoption in California, 71 districts committed to a long-term relationship with Benchmark Education. The students in the 71 districts represents about 6.5 percent of students enrolled in California elementary schools.
- The **3rd to 4th grade group** in the 71 districts contained **30,184 students tested on the 2016 CAASPP** and **29,950 students tested on the 2017 CAASPP**. The 3rd to 4th grade comparison group in the other California districts had **402,855 students tested on the 2016 CAASPP** and **401,296 students tested on the 2017 CAASPP**.
- The **4th to 5th grade group** in the 71 districts contained **31,069 students tested on the 2016 CAASPP** and **30,871 students tested on the 2017 CAASPP**. The 4th to 5th grade comparison group in the other California districts had **417,999 students tested on the 2016 CAASPP** and **415,791 students tested on the 2017 CAASPP**.
- The CAASPP does not report an aggregated standard deviation so calculating statistical significance or effect sizes is not possible, but this study looks at the movement of students into the Met and Above category on the CAASPP. **For the 3rd to 4th grade**, the 71 districts group started at 42.77 percent and move to 46.44 percent, **an increase of 3.67 percentage points** and the other California districts started at 42.45 percent and moved to 45.03, an increase of 2.58 percentage points. **For the 4th to 5th grade**, the 71 districts group started at 44.52 percent and move to 48.24 percent, **an increase of 3.72 percentage points** and the other California districts started at 43.74 percent and moved to 46.48, an increase of 2.74 percentage points. There is about one percentage point difference in growth for both grade groups, favoring students in the 71 districts.
- This California report also includes the analysis of the following **demographic subgroups: economically disadvantaged; not economically disadvantaged; Ethnicities, including Asian, Black or African American, Filipino, Hispanic or Latino, Two or more races, and White; and Gender, female and male.**

In the **North Carolina Study #1**, the salient information includes:

- The focus of this study is to **compare the results of instruction** using Benchmark Adelante in a **Dual Language Immersion (DLI) classroom with English-Only (EO) classrooms using English materials**, but not Benchmark Advance, in kindergarten and 1st grade classrooms in three schools in a district located in the eastern part of North Carolina.
- A total of 127 students, **64 in kindergarten and 63 in 1st grade received instruction using Benchmark Adelante** and assessments in Spanish (IDEL). There were **590 students, 283 in kindergarten and 307 in 1st grade, who received English instruction** and assessments in English (DIBLES). IDEL and DIBELS measures that assessed the same basic early literacy skills were compared.
- Looking at the effect sizes from initial to final scores, measuring **phonemic awareness in kindergarten (using FSF and PSF) for DLI students $d = 2.44$ and for EO students $d = 1.59$;**

measuring the **alphabetic principal in kindergarten (using FPS and NWF) for DLI students $d = 1.42$** and for EO students $d = 0.95$; measuring the **alphabetic principal in 1st grade (using FPS and NWF) for DLI students $d = 1.66$** and for EO students $d = 1.30$; and measuring **accurate and fluent reading of connect text in 1st grade (using FLO and DORF) for DLI students $d = 0.89$** and for EO students $d = 0.75$. Scores are **statistically significantly different from the initial to final** scores across the measures above.

- This report also includes the analysis of the following demographic subgroups include, among other categories, **English Learners** or not and **home language of English or Spanish**.

Summary

Benchmark Education will continue to look for research opportunities, including, but not limited to random control trials. Instead of putting all eggs in one basket, we will continue to look at almost any research opportunity as a good one, including working with: districts who want to understand how students are benefitting from *Benchmark Advance* instruction; classroom teachers who are in graduate school and interested in doing action research; state test results; and organizations and universities who are applying to federal grants and need partners. While the evidence structure provided by the “Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments”⁴ clarifies the definitions provided in section 8101(21)(A) of ESSA, it does not suggest the restriction to only strong evidence.

⁴ <https://www2.ed.gov/policy/elsec/leg/essa/guidanceuseinvestment.pdf>