# Report on the Supplemental Empirical Analyses of the Alternate English Language Proficiency Assessments for California

**Contents**

[Background and Purpose 3](#_Toc129934369)

[Data Sources and Student Sample 4](#_Toc129934370)

[Sample Limitations and Considerations 8](#_Toc129934371)

[Analytic Methods 9](#_Toc129934372)

[Findings 11](#_Toc129934373)

[EL Students’ Performance on California Alternate Assessments for English Language Arts and Alternate English Language Proficiency Assessments for California 11](#_Toc129934374)

[Results from Descriptive Box Plot Analyses 12](#_Toc129934375)

[Results from Logistic Regression Analysis 15](#_Toc129934376)

[Results from Decision Consistency Analysis 18](#_Toc129934377)

[Results from Exploratory Analysis Setting California Alternate Assessment for English Language Arts Target at a Level Other than “Proficient” 19](#_Toc129934378)

[Outreach to Interest Holders 24](#_Toc129934379)

[Summary and Considerations 25](#_Toc129934380)

[Recommendations for Future Analysis 25](#_Toc129934381)

[Appendix A: Additional Information About Coaching Support from the Regional Educational Laboratory, West 28](#_Toc129934382)

## Background and Purpose

State and federal laws require that all students whose primary language is a language other than English be assessed for English language proficiency (ELP). The legal basis for requiring ELP testing is that all students have the right to an equal and appropriate education, and any English language limitations left unidentified and/or unaddressed could preclude a student from accessing that right. This legal basis comes from *California Code of Regulations*, Title 5 (5 *CCR*), Section 11518; California *Education Code* (*EC*) sections 313 and 60810; and Titles I and Ill of the federal Every Student Succeeds Act, which reauthorized the Elementary and Secondary Education Act.

Between 2019 and 2021, the California Department of Education (CDE) and its testing contractor, ETS, developed the Alternate English Language Proficiency Assessments of California (Alternate ELPAC) to be used as a statewide, standardized assessment of ELP for English learner (EL) and potential EL students with the most significant cognitive disabilities.

The first administration of the Alternate ELPAC in the 2021–22 school year was an operational field test that included all students eligible to participate. For that administration, the test developer, ETS, applied threshold scores developed in a standard setting process involving teachers reviewing Alternate ELPAC operational field test results. These threshold scores for the Initial Alternate ELPAC and the Summative Alternate ELPAC were approved by the State Board of Education (SBE) in May 2022. The CDE planned to conduct a follow-up analysis to further evaluate and validate the thresholds.

Then in fall 2022, the Regional Educational Laboratory West (REL West) began providing the CDE with coaching and support around data analysis for this study as quality assurance of the results. Specifically, REL West supported the CDE's supplemental empirical analyses to examine the relationship of EL performance on the Alternate ELPAC relative to their performance on the California Alternate Assessments (CAA) for English Language Arts (ELA). These analyses are described in further detail in the Analytic Methods section.

The CDE has conducted a series of analyses to evaluate and validate the state’s current recommended thresholds for reclassifying EL students with the most significant cognitive disabilities.

This report and the analyses therein focus on the performance levels and thresholds on the Summative Alternate ELPAC, California’s new statewide assessment of ELP for EL students with the most significant cognitive disabilities as determined by their individualized education program team. The CDE conducted these analyses with coaching support from the REL West; REL West also advised the CDE on the content of this summary report.

The CDE has provided interim guidance to local educational agencies (LEAs) recommending that EL students with the most significant cognitive disabilities be eligible for reclassification when they earn an Overall Performance Level (PL) of 3 (out of three possible levels) on the Summative Alternate ELPAC. The findings of these supplemental empirical analyses support the continued use of an overall PL 3 to meet Criterion 1 for reclassification purposes. For more information on the reclassification process, including the four reclassification criteria, please visit the CDE Reclassification web page at <https://www.cde.ca.gov/sp/el/rd/>.

This recommendation is based on a series of analyses conducted by the CDE using industry-standard methods to identify and evaluate reclassification thresholds using scores on both ELP and academic content assessments. It is also informed by political and logistical considerations for state and local educators and leaders who will be tasked with implementing new standardized reclassification procedures for EL students with the most significant cognitive disabilities now and into the future.

The overall purpose of this study was to conduct supplemental analyses of test-based performance data to evaluate the placement of the thresholds as it pertains to the interim guidance for meeting Criterion 1 for reclassification. These analyses were conducted within a larger context, described above, with several notable features:

* The state has only recently transitioned (starting in the 2021–22) to a statewide alternate assessment for ELP for EL students with the most significant cognitive disabilities.
* Only one year of student performance on the Alternate ELPAC (from the 2021–22 operational field test) was available at the time these analyses were requested.

Based on this context, these results should be treated as one additional piece of evidence in the state’s ongoing validation and evaluation of its reclassification threshold score.

## Data Sources and Student Sample

Data for this study are based on the results of the Alternate ELPAC operational field test, which took place during the 2021–22 school year. The CDE gathered data from the Alternate ELPAC section of the 2021–22 ELPAC data file and from the CAA for ELA section of the 2022 California Assessment of Student Performance and Progress (CAASPP) data file. The two data files were merged based on each student’s statewide student identification number (SSID) and records were kept only if the observation had valid test scores from either source file and had matching enrollment and test grades for grades three through eight and eleven. This resulted in a sample comprised of 32,389 total student observations with valid CAA for ELA test scores. Among these 32,389 students, 8,077 were EL students who also had valid tests scores from the Summative Alternate ELPAC (meaning EL students with valid CAA for ELA scores but lacking valid Alternate ELPAC scores are also excluded from the analyses).

