California Department of Education

Executive Office

SBE-002 (REV. 11/2017)

memo-pptb-adad-oct18item02

# **MEMORANDUM**

**DATE:** October 15, 2018

**TO:** MEMBERS, State Board of Education

**FROM:** TOM TORLAKSON, State Superintendent of Public Instruction

**SUBJECT:** Summative English Language Proficiency Assessments for California: Update on the Threshold Score Validation Study and Supplemental Empirical Analyses.

## Summary of Key Issues

This Memorandum provides a summary of the Summative English Language Proficiency Assessments for California (ELPAC) threshold score validation study and supplemental empirical analyses. Results from these studies will be used to inform proposed recommendations to the State Board of Education (SBE) in November 2018.

### Purpose of the Studies

In December 2017, the California Department of Education (CDE) submitted a Memorandum to the SBE that outlined the expectations of the threshold score validation study and supplemental empirical analyses upon completion of the first operational Summative ELPAC. These analyses align with the current practice that has been conducted by other states to determine their threshold scores. The Memorandum can be found on the CDE SBE information memorandum web page at <https://www.cde.ca.gov/be/pn/im/documents/memo-pptb-adad-dec17item03.docx>.

The purpose of the threshold score validation study was to evaluate the degree to which the threshold scores and performance levels of the Summative ELPAC accurately distinguish between levels of students’ English proficiency. Included in Attachment 1 is an executive summary of the study from the ELPAC testing contractor, Educational Testing Service (ETS), of the results based on students in kindergarten through grade twelve.

The purpose of the supplemental empirical analyses was to examine the relationship between student test scores and performance levels on the first operational Summative ELPAC relative to student test scores and performance levels on the California Assessment of Student Performance and Progress (CAASPP) English language arts/literacy (ELA). The primary purpose was to examine the ELPAC level 4 students’ performance on the CAASPP ELA as one way to evaluate if those students performed similarly to English only (EO) students on the CAASPP ELA. Both assessments were administered in the spring of 2018. Included in Attachment 2 is an executive summary from the California Comprehensive Center at WestEd of the results from their supplemental empirical analyses of the ELPAC based on grades three through eight and grade eleven.

Validation of threshold scores and supplemental empirical analyses of the relationship of ELPAC to CAASPP ELA are critical for appropriate test use and state-level policy discussions on establishing an ELPAC criterion for English learner (EL) reclassification decisions.

### Next Steps

The CDE is currently reviewing the results of the threshold score validation study and supplemental empirical analyses to consider whether to maintain or adjust 2017–18 threshold scores approved on November 8, 2017. The CDE will provide the SBE with a recommendation in November 2018. If the SBE approves threshold score changes, the CDE will notify LEA Superintendents and charter school administrators of those changes to the 2018-19 threshold scores.

In addition, the CDE has begun recruitment of teachers to participate in the Initial ELPAC threshold validation review process this fall. The CDE is anticipating collecting surveys from teachers for approximately 1,500 students of which 50 percent will be ELs and 50 percent will be initial fluent English proficient (IFEP). Teachers will use the approved general performance level descriptors to agree or disagree with the classification of their students as either EL or IFEP. After surveys are collected, an analysis of the responses will be conducted by ETS. CDE will utilize the results from the analysis to inform a recommendation to either maintain or revise the Initial ELPAC threshold scores for the 2019–20 school year. The recommendation related to the Initial ELPAC threshold scores is anticipated to be brought before the SBE in March 2019.

## Attachment(s)

* Attachment 1: English Language Proficiency Assessments for California Threshold Score Validation Study Executive Summary (3 Pages)
* Attachment 2: Supplemental Empirical Analyses of the English Language Proficiency Assessments for California (3 Pages)

# **Attachment 1: English Language Proficiency Assessments for California Threshold Score Validation Study Executive Summary**

## Overview

At the request of the California Department of Education (CDE), Educational Testing Service (ETS) conducted a threshold score validation study to provide additional validity evidence of the Summative English Language Proficiency Assessments for California (ELPAC) 2017–18 threshold scores that were approved by the State Board of Education in November 2017. The methodology used for the study is known as the Contrasting Groups Method[[1]](#footnote-1) and is described below.

