

Standard Setting Technical Report for the Alternate English Language Proficiency Assessments for California

**Contract #CN150012**

**Prepared for the California Department of Education by ETS**

**Presented July 18, 2022**



Table of Contents

[Chapter 1: Introduction 1](#_Toc107300819)

[Purpose and General Description of the Standard Setting Workshop 3](#_Toc107300820)

[Chapter 2: Method 4](#_Toc107300821)

[Modified and Extended Angoff Methods 4](#_Toc107300822)

[Standard Setting Panels 5](#_Toc107300823)

[Chapter 3: Materials 11](#_Toc107300824)

[Judgment Training Materials 11](#_Toc107300825)

[Evaluation Forms 12](#_Toc107300826)

[Chapter 4: Process 13](#_Toc107300827)

[Preparation and Training 13](#_Toc107300828)

[General Overview of Training and Process 13](#_Toc107300829)

[Test Familiarization 14](#_Toc107300830)

[Threshold Student Definitions 14](#_Toc107300831)

[Panelist Judgments 14](#_Toc107300832)

[Item Scoring, Judgments, and Rating Scales 15](#_Toc107300833)

[Modified and Extended Angoff Judgments 15](#_Toc107300834)

[Feedback and Discussion 17](#_Toc107300835)

[Vertical Articulation 17](#_Toc107300836)

[Chapter 5: Results 20](#_Toc107300837)

[Five Types of Data Tables 20](#_Toc107300838)

[Data Presentation 20](#_Toc107300839)

[Alternate ELPAC Threshold Score Results 21](#_Toc107300840)

[Incorporating Additional Considerations in Setting Threshold Scores 28](#_Toc107300841)

[Evaluation Results from the Training Evaluation of the Standard Setting Process 28](#_Toc107300842)

[Evaluation Results from the Alternate ELPAC Standard Setting Final Evaluations 29](#_Toc107300843)

[Chapter 6: Post Standard Setting Results 39](#_Toc107300844)

[Conclusion 39](#_Toc107300845)

[References 40](#_Toc107300846)

[Appendix 1: Attachments 41](#_Toc107300847)

[Attachment A: Panelist Pre-Workshop Assignment, Grades Nine Through Twelve   
Sample 41](#_Toc107300848)

[Attachment B: Final Threshold Student Definitions 45](#_Toc107300849)

[Alternate ELPAC Kindergarten Threshold Student Definitions by Connector 45](#_Toc107300850)

[Alternate ELPAC Grade One Threshold Student Definitions by Connector 46](#_Toc107300851)

[Alternate ELPAC Grade Two Threshold Student Definitions by Connector 47](#_Toc107300852)

[Alternate ELPAC Grades Three Through Five Threshold Student Definitions by Connector 50](#_Toc107300853)

[Alternate ELPAC Grades Six Through Eight Threshold Student Definitions by Connector 52](#_Toc107300854)

[Alternate ELPAC Grades Nine Through Twelve Threshold Student Definitions by Connector 54](#_Toc107300855)

[Attachment C: Final Evaluation Form 58](#_Toc107300856)

[Attachment D: Nondisclosure Agreement Form 60](#_Toc107300857)

List of Tables

[Acronyms and Initialisms Used in the Standard Setting Technical Report for the Alternate ELPAC iv](#_Toc109039490)

[Table 1. Number of Panelists 6](#_Toc109039491)

[Table 2. Panelist Gender 6](#_Toc109039492)

[Table 3. Panelist Primary Ethnicity or Race 7](#_Toc109039493)

[Table 4. Geographical Region of Panelists 7](#_Toc109039494)

[Table 5. Panelist Teaching Experience with Students 8](#_Toc109039495)

[Table 6. Panelist Years of Experience in Special Education 8](#_Toc109039496)

[Table 7. Panelist Grade Level(s) Currently Teaching 9](#_Toc109039497)

[Table 8. Panelist Experience Working with the California 2012 ELD Standards 9](#_Toc109039498)

[Table 9. Panelist Experience Working with the ELD Connectors 10](#_Toc109039499)

[Table 10. Number of Panelists in the Vertical Articulation Meeting 18](#_Toc109039500)

[Table 11. Mean Raw Score Threshold Scores at the End of Each Round: Kindergarten 21](#_Toc109039501)

[Table 12. Mean Raw Score Threshold Scores at the End of Each Round: Grade One 21](#_Toc109039502)