Table 1 summarizes the study sample by English language proficiency status, primary disability, and grade-level. The table shows the following:

1. Among students who had valid test scores on the CAA for ELA, three quarters (24,312 out of 32,389 students) were not currently EL students, meaning they were English only (EO), initially fluent English proficient (IFEP), or reclassified fluent English proficient (RFEP). A small number of students (fewer than 10) were also labeled as “to be determined” because they had not yet completed testing to determine if they were at Alternate ELPAC level 1, 2, 3, or IFEP. Students who could not be classified into one of these categories (EO, EL, IFEP, RFEP, or TBD) were excluded from the analytic sample (816 students in all). These students are included in the descriptive box plot analyses, but not in the logistic regression or decision consistency analyses.
2. Among students who had valid test scores on the CAA for ELA, roughly one in four also had valid test scores on the Alternate ELPAC (8,077 out of 32,389 students). These students are included in all three analyses.
3. The distribution of non-EL students was fairly even across grades (between 3,300 and 3,600 students per tested grade, ranging from 13.8 to 14.9 percent of the overall sample), whereas the distribution of EL students decreased as grades increased, from a high of 1,365 participants in grade three (16.9 percent of the overall sample) to a low of 792 participants in grade eleven (9.8 percent of the overall sample). This distribution of the EL sample across grade-levels is typical in the general EL student population as well and reflects the goal of having EL students transition to RFEP status as a result of receiving EL programs and services.
4. Although the most common disability categorizations in both groups are autism and intellectual disabilities—a trend that aligns with other national findings and surveys about students who participate in alternate assessments[[1]](#footnote-2)—the relative ranking of these two disability groups is reversed between EL students and non-EL students. Specifically, autism is the most common disability categorization for the non-EL students (47.3 percent of students fall in this category), followed by intellectual disabilities (2.5 percent), whereas intellectual disabilities are most common among EL students (44.34 percent), followed by autism (36.58 percent).

Overall, Table 1 suggests that the EL students in the analytic sample differ from the non-EL students in terms of their distribution across grades and their prevalence of autism and intellectual disabilities. This information is particularly relevant for the results from the descriptive box plot method, which includes CAA for ELA results for EO, IFEP, and RFEP students as a point of comparison for the Alternate ELPAC students. It is also for the exploratory logistic regression analyses using the mean CAA for ELA scores of EO students as a reference point (not shown in this report).

Table 1: Number and Percentage of Students in the Study Sample by English Language Proficiency Status, Primary Disability, and Grade Level

| Categories:Primary Disability and Grade Level | Number of Students with Valid CAA ELA Scores | Percent Students with Valid CAA ELA Scores | Number of EO, IFEP, and RFEP Students with Valid CAA ELA Scores | Percent of EO, IFEP, and RFEP Students with Valid CAA ELA Scores | Number of EL Students with Valid Scores on Both Alternate ELPAC and CAA for ELA  | Percent EL Students with Valid Scores on Both Alternate ELPAC and CAA for ELA  |
| --- | --- | --- | --- | --- | --- | --- |
| **Total** | 32,389 | – | 24,312 | – | 8,077 | – |
| Autism | 14,441 | 44.6 | 11,489 | 47.3 | 2,952 | 36.6 |
| Intellectual Disability | 11,483 | 35.5 | 7,904 | 32.5 | 3,579 | 44.3 |
| Other Disability | 6,452 | 19.9 | 4,912 | 20.2 | 1,540 | 19.1 |
| Missing Primary Disability | 13 | 0.0 | 7 | 0.0 | 6 | 0.1 |
| Grade Three | 4,730 | 14.6 | 3,365 | 13.8 | 1,365 | 16.9 |
| Grade Four | 4,771 | 14.7 | 3,483 | 14.3 | 1,288 | 16.0 |
| Grade Five | 4,644 | 14.3 | 3,383 | 13.9 | 1,261 | 15.6 |
| Grade Six | 4,724 | 14.6 | 3,567 | 14.7 | 1,157 | 14.3 |
| Grade Seven | 4,651 | 14.4 | 3,471 | 14.3 | 1,180 | 14.6 |
| Grade Eight | 4,646 | 14.3 | 3,612 | 14.9 | 1,034 | 12.8 |
| Grade Eleven | 4,223 | 13.0 | 3,431 | 14.1 | 792 | 9.8 |

**NOTE:** The study sample excludes 816 non-EL students who were missing a flag for EO, IFEP, or RFEP status. It also excludes EL students who had valid CAA for ELA scores but lacked valid Alternate ELPAC scores.

Table 2 summarizes the EL student study sample relative to all students who participated in the Alternate ELPAC field test. Because the Alternate ELPAC is administered to all grades from kindergarten through grade twelve (K–12), this table illustrates representativeness of the study sample compared to the full K–12 population of EL students with the most significant cognitive disabilities.

Table 2 shows the following:

* Out of 19,697 students who were eligible to participate in the Alternate ELPAC, only about 40 percent (8,077 students) were in the CAA for ELA tested grades (three through eight and eleven) and had valid scores on both assessments. This is likely because, as with the general EL student population, a large number of EL students with the most significant cognitive disabilities are in the earliest grades (K through two), which do not participate in large-scale summative academic content assessments.
* Only about 85 percent of eligible students participated in the Alternate ELPAC operational field test (16,669 students out of 19,697), but participation rates were fairly consistent across home language groups and grades.
* The distribution of home languages and grade-levels in the analytic sample was comparable to the total eligible sample, other than the fact that grades three through eight and eleven make up larger percentages of the analytic sample due to the study design.

