

# Initial English Language Proficiency Assessments for California (ELPAC) Threshold Score Review Study Final Report

**Contract #CN140284**

**Prepared for the California Department of Education by Educational Testing Service**

**Presented February 15, 2019**

Table of Contents

[Background 1](#_Toc428532)

[Overview 1](#_Toc428533)

[Contrasting Groups Method 2](#_Toc428534)

[Description of the Sample of Participating Teachers and Students 2](#_Toc428535)

[Study Sample: Representativeness 3](#_Toc428536)

[Instructions to Participating Teachers 6](#_Toc428537)

[Analytic Methods 7](#_Toc428538)

[Results 10](#_Toc428539)

[Summary 11](#_Toc428540)

[Next Steps 11](#_Toc428541)

List of Tables

[Table 1. Percentage of Students by Geographic Region: Initial ELPAC Participants in California and in Study Sample 3](#_Toc533290)

[Table 2. Number of Students Rated by Grade 4](#_Toc533291)

[Table 3. Student Characteristics of Study Sample, Overall 4](#_Toc533292)

[Table 4. Student Characteristics of Study Sample, Combined Kindergarten 5](#_Toc533293)

[Table 5. Student Characteristics of Study Sample, Grade One 5](#_Toc533294)

[Table 6. Student Characteristics of Study Sample, Grade Two 5](#_Toc533295)

[Table 7. Student Characteristics of Study Sample, Grades Three through Five 6](#_Toc533296)

[Table 8. Student Characteristics of Study Sample, Grades Six through Eight 6](#_Toc533297)

[Table 9. Student Characteristics of Study Sample, Grades Nine through Twelve 6](#_Toc533298)

[Table 10. ELAS Agreement for Combined Kindergarten\* 7](#_Toc533299)

[Table 11. ELAS Agreement for Grade One 8](#_Toc533300)

[Table 12. ELAS Agreement for Grade Two 8](#_Toc533301)

[Table 13. ELAS Agreement for Grades Three through Five 8](#_Toc533302)

[Table 14. ELAS Agreement for Grades Six through Eight 9](#_Toc533303)

[Table 15. ELAS Agreement for Grades Nine through Twelve 9](#_Toc533304)

[Table 16. ELAS Agreement for All Grades 9](#_Toc533305)

[Table 17. Interrater Agreement by ELAS for Grades Six through Twelve 10](#_Toc533306)



## Background

The Initial English Language Proficiency Assessments for California (ELPAC) field test administration was conducted in fall 2017, during the 2017–18 school year. The first operational administration of the Initial ELPAC was conducted in fall 2018. The assessments, given in paper and pencil, were administered at six grades or grade spans (kindergarten [K], one, two, three to five, six to eight, nine to twelve) and assessed four domains (Listening, Speaking, Reading, and Writing). The task types and domain descriptions are described on the California Department of Education (CDE) ELPAC web page at <https://www.cde.ca.gov/ta/tg/ep/>. The score-reporting hierarchy for the Initial ELPAC was approved in September 2017 by the State Board of Education (SBE); three performance levels are reported for three composite scores: Oral Language, Written Language, and Overall Score.

To develop threshold-score recommendations aligned to the score-reporting hierarchy, Educational Testing Service (ETS) conducted standard-setting workshops in Sacramento, California, in February 2018. All four domains and the overall score were considered in the process of standard setting. Teachers who are familiar with the 2012 California English Language Development (ELD) Standards and familiar with the students taking the Summative ELPAC participated in a four-day workshop, during which they reviewed and discussed the items on the test, the general and specific performance level descriptors (PLDs), and the knowledge and skills measured by the ELPAC that differentiate student performance levels. Teachers recommended threshold scores for all grades and grade spans. The SBE approved preliminary threshold scores on May 9, 2018.

## Overview

At the CDE’s request, ETS conducted a threshold score review study to provide additional validity evidence for the Initial ELPAC preliminary threshold scores that were approved by the SBE. The Initial ELPAC assesses all students whose primary language is a language other than English to determine English Language Acquisition Status (ELAS)—students are designated as either an English learner (EL) or initial fluent English proficient (IFEP).

During the study, ETS gathered information about the extent to which California educators with students who took the Initial ELPAC agree that their students were correctly classified as EL or IFEP. The results provide the CDE with

* educator judgments from a multistep process;
* recommendations from the panel-based standard setting, conducted in February 2018, using the Bookmark standard setting method; and
* educator judgments from the validation study.

Implementing a multistep process offers increased confidence in decisions made using threshold scores based on ELPAC results.