The purpose of conducting the threshold score validation study was to evaluate the degree to which the threshold scores and performance levels of the Summative ELPAC consistently distinguish between levels of students’ English language proficiency, based on teacher ratings from a multi-step process (standard setting in October 2017 and the validation study). Implementing a multistep process offers increased confidence in decisions utilizing threshold scores based on ELPAC results. Because the classification of the proficiency levels for English learners (ELs) entails relatively high-stakes decisions for individual students’ academic paths, school program funding, and resource plans, it is crucial to cross-validate the threshold scores for each proficiency level to the extent possible.

## Contrasting Groups Method

Using a methodology known as Contrasting Groups, teacher judgments of students’ performance levels were collected. At a point approximately seven months into the school year, teachers familiar with students in their classroom were asked to classify students according to the approved ELPAC General Performance Level Descriptors (PLDs) and ELPAC domain and grade/grade-span–specific PLDs. Selection of local educational agencies (LEAs) participating in the contrasting groups study was targeted to allow for a statewide demographic representation of LEAs and students, a wide range of performance and to require a reasonable number of student ratings from each teacher.

The judgment of the teachers was based on their knowledge and understanding of their own students’ levels of proficiency, relative to the California-approved final PLDs. Note that California-approved PLDs were the starting point for this contrasting groups study, thereby maintaining the meaning of the performance levels from the standard setting studies for consistency and standardization. A statistical analysis was conducted comparing students’ ELPAC scores to teachers’ judgments.

The results of the statistical analyses and teacher judgments can be used in concert with other information including the results from the panel-based standard setting and post-standard setting considerations and the impact of threshold scores on the ELPAC score distributions.

### Description of sample of participating teachers and students

The selection of LEAs was done in concert with input from the CDE and the ELPAC Technical Advisory Group (TAG)*.* The goals of the teacher recruitment were (1) to obtain teacher ratings on ELs across the state who represent the full range of English language proficiency levels and (2) to select teachers who had not administered the operational ELPAC to their students in order to eliminate potential bias in their ratings. In the final analysis, 1,521 teachers and 11,128 students, from 154 LEAs across the state were included. The student sample characteristics were representative of the 2017–18 operational ELPAC test taker population.

### Instructions to participating teachers

Teachers were provided the domain- and grade-specific Performance Level Descriptors for each of the four domains, and the General PLDs, and asked to become familiar with these documents for the their students’ grade level.

Participating teachers received an overview of the ELPAC, the score reporting hierarchy and how the overall score is calculated, and a review of the rating form and how to complete the form. Teachers were asked to complete one form per student and to provide the expected overall score performance level.

## Results

Results from the rater agreement analysis of the contrasting groups study indicated a trend for most grades for most performance levels. Teacher ratings of expected performance level were compared to the students’ ELPAC performance levels based on the 2017–18 threshold scores. The comparison indicated an exact or adjacent agreement rate between 85 percent and 90 percent in kindergarten and grades one and two, and a rate at 90 percent or above for grades three through twelve. This means, for example, where teacher ratings of students indicated a Level 3, exact or adjacent agreement was met if the students’ ELPAC performance level was a Level 2 or 3 or 4. For most grades, it appears teachers who rated the ELs in their classrooms placed their students at a lower performance level than what the ELPAC 2017–18 threshold score indicated.

## Next steps

Final analyses based on the 2018 ELPAC operational data are underway by the CDE. The impact on the students who will be classified into each performance level must be considered prior to making recommendations for adjusting the 2017–18 Summative ELPAC threshold scores. Analyses will take into account the results from the threshold score validation study, score scales and score distributions, as well as the four performance levels reported for three composite scores, Oral Language, Written Language, and an Overall Score.

# **Attachment 2: Supplemental Empirical Analyses of the English Language Proficiency Assessments for California**

## Background

In addition to a threshold score validation study conducted by the Educational Testing Service, the CDE requested the California Comprehensive Center at WestEd (CA CC) to conduct supplemental analyses to examine the relationship of English Learners’ (ELs) performance on the California Assessment of Student Performance and Progress (CAASPP) English language arts/literacy (ELA) relative to their English Language Proficiency Assessments for California (ELPAC) performance. Validation of threshold scores and empirical analyses of the relationship of ELPAC to CAASPP ELA are critical for appropriate test use and state-level policy discussions on establishing an ELPAC criterion for EL reclassification decisions.