Table 2: Comparison of the Full Alternate ELPAC Population and the Study Sample by Home Language and Grade Level

| Categories:Home Language and Grade Level | Number Eligible for Alternate ELPAC | PercentAll Eligible Alternate ELPAC | Number Tested with Valid Alternate ELPAC Scores | PercentAll Eligible Alternate ELPAC | Number Matched Valid Alternate ELPAC and CAA for ELA Scores | PercentAll Matched Valid Alternate ELPAC and CAA for ELA Scores |
| --- | --- | --- | --- | --- | --- | --- |
| **Total** | 19,697 | – | 16,669 | 84.6 | 8,077 | – |
| Spanish | 15,172 | 77.0 | 12,884 | 84.9 | 6,325 | 78.3 |
| Chinese (Cantonese) | 293 | 1.5 | 217 | 74.1 | 112 | 1.4 |
| Chinese (Mandarin) | 298 | 1.5 | 257 | 86.2 | 129 | 1.6 |
| Other Home language | 3,934 | 20.0 | 3,311 | 84.2 | 1,511 | 18.7 |
| Grade Three | 1,664 | 8.5 | 1,475 | 88.6 | 1,365 | 16.9 |
| Grade Four | 1,589 | 8.1 | 1,388 | 87.4 | 1,288 | 16.0 |
| Grade Five | 1,546 | 7.9 | 1,368 | 88.5 | 1,261 | 15.6 |
| Grade Six | 1,464 | 7.4 | 1,252 | 85.5 | 1,157 | 14.3 |
| Grade Seven | 1,497 | 7.6 | 1,258 | 84.0 | 1,180 | 14.6 |
| Grade Eight | 1,354 | 6.9 | 1,114 | 82.3 | 1,034 | 12.8 |
| Grade Eleven | 1,165 | 5.9 | 896 | 76.9 | 792 | 9.8 |

Table 2 suggests that the analytic sample of EL students used in this study is not meaningfully different from the population of EL students with the most significant cognitive disabilities in terms of their home languages, disabilities, or distribution across grades. That said, there remain several limitations to this study’s sample, which are discussed in the next section.

### **Sample Limitations and Considerations**

Because the methods used scores from both the Alternate ELPAC and the CAA for ELA, they are necessarily limited to the tested grades for the CAA for ELA (grades three through eight, and grade eleven). Since the Alternate ELPAC is administered to all grades, K–12, and because the EL population is generally larger in the lower grades, this limitation means that these analyses are based on less than half of the student population who participated in the Alternate ELPAC operational field test (8,077 students out of 16,669 tested). This report is being submitted to the SBE along with results from a contrasting groups study, conducted by ETS, that uses teacher judgment about students’ language acquisition status, and allows for the inclusion of all tested grades.

Additionally, it is significant that these analyses are based on a single year of data, which was also the first year the Alternate ELPAC was administered statewide. Because of the policy shift from locally determined instruments to a statewide one, it is likely that the performance distribution of students in the first year may differ from future years in important ways. For example:

* The operational field test census sample may include a “backlog” of students who would have been reclassified sooner had the Summative Alternate ELPAC been in place instead of locally-determined instruments.
* The operational field test marked the first time that the vast majority of teachers and students participated in the administration of the Alternate ELPAC, and both groups will likely become more familiar and confident with the assessment as time goes on.

To the extent that this is the case, future score distributions on the assessment may not look similar to the distributions observed in the first year. For this reason, the CDE recommends several important follow-up analyses to be conducted in future years as more data and information become available.

## Analytic Methods

For this supplemental study, REL West coached the CDE to apply three analytic methods that were developed by the US Department of Education as part of a National Evaluation of Title III (see: Cook, [Linquanti, Chinen & Jung [2012]](https://www2.ed.gov/rschstat/eval/title-iii/implementation-supplemental-report.pdf)). These methods were designed specifically to support states to empirically identify appropriate thresholds for ELP, and they have since been used by many states—including California[[2]](#footnote-3)—to identify and evaluate thresholds on both general and alternate[[3]](#footnote-4) summative ELP assessments used for accountability.

All three methods are based on the facts that (1) English learner status is intended to address language barriers that may affect students’ academic learning and performance, and (2) monolingual English students demonstrate a range of performance on assessments of academic content like ELA and mathematics—they are not, in other words, uniformly proficient in these areas.

Accordingly, the methods’ developers argue, EL students should not be expected to achieve academic proficiency as a condition of exiting EL status. Rather, researchers and decision-makers should look for evidence that students’ ELP performance no longer strongly predicts or correlates with their academic content performance. This lack of a strong relationship serves as evidence that students’ language proficiency is no longer affecting or driving their academic performance; a sign that the removal of the language services and supports of EL status are less likely to be harmful for students.

In the words of the authors:

“…there is a point at which EL students have sufficient English language skills to adequately function in English on content assessments; accordingly, there should be observable decreases in the relationship between the two assessments. At or beyond this point is where the ELP performance standard might be considered...” (Cook et al., 2012, p. 8). The methods operationalize this concept by generally seeking the point at which students’ ELP performance predicts their academic content achievement (proficient, or not proficient) at the level of chance. In other words, if students with the same level of ELP have a 50-50 chance of being proficient on an academic content assessment, this suggests that their ELP is not making a difference in their academic achievement. A stronger level of prediction in either direction would suggest that their ELP is driving their academic content achievement (so removal of services could be harmful).