## Contrasting Groups Method

Using a methodology known as contrasting groups[[1]](#footnote-1), teachers familiar with students who were classified as EL or IFEP in their classroom were asked to make their own judgments about the students’ classification based on the approved Initial ELPAC PLDs. Selection of local educational agencies (LEAs) participating in the contrasting groups study was targeted to provide statewide demographic representation of LEAs and students, a wide range of performance, and a sufficient 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. These judgments were made between one and three months after the Initial ELPAC was administered. For the Initial ELPAC study, the focus was on students’ performance relative to ELAS. A statistical analysis was conducted comparing students’ ELPAC scores to teachers’ judgments.

Note that the California-approved PLDs were the starting point for this contrasting groups study, thereby maintaining the meaning of the performance levels from the two studies for consistency and standardization.

### Description of the Sample of Participating Teachers and Students

The selection of LEAs was based on the goal of obtaining a large, representative sample of California students classified as EL and IFEP, and the teachers working with those students. Sampling was conducted with input from the CDE and the ELPAC Technical Advisory Group*.*

The goal of the recruitment was to obtain a sample of approximately 100 teacher ratings per grade. A sample of 50 IFEP and 50 EL student ratings was targeted. Additional grade-range specific recruitment goals for grade six and above were specified based on feedback from the ELPAC Technical Advisory Group.

Because the number of students is higher proportionally for the Initial ELPAC in K and grade one, a sample of approximately 200 students was targeted for these two grades, i.e., 100 IFEP and 100 EL students. For grades two through twelve, a sample of 50 IFEP and 50 EL student ratings was targeted. For grades six through twelve, recruitment efforts targeted ratings from two teachers—an English language arts (ELA)/ELD teacher and a content-area teacher—and the goal was to collect two judgments for each student.

The overall sample was to have the following characteristics:

* Representative geographically
* Representative by LEA type

### Study Sample: Representativeness

A total of 95 LEAs across California participated in the initial ELPAC validation process, with a total of 1,919 teachers providing ratings for 2,844 students on their students’ English-language proficiency based on the initial ELPAC PLDs. Taking into account school size, represented by the locked scores data file using the Local Scoring Tool as of December 16, 2018, compared to the schools in the study (study sample), results suggest that the study sample includes a reasonable distribution of school sizes.

The percentages of students who are ELs from each geographic region in California—North, Central, and South—and for the study sample are presented in Table 1. The study sample has the highest percentage of students in the southern region, as is the case in California. The percentages of students in the study sample also represented the variety of LEA types in the state, including charter schools.

Table 1. Percentage of Students by Geographic Region: Initial ELPAC Participants in California and in Study Sample

|  |  |  |
| --- | --- | --- |
| Region | Initial ELPAC Participants | Study Sample |
| North | 9.0% | 10.4% |
| Central | 30.5% | 30.3% |
| South | 60.5% | 59.3% |

Recruiting resulted in the inclusion of ratings for more than 400 students in K (which includes transitional kindergarten [TK] students who are in the first year of a two-year K program) and more than 250 students in grade one. Table 2 shows the number of students by grade in the study sample and provides evidence of meeting or exceeding the targeted sample.

Table 2. Number of Students Rated by Grade

|  |  |
| --- | --- |
| Grade | Number of Students |
| K\* | 426 |
| 1 | 273 |
| 2 | 184 |
| 3 | 173 |
| 4 | 174 |
| 5 | 179 |
| 6 | 167 |
| 7 | 222 |
| 8 | 199 |
| 9 | 318 |
| 10 | 202 |
| 11 | 187 |
| 12 | 140 |
| Total | 2,844 |

\* Data includes 69 TK students.

Table 3 provides a view of the student characteristics of the full study sample of 2,844 students rated by their teachers in the contrasting groups study.

Table 3. Student Characteristics of Study Sample, Overall

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 2,844 |
| Students designated as IFEP | 38.2% |
| Students designated as EL | 61.8% |
| Male | 51.1% |
| Female | 48.9% |

Table 4 through Table 9 provide information about the demographic characteristics of the study sample. Recruiting efforts for the study focused on students who are EL and IFEP. Table 4 presents data combined for students in TK and K. Based on data available on the Initial ELPAC test takers as of mid-January 2019, these characteristics are representative of the overall student test taking populations. Further comparisons of the study sample and full Initial ELPAC test taking sample can be conducted in fall 2019, when the final Initial ELPAC data is available after the testing window for the Initial ELPAC has closed.