Specifically, these empirical analyses examined the relationship of student test scores and performance levels on the first operational Summative ELPAC to student test scores and performance levels on the 2018 CAASPP ELA. Both assessments were administered in spring 2018.

## Data Sources

The CDE supplied the CA CC with a matched student level data file containing: a) 2018 CAASPP ELA assessment data for **all students** in grades three through eight and grade eleven, including overall scale score and performance level results for each subject area; and b) 2018 Summative ELPAC assessment data for **all EL students** in grades three through eight and grade eleven, including scale score and performance level results at the overall, oral and written, and individual (listening, speaking, reading, and writing) domain levels.

## Analytic Methods

Three analytical methods1 were undertaken:

1. **Descriptive box plot analysis**[[2]](#footnote-2)examines the distribution of overall scale scores on CAASPP ELA for ELs by each performance level on ELPAC, for each applicable grade level. This analysis also includes the distribution of overall scale scores on CAASPP ELA for reclassified fluent English proficient (RFEP), initially fluent English proficient (IFEP), and monolingual English (English only, or EO) students for comparison. The purpose of the analysis is to identify an ELPAC performance level where a) at least half the EL students score at or above the CAASPP ELA Level 3 threshold (Standard Met); or b) the distribution of EL students’ overall scale scores on CAASPP ELA is very similar to that of EO students statewide.
2. **Logistic regression analysis** estimates the probability of reaching Level 3 (Standard Met) on CAASPP ELA for each ELPAC overall scale score. This approach helps to identify the ELPAC overall scale score range in which EL students have a probability equal to or greater than 50 percent (0.5) of attaining that standard on the CAASPP ELA.
3. **Decision consistency analysis** analyzes ELPAC and CAASPP proficient-level categorizations and optimizes consistent categorization of EL students at or above the current CAASPP ELA threshold score for Level 3 (Standard Met). The analysis determines the ELPAC overall score range that maximizes the amount of agreement between achieving ELPAC proficiency and CAASPP ELA proficiency.[[3]](#footnote-3)

## Findings and Next Steps

For grades three through eight and eleven, 24 percent of ELPAC examinees attained the 2018 ELPAC Overall performance level (PL) 4, which is currently the ELPAC threshold level performance for the English-proficient criterion based on the 2017–18 threshold scores approved by the State Board in 2017.

Grade-level ELPAC-CAASPP empirical analyses largely converged across methods. Analyses suggest that, for grades four through eight and eleven, ELs at ELPAC Overall PL 4 do not approximate the 50 percent level/0.5 probability of attaining Level 3 on CAASPP ELA, nor do they approximate EO student performance. This is particularly so for ELs in grades seven, eight, and eleven. In grade three, ELs at ELPAC Overall PL 4 outperformed EO students and had a greater than 50 percent probability of reaching Level 3 on CAASPP ELA.

In light of these initial findings, the CA CC team examined the difference in CAASPP ELA performance for ELs at the minimum scale score for ELPAC Overall PL 4 to just below the midpoint of the Overall PL 4 scale score range, compared to those at the midpoint or higher. The team also analyzed options for adding conjunctive rules related to the oral and written composite domains for EL students at ELPAC Overall PL 4. These analyses did not substantially change the initial findings, and will be summarized in the final report.

1. Zieky, M.J, Perie, M. & Livingston, S.A. (2008). *Cutscores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*. Princeton, NJ: Educational Testing Service. [↑](#footnote-ref-1)
2. 1 These methods, described in Cook, Linquanti, Chinen & Jung (2012), <https://www2.ed.gov/rschstat/eval/title-iii/implementation-supplemental-report.pdf>, are used by many states to support decision making on setting an optimal English-proficient performance standard on their state English Language Proficiency (ELP) assessments.

   2 A box plot shows graphically five-number summaries—the smallest observation, lower quartile, median, upper quartile, and largest observation—as well as individual outliers, if applicable. [↑](#footnote-ref-2)
3. 3 The analysis takes the CAASPP ELA Level 3 as given, and determines the ELPAC overall score range that maximizes the percentage of true positives (proficient on ELPAC and CAASPP ELA 3) and true negatives (not proficient on both assessments) and minimizes the percentage of false positives (proficient on ELPAC, not proficient on CAASPP) and false negatives (not proficient on ELPAC, proficient on CAASPP). [↑](#footnote-ref-3)