Accordingly, all three methods use scores from both an ELP assessment and academic content assessments. For the current study, the CDE used scores from the CAA for ELA as the academic content assessment in the analysis, alongside the Alternate ELPAC. Each of the three methods seeks to identify the point of separation between performance in language and academic content in a different way:

1. **Descriptive box plot analysis** examines the distribution of overall scale scores on the CAA for ELA for students at each performance level on the Summative Alternate ELPAC, for each applicable grade level. This analysis also includes the distribution of overall scale scores on CAA for ELA for RFEP, IFEP, and m EO students for comparison. The purpose of the analysis is to identify an Alternate ELPAC performance level where a) EL students have an equal likelihood of scoring above or below the CAA for ELA Level 3 threshold (Understanding); or b) EL student’s score distribution on the CAA for ELA is very similar to that of EO students statewide.
2. **Logistic regression analysis** estimates the probability of reaching Level 3 (Understanding) on CAA for ELA for each Alternate ELPAC overall scale score. This approach helps to identify the Alternate ELPAC overall scale score range in which EL students have a probability equal to or greater than 50 percent of attaining that standard on the CAA for ELA.
3. **Decision consistency analysis** analyzes Alternate ELPAC and CAA for ELA proficient-level categorizations and optimizes consistent categorization of EL students at or above the current CAA for ELA threshold score for Level 3 (Understanding). For this analysis, consistency is defined as being proficient either on both the CAA for ELA and Alternate ELPAC, or on neither. Because states are expected to set their reclassification thresholds at a place where students have developed the linguistic knowledge and skills they need to function independently in classrooms where English is the language of instruction, a cut score that results in large numbers of “inconsistent” classification scenarios (i.e., students who are considered proficient in ELP but who overwhelmingly fail to reach proficiency in ELA, or who are proficient in ELA but still classified as English learners) would suggest that the ELP threshold is not set in the right place. Accordingly, this analysis determines the Alternate ELPAC overall scale score range that maximizes the amount of agreement between achieving Alternate ELPAC proficiency and CAA for ELA proficiency.

The CDE applied all three methods separately to each grade-level of the Alternate ELPAC for students in grades three through eight, and eleven. Additionally, the CDE separately ran models for all methods and grade levels for the two largest disability groups in the student sample – students with autism, and students with intellectual disabilities (see Table 1). The CDE also ran the logistic regression and decision consistency models using three different performance thresholds on the CAA for ELA as a performance target. In all, these different approaches resulted in nearly 100 different results to consider collectively in making the recommendations contained in this report.

## Findings

This section provides selected data visualizations for each method, as well as summary information about the overall results for each method, including the analyses conducted for specific student groups or using different CAA for ELA scores as outcomes. In other words, the grades and groups that are not shown have results that point to the same achievement level as those shown in the report.

### **EL Students’ Performance on California Alternate Assessments for English Language Arts and Alternate English Language Proficiency Assessments for California**

Table 3 shows a cross tabulation of the number of students at each performance level on the Alternate ELPAC (under “Row Total”) and achievement level on the CAA for ELA (under “Column Total”). The Table shows that students generally scored at a higher proficiency level on the Alternate ELPAC than they did on the CAA for ELA. The number of students at each of the three Alternate ELPAC performance levels are more evenly distributed than the number of students at each of the CAA for ELA proficiency levels, with more students scoring at Alternate ELPAC Levels 2 and 3 (3,345 and 2,711, respectively) compared to Level 1.

On the CAA for ELA, by contrast, half of all students (4,032) scored at the lowest achievement level, and approximately one in seven students scored at the highest level. This imbalance is further evident in the interior cells of the table, which show the number of students who earned each combination of performance levels across the two assessments (e.g., Alternate ELPAC Level 1 and CAA for ELA Level 2). While most students who scored at Alternate ELPAC Level 1 also scored at CAA for ELA Level 1 (1,821 out of 2,021 students), the Table shows that students who scored at Alternate ELPAC Level 2 were most likely to score at CAA for ELA Level 1 (1,887 out of 3,345). Similarly, students who scored at Alternate ELPAC Level 3 were most likely to score at CAA for ELA Level 2. In other words, students generally performed at one level lower on the CAA for ELA than they did on the Alternate ELPAC.

Table 3: Alternate ELPAC Performance Level by CAA for ELA Achievement Level

| **Alternate ELPAC Performance** | **CAA for ELA Level 1** | **CAA for ELA Level 2** | **CAA for ELA Level 3** | **Row** **Total** |
| --- | --- | --- | --- | --- |
| **Level 1** | 1,821 | 192 | 8 | 2,021 |
| **Level 2** | 1,887 | 1,262 | 196 | 3,345 |
| **Level 3** | 324 | 1,448 | 939 | 2,711 |
| **Column Total** | 4,032 | 2,902 | 1,143 | ***8,077*** |

### **Results from Descriptive Box Plot Analyses**

Figures 1 through 3 display box plots showing the CAA for ELA score distributions for all EL, EO, IFEP, and RFEP students with valid scores in grades four, seven, and eleven, respectively. Score distributions for current EL students are further broken out by overall performance level on the Alternate ELPAC (1, 2, or 3). Results for the other tested grades, as well as for students with autism and students with intellectual disabilities, align with the results shown here.

The figures show three important results:

1. As expected, the mean and distribution of CAA for ELA scores are progressively higher for students at each Alternate ELPAC performance level. In other words, students who do better on the Alternate ELPAC, also generally do better on the CAA for ELA.
2. With the exception of Grade 3, none of the distributions reflect a situation where half the students score above, and half below, the proficiency cut score on the CAA for ELA, the students at Alternate ELPAC level 3 are consistently closest to this goal, particularly relative to the other groups. Across all grades, between a half and quarter of the students at Alternate ELPAC Level 3 score at or above proficient on the CAA for ELA, and the proportion for this group is consistently the largest relative to all other student groups.
3. The figures show that students at Alternate ELPAC Level 3 generally outperform students who are EO, RFEP, and IFEP on the CAA for ELA. In addition, across all the grades (three through eight, and eleven) EL students at Alternate ELPAC Overall Level 2 had about same mean and median scores as EO and RFEP students on the CAA for ELA.