Table 4. Student Characteristics of Study Sample, Combined Kindergarten

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 426 |
| Students designated as IFEP | 40.6% |
| Students designated as EL | 59.4% |
| Male | 45.8% |
| Female | 54.2% |

Table 5. Student Characteristics of Study Sample, Grade One

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 273 |
| Students designated as IFEP | 38.8% |
| Students designated as EL | 61.2% |
| Male | 54.2% |
| Female | 45.8% |

Table 6. Student Characteristics of Study Sample, Grade Two

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 184 |
| Students designated as IFEP | 31.0% |
| Students designated as EL | 69.0% |
| Male | 52.2% |
| Female | 47.8% |

Table 7. Student Characteristics of Study Sample, Grades Three through Five

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 526 |
| Students designated as IFEP | 30.6% |
| Students designated as EL | 69.4% |
| Male | 53.0% |
| Female | 47.0% |

Table 8. Student Characteristics of Study Sample, Grades Six through Eight

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 588 |
| Students designated as IFEP | 36.1% |
| Students designated as EL | 63.9% |
| Male | 54.6% |
| Female | 45.4% |

Table 9. Student Characteristics of Study Sample, Grades Nine through Twelve

|  |  |
| --- | --- |
| Student Characteristics | Study Sample |
| N Count | 847 |
| Students designated as IFEP | 44.6% |
| Students designated as EL | 55.4% |
| Male | 48.8% |
| Female | 51.2% |

### Instructions to Participating Teachers

Teachers were provided the Initial ELPAC General PLDs describing the range of performance for IFEP and ELs and were asked to refer to these descriptors. The rating form for each student indicated the student classification based on the Initial ELPAC (IFEP or EL). Teachers were asked to evaluate the degree to which they agree or disagree with the student ELAS resulting from the Initial ELPAC.

### Analytic Methods

#### Teacher Agreement Analysis

Table 10 through Table 16 show the number of students classified as EL or IFEP based on the Initial ELPAC Threshold Overall Score and the degree of agreement with that classification based on the teacher ratings in the contrasting groups study. Each table presents, for a grade or grade-span test, the number and percent of students for whom teacher ratings were in agreement (*strongly agree* or *agree*) or in disagreement (*disagree* or *strongly disagree*) with the student classification. A summary across all grades for the full sample of teacher ratings of students is displayed in Table 16.

Table 10 describes agreement for both TK and K.

Table 10. English Language Acquisition Status––Teacher Agreement for Combined Kindergarten

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| TK Strongly agree | 36 | 64% | 13 | 100% |
| TK Agree | 15 | 27% | 0 | 0% |
| TK Disagree | 5 | 9% | 0 | 0% |
| TK Strongly disagree | 0 | 0% | 0 | 0% |
| **TK Total** | **56** | **100%** | **13** | **100%** |
| K Strongly agree | 101 | 51% | 101 | 63% |
| K\* Agree  | 80 | 40% | 52 | 33% |
| K Disagree | 11 | 6% | 6 | 4% |
| K Strongly disagree | 6 | 3% | 1 | 0% |
| **K Total** | **198** | **100%** | **160** | **100%** |
| TK/K Strongly agree | 137 | 54% | 114 | 66% |
| TK/K Agree | 95 | 37% | 52 | 30% |
| TK/K Disagree | 16 | 6% | 6 | 3% |
| TK/K Strongly disagree | 6 | 2% | 1 | 1% |
| **TK/K Total** | **254** | **NA** | **173** | **NA** |

\*One K student has two ratings.

Table 11. English Language Acquisition Status––Teacher Agreement for Grade One

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 118 | 71% | 65 | 61% |
| Agree  | 41 | 25% | 37 | 35% |
| Disagree  | 5 | 3% | 4 | 4% |
| Strongly disagree  | 3 | 2% | 0 | 0% |
| Total | 167 | 100% | 106 | 100% |

Total percent may not sum to 100 due to rounding.

Table 12. English Language Acquisition Status––Teacher Agreement for Grade Two

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 82 | 65% | 42 | 74% |
| Agree  | 41 | 32% | 14 | 25% |
| Disagree  | 2 | 2% | 1 | 2% |
| Strongly disagree  | 2 | 2% | 0 | 0% |
| Total | 127 | 100% | 57 | 100% |

Total percent may not sum to 100 due to rounding.

Table 13. English Language Acquisition Status––Teacher Agreement for Grades Three through Five

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 274 | 75% | 96 | 60% |
| Agree  | 64 | 18% | 50 | 31% |
| Disagree  | 17 | 5% | 12 | 7% |
| Strongly disagree  | 10 | 3% | 3 | 2% |
| Total | 365 | 100% | 161 | 100% |

Total percent may not sum to 100 due to rounding.