Taken together, these results generally confirm Alternate ELPAC Level 3 as an appropriate threshold for reclassification, given the observed relationship between students’ Alternate ELPAC scores and CAA for ELA at this performance level. The achievement gap between Alternate ELPAC Level 3 students and EO and RFEP students on the CAA for ELA is notable and warrants additional consideration. Notes on how the CDE considered this in its analyses for the other methods are included in the summary of those results, and a comprehensive accounting of this gap in the overall recommendation is provided in the “Summary and Conclusions” section.

Figure 1: Descriptive Box Plot Grade Four



Caption: The chart shows the descriptive box plot analysis for grade four. The Y-axis is the CAA for ELA scale score. The X-axis contains the groups of EL students categorized by the proficiency level (Levels 1, 2, and 3) on Alternate ELPAC and the groups of non-EL students such as EO, IFEP, and RFEP . For each group of students, a box and whisker plot represents the minimum, mean, median, and maximum scale scores on CAA for ELA. A horizontal reference line represents the scale score of 460 at proficiency (or Level 3) on CAA for ELA. In the chart, about 40 percent of EL students at Level 3 on Alternate ELPAC achieve proficiency on CAA for ELA and they outperform all the other student groups. The EL students at Level 2 on Alternate ELPAC approximate the non-EL students on CAA for ELA.

Figure 2: Descriptive Box Plot Grade Seven



Caption: The chart shows the descriptive box plot analysis for grade seven. The Y-axis is the CAA for ELA scale score. The X-axis contains the groups of EL students categorized by the proficiency level (Levels 1, 2, and 3) on Alternate ELPAC and the groups of non-EL students such as EO, IFEP, and RFEP. For each group of students, a box and whisker plot represents the minimum, mean, median, and maximum scale scores on CAA for ELA. A horizontal reference line represents the scale score of 760 at proficiency (or Level 3) on CAA for ELA. In the chart, about 25 percent of EL students at Level 3 on Alternate ELPAC achieve proficiency on CAA for ELA and they outperform all the other student groups. The EL students at Level 2 on Alternate ELPAC approximate the non-EL students on CAA for ELA.

Figure 3: Descriptive Box Plot Grade Eleven



Caption: The chart shows the descriptive box plot analysis for grade eleven. The Y-axis is the CAA for ELA scale score. The X-axis contains the groups of EL students categorized by the proficiency level (Levels 1, 2, and 3) on Alternate ELPAC and the groups of non-EL students such as EO, IFEP, and RFEP. For each group of students, a box and whisker plot represents the minimum, mean, median, and maximum scale scores on CAA for ELA. A horizontal reference line represents the scale score of 960 at proficiency (or Level 3) on CAA for ELA. In the chart, about 25 percent of EL students at Level 3 on Alternate ELPAC achieve proficiency on CAA for ELA and they outperform all the other student groups. The EL students at Level 2 on Alternate ELPAC approximate the non-EL students on CAA for ELA.

*Note: In Figures 1 through 3, “TBD” refers to potential English learners who have not yet been tested to determine if they are at ELPAC level 1, 2, 3, or IFEP.*

### **Results from Logistic Regression Analysis**

Selected results for the logistic regression analysis are displayed in Figures 4–6, which show the probability of reaching Level 3 (Understanding) on the CAA for ELA for each Alternate ELPAC overall scale score value. As noted in the “Analytic Methods” section, the goal for this method is to identify the Alternate ELPAC scale score at which students have a 50 percent probability of scoring at Level 3 (Understanding) on the CAA for ELA. The figures show three notable results:

1. For all grades and groups, the scale score at which the probability is 50 percent falls in performance level 3 on the Alternate ELPAC.
2. The score at which the probability reaches 50 percent is higher in higher grades. In the grade eleven sample, there is no score at which the probability reaches the 50 percent target (see Figure 6).
3. The probability of achieving a Level 3 on the CAA for ELA at lower Alternate ELPAC performance levels is very low for all grades and does not vary by more than ten percentage points across the range of scores within the performance level. For example, the probability that a student at Alternate ELPAC Level 2 would score at Level 3 on the CAA for ELA ranges from approximately 5 to 15 percent for all grade levels shown here. This suggests there is a strong relationship between the ELP and ELA scores at these lower levels (i.e., students who attain any scale score within Alternate ELPAC Level 2 are very likely to not be proficient on the CAA for ELA).

Taken together, these results also affirm Alternate ELPAC Level 3 as the level at which the relationship between ELP and ELA decreases. Given the observed distributions of CAA for ELA performance among EO and RFEP students in the descriptive box plots methods (see figures 1 to 3 above), the CDE also ran logistic regression models using the mean EO achievement on CAA for ELA (not shown in this report), as well as the midpoint of Level 2 on the CAA for ELA, as alternate targets for the Alternate ELPAC to predict. Some of these results are shared below in the “Results from exploratory analyses” section.

Figure 4: Logistic Regression Grade Four, Proficiency Cut



Caption: The chart shows the logistic regression analysis for grade four. The Y-axis is the probability of EL students reaching proficiency (or Level 3) on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching proficiency on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching proficiency on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 544) and 3 (or scale score 560) on the Alternate ELPAC, respectively. The curved line crosses the horizontal reference line to the right of the vertical reference line for Level 3 on Alternate ELPAC (approximately at scale score 580), which means EL students must be beyond the Level 3 threshold on Alternate ELPAC to have at least a fifty percent chance of reaching proficiency on CAA for ELA.

Figure 5: Logistic Regression Grade Seven, Proficiency Cut



Caption: The chart shows the logistic regression analysis for grade seven. The Y-axis is the probability of EL students reaching proficiency (or Level 3) on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching proficiency on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching proficiency on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 644) and 3 (or scale score 660) on Alternate ELPAC, respectively. The curved line crosses the horizontal reference line to the right of the vertical reference line for Level 3 (approximately at scale score 690) on Alternate ELPAC, which means EL students must be beyond Level 3 threshold on Alternate ELPAC to have at least fifty percent chance of reaching proficiency on CAA for ELA.

Figure 6: Logistic Regression Grade Eleven, Proficiency Cut



Caption: The chart shows the logistic regression analysis for grade eleven. The Y-axis is the probability of EL students reaching proficiency (or Level 3) on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching proficiency on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching proficiency on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 844) and 3 (or scale score 860) on Alternate ELPAC, respectively. The curved line does not cross the horizontal reference line at all, which means EL students at Level 3 on Alternate ELPAC are not able to have a fifty percent chance of reaching proficiency on CAA for ELA.

### **Results from Decision Consistency Analysis**

Figure 7 shows the results of the decision consistency analysis for all tested grades in the study sample. As described above under “Analytic Methods,” this method optimizes consistent categorization of EL students as either proficient in both ELP and ELA, or in neither. For this analysis, students are further sorted into smaller groups based on whether their scale scores place them in the lower half (L) or higher half (H) of the score range for each Alternate ELPAC achievement level.

As Figure 7 illustrates, the results of these analyses generally corroborate earlier findings from the box plot and logistic regression analyses. The results are clearest for grade three, which shows a peak in consistent classifications at the lower end (threshold to below the midpoint) of the Alternate ELPAC Overall Level 3 scale score range, after which the percentage of consistent decisions decreases again. The other grades do not reach a similar peak (i.e., the curves at the median value of both the lower and higher half of Alternate ELPAC Overall Level 3 continue to have increasing slopes through the scale score range), but they do affirm that, across the performance continuum, consistent classifications are highest at Alternate ELPAC Level 3.

It is worth noting that this analysis is the most impacted, of the three, by the uneven score distributions on the Alternate ELPAC and CAA for ELA that are discussed above and summarized in Table 3. Because it was common for students to score one performance level higher on the Alternate ELPAC versus the CAA for ELA, there are very few instances of inconsistent classifications in the other direction (students who are proficient in ELA but not in ELP), which likely explains why most grades do not show a point of inflection and subsequent decrease that generally occurs in this analysis.

Figure 7: Using CAA ELA Level 3 as Proficiency Cutoff



Caption: The chart shows the decision consistency analysis results for all grade levels (3–8, and 11). The Y-axis is the percent of students who were consistently classified as proficient or not on CAA for ELA and the Alternate ELPAC, ranging from 0% to 100%. The X-axis is the levels on Alternate ELPAC broken down by lower half and higher half: Level 1 low, Level 1 high, Level 2 low, Level 2 high, Level 3 low, and Level 3 high. The curved line for each grade represents the consistency rate as categorized by each level. A peak on the curved line indicates that there is a point on the Alternate ELPAC scale where EL students reach the maximum consistency rate . The chart shows only students in grade 3 reach the peak at performance level 3 low. All other grades are increasing across the figure such that higher consistency occurs with higher performance levels. The highest consistency for grades 4–8 and 11 occurs at performance level 3 high and is between 80–90 percent consistent.

### **Results from Exploratory Analysis Setting California Alternate Assessment for English Language Arts Target at a Level Other than “Proficient”**

One function of EL status is to prepare EL students – including those with the most significant cognitive disabilities – to meet the same rigorous academic achievement standards set for all students. At the same time, it is also important–particularly given the language of the CDE’s reclassification Criterion 4–to try to avoid a situation wherein EL students appear to be held to a higher standard than their EO peers. These two goals appear to be somewhat in tension in the results above, specifically given that

1. As shown in the descriptive box plot results (see figures 1 to 3), students at Alternate ELPAC Level 3 generally outperformed EO and RFEP students, and
2. The results from all three methods, as well as the descriptive results for the sample generally, suggest that Level 3 on the CAA for ELA is a high standard to meet.

For this reason, as noted above, the CDE also conducted several exploratory analyses using different CAA for ELA scores as targets for proficiency. Specifically, the CDE ran logistic regression and decision consistency analyses using two alternative scores as a definition for acceptable performance on the CAA for ELA. These were:

1. The average score of EO students on the CAA for ELA in each grade level, and
2. The mid-point of Level 2 on the CAA for ELA

Selected results for the logistic regression models using the mid-point of Level 2 on the CAA for ELA as a target are shown in Figures 8–10. The results show that, with this score as the performance target, the point at which EL students’ probability of attaining proficiency on the CAA for ELA reaches the 50 percent target (suggesting equal odds of being proficient or non-proficient) is at or above the midpoint of Alternate ELPAC Level 2. Results are similar in other grades, and when using the mean CAA for ELA score of EO students as a target.

Figure 8: Logistic Regression Grade Four, CAA ELA L2 Mid-Point Cut



Caption: The chart shows the logistic regression analysis for grade four using the midpoint of Level 2 on CAA for ELA as the definition of proficiency. The Y-axis is the probability of EL students reaching midpoint of Level 2 on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching midpoint of Level 2 on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching midpoint of Level 2 on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 544) and 3 (or scale score 560) on Alternate ELPAC, respectively. The curved line crosses the horizontal reference line at approximately 556 (in between the two vertical reference lines), indicating that students must be beyond the Level 2 threshold but lower than the Level 3 threshold on Alternate ELPAC to have at least a fifty percent chance of reaching the midpoint of Level 2 on CAA for ELA.