Table 14. English Language Acquisition Status––Teacher Agreement for Grades Six through Eight

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 489 | 75% | 222 | 59% |
| Agree  | 133 | 20% | 128 | 34% |
| Disagree  | 25 | 4% | 19 | 5% |
| Strongly disagree  | 9 | 1% | 7 | 2% |
| Total | 656 | 100% | 376 | 100% |

Total percent may not sum to 100 due to rounding.

Table 15. English Language Acquisition Status––Teacher Agreement for Grades Nine through Twelve

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 614 | 70% | 374 | 55% |
| Agree  | 239 | 27% | 249 | 37% |
| Disagree  | 18 | 2% | 43 | 6% |
| Strongly disagree  | 2 | 0% | 11 | 2% |
| Total | 873 | 100% | 677 | 100% |

Total percent may not sum to 100 due to rounding.

Table 16. English Language Acquisition Status––Teacher Agreement for All Grades

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Teacher Rating | EL (N) | EL | IFEP (N) | IFEP |
| Strongly agree  | 1,714 | 70% | 913 | 59% |
| Agree  | 613 | 25% | 530 | 34% |
| Disagree  | 83 | 3% | 85 | 5% |
| Strongly disagree  | 32 | 1% | 22 | 1% |
| Total | 2,442 | 100% | 1,550 | 100% |

Total percent may not sum to 100 due to rounding.

## Results

Results from the teacher-to-teacher agreement analysis of the contrasting groups study, summarized in Table 17, indicated strong agreement with ELAS based on the Initial ELPAC across all grades:

* In K through grade five, 90 percent to 97 percent of teachers indicated agreement or strong agreement.
* In grades six through twelve, 91 percent to 98 percent of teachers indicated agreement or strong agreement.
* Between 86 and 100 percent of teacher ratings in grades six through high school indicated *agree* or *strongly agree* with the ELAS classification of the students in the classroom, for both the ELD/ELA and the content classroom teachers. In general, there was strong agreement as to the classification of the students based on two teachers.
	+ For grades six through high school, where possible, two teacher ratings were collected for each student.
	+ The specific goal was to collect ratings from an ELA/ELD teacher and a content-area teacher.
	+ Teachers were asked the same question, to rate the extent analyses of the results from teachers working on different content areas in their classrooms (e.g., ELA/ELD versus mathematics or science teachers) were conducted.
	+ The ratings were similar to the data reported previously in Table 10 through Table 16 and supported the Initial ELPAC classification.

Table 17. Teacher-to-Teacher Agreement by ELAS for Grades Six through Twelve

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade | Students with Two Ratings (N) | Agreement in Ratings (N) | Percent of Agreement in Ratings | Disagreement in Ratings (N) | Percent of Disagreement in Ratings |
| 6 | 86 | 84 | 98% | 2 | 2% |
| 7 | 187 | 172 | 92% | 15 | 8% |
| 8 | 171 | 155 | 91% | 16 | 9% |
| 9 | 261 | 249 | 95% | 12 | 5% |
| 10 | 174 | 162 | 93% | 12 | 7% |
| 11 | 161 | 146 | 91% | 15 | 9% |
| 12 | 107 | 100 | 93% | 7 | 7% |
| Combined Total | 1,147 | 1,068 | 98% | 79 | 7% |

## Summary

The standard-setting workshop conducted in February 2018 resulted in recommendations for preliminary threshold scores, which, in turn, resulted in performance levels for students who took the fall 2017–18 Initial ELPAC field test. Subsequently, a threshold score review study was conducted, using a contrasting groups standard-setting method.

Teachers considered the performance of EL and IFEP students in their classrooms and provided agreement ratings of the classification of their students. An analysis of the extent of agreement or disagreement for classification of those students based on the preliminary threshold scores was conducted. The conclusion from the analysis indicates that there is a high degree of agreement among the teachers across grades that their students who were classified as EL or IFEP based on the Initial ELPAC threshold scores are classified appropriately.

## Next Steps

Upon review of this final report by the CDE, the results of this study can be used to inform a recommendation to maintain or revise the Initial ELPAC threshold scores approved by the SBE in May 2018.

ETS will replicate the Initial ELPAC threshold score study using the Initial ELPAC results from the 2020–21 computer-based assessment administration. In spring 2021, the CDE will provide the SBE with the results of the replication and any recommendations for changes to the threshold scores following the review study.

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)