Figure 9: Logistic Regression Grade Seven, CAA ELA L2 Mid-Point Cut



Caption: The chart shows the logistic regression analysis for grade seven using the midpoint of Level 2 on CAA for ELA as the definition of proficiency. The Y-axis is the probability of EL students reaching midpoint of Level 2 on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching midpoint of Level 2 on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching midpoint of Level 2 on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 644) and 3 (or scale score 660) on Alternate ELPAC, respectively. The curved line crosses the horizontal reference line at approximately 658 (in between two vertical reference lines), which means students must be beyond the Level 2 threshold but lower than the Level 3 threshold on Alternate ELPAC to have at least a fifty percent chance of reaching the midpoint of Level 2 on CAA for ELA.

Figure 10: Logistic Regression Grade Eleven, CAA ELA L2 Mid-Point Cut



Caption: The chart shows the logistic regression analysis for grade eleven using the midpoint of Level 2 on CAA for ELA as the definition of proficiency. The Y-axis is the probability of EL students reaching midpoint of Level 2 on CAA for ELA and ranges from 0 to 1. The X-axis is the scale score on Alternate ELPAC. A curved line represents the probability of each scale score on Alternate ELPAC reaching midpoint of Level 2 on CAA for ELA. A horizontal reference line represents the fifty percent chance (or 0.5) of reaching midpoint of Level 2 on CAA for ELA. Two vertical reference lines represent the threshold scores for Level 2 (or scale score 844) and 3 (or scale score 860) on Alternate ELPAC, respectively. The curved line crosses the horizontal reference line at approximately 852 (in between the two vertical reference lines), indicating that EL students must be beyond the Level 2 threshold but lower than the Level 3 threshold on Alternate ELPAC to have at a fifty percent chance of reaching midpoint of Level 2 on CAA for ELA.

Results for the exploratory decision consistency analyses using the mid-point of CAA for ELA Level 2 are shown in Figure 11 for all grades. With this target, all grades now show a clear point of inflection, beyond which the percentage of consistent classifications decreases (meaning more students are proficient in ELA but not ELP). The Figure shows that, in all grades except grade three, the peak in consistent classifications occurs in the lower half of Alternate ELPAC Level 3. (For grade three, consistency peaks in the higher half of Alternate ELPAC Level 2). These exploratory results affirm Alternate ELPAC Level 3 as an appropriate reclassification threshold.

Figure 11: Using Midpoint of CAA ELA Level 2 as Proficiency Cutoff



Caption: The chart shows the decision consistency analysis results for all grade levels (3–8, and 11) using midpoint of Level 2 on CAA for ELA as the definition of proficiency . The Y-axis is the percent of students who were consistently classified as proficient or not on CAA for ELA and the Alternate ELPAC, ranging from 0% to 100%. The X-axis is the levels on Alternate ELPAC broken down by lower half and higher half: Level 1 low, Level 1 high, Level 2 low, Level 2 high, Level 3 low, and Level 3 high. The curved line for each grade represents the consistency rate as categorized by each level. A peak on the curved line indicates that there is a point on the Alternate ELPAC scale where EL students reach the maximum consistency rate. The chart shows all grades peaking at performance level 2 high or performance level 3 low, at approximately 75–80 percent consistency.

Taken together, the results from these exploratory analyses suggest Alternate ELPAC Level 3 as an appropriate reclassification threshold. Choosing between Levels 2 and 3 requires a careful consideration of the trade-offs between potentially removing EL services earlier in a student’s language development trajectory (which is more likely to occur if the threshold is set at Level 2), versus potentially maintaining EL services past the point of their being useful (which is more likely to occur if the threshold is set at Level 3). As noted in the “Recommendations for Future Analysis” section below, an informed decision on this front requires understanding the nature and quality of English language development instruction that EL students with the most significant cognitive disabilities currently receive, and whether this instruction has any negative consequences in terms of access to content instruction or other learning opportunities. Recommending a Level 2 threshold would also constitute a policy change from the CDE’s current recommended threshold, which could have implementation implications for the field.

### **Outreach to Interest Holders**

As a final step in the analysis process, the CDE also shared information about this study, including the preliminary results, with several groups from across the state with interest and knowledge about English learners, students with disabilities, assessment, policy, and various combinations thereof. These groups included:

* The state’s Assessment Interest Holder representatives
* The state’s Special Education Local Plan Area representatives
* The state’s CAASPP and ELPAC TAG members
* Regional Assessment Network representatives

All groups were shown the same presentation and invited to ask questions and share feedback and reactions. In general, none of the groups expressed major concerns about maintaining the current Level 3 threshold, but almost all of the groups sought to contextualize the recommendation by asking questions about services and instruction for EL students with the most significant cognitive disabilities. Broadly speaking, special education and English learner interest holder groups were more concerned about students losing access to services, while the state TAG members were more concerned about students being retained in EL status. The questions and concerns that these groups raised informed the CDE’s final recommendation.

## Summary and Considerations

In summary, the CDE recommends that the SBE maintain the current threshold of Overall Performance Level 3 on the Summative Alternate ELPAC for use in meeting Criterion 1 for reclassification decisions for EL students with the most significant cognitive disabilities. This recommendation is based on the following evidence and considerations:

1. The results for all methods, all grade-levels, and all student groups (including the two largest disability groups within the overall sample) identify Alternate ELPAC Level 3 as the performance level at which the relationship between ELP and ELA decreases relative to lower Alternate ELPAC achievement levels. This suggests that language proficiency is becoming less likely to drive academic achievement at this performance level compared to other performance levels.
2. Exploratory decision consistency analyses using lower achievement targets on the CAA for ELA also affirm Level 3 as an appropriate target for reclassification.
3. The results of the analyses conducted by the CDE consistently affirm this standard as the performance level at which language and academic achievement become less strongly related to each other, which is an appropriate range for reclassifying students out of EL status and ceasing their EL services.

### **Recommendations for Future Analysis**

Although the recommendations in this report are strongly backed by data, there are certain unavoidable limitations to these analyses, which are acknowledged above under “Sample Limitations & Considerations.” Specifically, the current analyses are based on only a single cohort of students – who, importantly, are the very first cohort of students ever to be recommended for reclassification using the Alternate ELPAC, rather than a local measure. As noted above, interest holders in the field also expressed a desire for more information about programming and services for EL students with the most significant cognitive disabilities – a desire shared by the CDE.

In light of these considerations, the CDE recommends the following activities and analyses for future work:

1. **Replicate the current analyses with additional years of data as they become available.** As more students take the Summative Alternate ELPAC, and as students and educators become more familiar and confident with the Summative Alternate ELPAC’s content and administration, there is potential for the relationships and distributions to shift over time. By continuing to replicate these analyses as future years of data become available, the CDE will be able to monitor the findings of these analyses.
2. **Conduct additional analyses on current and future data.** In addition to replicating the current analyses with future data, the CDE could also expand on these analyses by, for example, replicating them with CAA for math scores, exploring the decisions and outcomes from reclassification decisions that have already been made for students who participated in the first two years of the Alternate ELPAC, or more closely evaluating the cut scores between the Alternate ELPAC achievement levels to see if these would benefit from minor adjustments (as was done for the general Summative ELPAC).
3. **Collect more detailed information about language instruction services for EL students with significant cognitive disabilities**. As noted above, the current recommendation errs on the side of maintaining services for classified students, which reflects the priorities of most advocacy and interest holder groups for this population. With more detailed information, the CDE could develop a more nuanced understanding of the benefits and trade-offs of continuing to receive services (such as whether students receiving ELD instruction miss out on other learning opportunities). This information can provide both the CDE and the field with a more detailed sense of what reclassification really means for students and ensure that the Summative Alternate ELPAC threshold is serving students’ best interest.
4. **Identify opportunities for data collection, professional learning, and technical assistance related to academic instruction for all students with significant cognitive disabilities.** As the descriptive box plot results showed, EO and RFEP students generally are not performing at high levels on the CAA for ELA. It may be the case that special educators and leaders need more supports or resources to implement instruction that prepares their students to meet the performance standards of the assessment. There may be opportunities for the CDE to bolster supports for special educators that may improve the achievement of students with the most significant cognitive disabilities.

With support from the SBE, the CDE will pursue these analyses to the extent practicable – particularly by leveraging work and relationships with their testing contractor, ETS, and two of the state’s federally-funded technical assistance centers, REL West and the Region 15 Comprehensive Center.

## Appendix A: Additional Information About Coaching Support from the Regional Educational Laboratory, West

The Regional Educational Laboratory (REL) West provided intensive coaching to CDE personnel as they engaged in the methods and analyses described in the previous section. REL West is one of ten RELS funded by the Institute of Education Sciences at the U.S. Department of Education. RELs partner and collaborate with school districts, state departments of education, and other education stakeholders to help generate and apply evidence, with the goal of improving learner outcomes. One way that RELs collaborate with partners is via Training, Coaching, and Technical Support (TCTS) projects.

TCTS projects leverage RELs’ unique expertise in designing and interpreting rigorous, relevant research, as well as the identification and application of evidence-based practices. TCTS includes:

* intensive training involving hands-on, direct instruction from experts in research or practice;
* coaching, or thought partnering, that supports decisionmakers in applying research evidence to inform high-leverage decisions and actions; and
* technical support to build partners’ capacity to identify, collect, analyze, and visualize data.

REL West engaged in TCTS activities with the CDE to support their completion of the study described above. Between October 2022 and May 2023, REL West met regularly with a team of collaborators from the CDE’s Assessment Development and Administration and Multilingual Support Divisions (ADAD and MSD, respectively) to help them prepare for, conduct, and interpret the analyses described above. Overall, this project included:

* Fifteen coaching meetings with MSD and ADAD to discuss the goals, methods, results, interpretations, and dissemination of the study.
* Five working sessions with ADAD staff to support their application of the methods to the CDE’s data and discuss results.
* Regular email communication to discuss questions, ideas, and concerns throughout the collaboration.

All materials for the coaching and data session were reviewed and approved by REL West leadership and by the Institute of Education Sciences.

Additionally, REL West entered into a data sharing agreement with the CDE to request and receive the same data set used by the CDE for these analyses so that REL West could replicate the analyses and results independently as a quality assurance check.

1. Christensen, L, L., Mitchell, J. D., Shyyan, V. V., & Ryan, S. (2018). Characteristics of English learners with significant cognitive disabilities: Findings from the Individual Characteristics Questionnaire. Alternate English Language Learning Assessment (ALTELLA). <https://altella.wceruw.org/pubs/ICQ-Report.pdf> ; Cook, H. G. (2014). Examining relationships between alternate ACCESS and state alternate assessments: Exploring notions of English language proficiency. WIDA Research Report. University of Wisconsin, Madison, Wisconsin Center for Education Research. <https://wida.wisc.edu/sites/default/files/resource/Report-ExaminingRelationshipsBetweenAlternateAccessandStateAlternateAssessmentsAL.pdf>​ [↑](#footnote-ref-2)
2. Linquanti, R., Huang, M., & Crane, E. (2018) *Report on Supplemental Empirical Analyses of the English Language Proficiency Assessments for California.*  Sacramento, CA: California Comprehensive Center at WestEd. [↑](#footnote-ref-3)
3. Cook, H. G. (2014). Examining relationships between alternate ACCESS and state alternate assessments: Exploring notions of English language proficiency. WIDA Research Report. University of Wisconsin, Madison, Wisconsin Center for Education Research. <https://wida.wisc.edu/sites/default/files/resource/Report-ExaminingRelationshipsBetweenAlternateAccessandStateAlternateAssessmentsAL.pdf> [↑](#footnote-ref-4)