[Table 13. Mean Raw Score Threshold Scores at the End of Each Round: Grade Two 21](#_Toc109039503)

[Table 14. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Three Through Five 21](#_Toc109039504)

[Table 15. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Six Through Eight 21](#_Toc109039505)

[Table 16. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Nine Through Twelve 22](#_Toc109039506)

[Table 17. Raw Score SEJs by Round: Kindergarten 22](#_Toc109039507)

[Table 18. Raw Score SEJs by Round: Grade One 22](#_Toc109039508)

[Table 19. Raw Score SEJs by Round: Grade Two 22](#_Toc109039509)

[Table 20. Raw Score SEJs by Round: Grade Span Three Through Five 22](#_Toc109039510)

[Table 21. Raw Score SEJs by Round: Grade Span Six Through Eight 22](#_Toc109039511)

[Table 22. Raw Score SEJs by Round: Grade Span Nine Through Twelve 23](#_Toc109039512)

[Table 23. Kindergarten Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 23](#_Toc109039513)

[Table 24. Grade One Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 23](#_Toc109039514)

[Table 25. Grade Two Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 23](#_Toc109039515)

[Table 26. Grade Span Three Through Five Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 24](#_Toc109039516)

[Table 27. Grade Span Six Through Eight Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 24](#_Toc109039517)

[Table 28. Grade Span Nine Through Twelve Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs 24](#_Toc109039518)

[Table 29. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Kindergarten 24](#_Toc109039519)

[Table 30. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade One 25](#_Toc109039520)

[Table 31. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Two 25](#_Toc109039521)

[Table 32. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Three Through Five 25](#_Toc109039522)

[Table 33. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Six Through Eight 25](#_Toc109039523)

[Table 34. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Nine Through Twelve 25](#_Toc109039524)

[Table 35. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Kindergarten 26](#_Toc109039525)

[Table 36. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade One 26](#_Toc109039526)

[Table 37. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Two 26](#_Toc109039527)

[Table 38. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Three Through Five 27](#_Toc109039528)

[Table 39. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Six Through Eight 27](#_Toc109039529)

[Table 40. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Nine Through Twelve 27](#_Toc109039530)

[Table 41. Final Evaluation on the Usefulness of Materials, Kindergarten 30](#_Toc109039531)

[Table 42. Final Evaluation on the Influence of Process Components, Kindergarten 30](#_Toc109039532)

[Table 43. Final Evaluation on Timing, Kindergarten 31](#_Toc109039533)

[Table 44. Final Evaluation on the Usefulness of Materials, Grade One 31](#_Toc109039534)

[Table 45. Final Evaluation on the Influence of Process Components, Grade One 32](#_Toc109039535)

[Table 46. Final Evaluation on Timing, Grade One 32](#_Toc109039536)

[Table 47. Final Evaluation on the Usefulness of Materials, Grade Two 33](#_Toc109039537)

[Table 48. Final Evaluation on the Influence of Process Components, Grade Two 33](#_Toc109039538)

[Table 49. Final Evaluation on Timing, Grade Two 34](#_Toc109039539)

[Table 50. Final Evaluation on the Usefulness of Materials, Grade Span Three Through Five 34](#_Toc109039540)

[Table 51. Final Evaluation on the Influence of Process Components, Grade Span Three Through Five 35](#_Toc109039541)

[Table 52. Final Evaluation on Timing, Grade Span Three Through Five 35](#_Toc109039542)

[Table 53. Final Evaluation on the Usefulness of Materials, Grade Span Six Through Eight 36](#_Toc109039543)

[Table 54. Final Evaluation on the Influence of Process Components, Grade Span Six Through Eight 36](#_Toc109039544)

[Table 55. Final Evaluation on Timing, Grade Span Six Through Eight 37](#_Toc109039545)

[Table 56. Final Evaluation on the Usefulness of Materials, Grade Span Nine Through Twelve 37](#_Toc109039546)

[Table 57. Final Evaluation on the Influence of Process Components, Grade Span Nine Through Twelve 38](#_Toc109039547)

[Table 58. Final Evaluation on Timing, Grade Span Nine Through Twelve 38](#_Toc109039548)

Acronyms and Initialisms Used in the Standard Setting Technical Report for the Alternate ELPAC

| Abbreviation | Definition |
| --- | --- |
| CAASPP | California Assessment of Student Performance and Progress |
| CDE | California Department of Education |
| CR | constructed response |
| CRT | Content Review Tool |
| CSEM | conditional standard errors of measurement |
| *DFA* | *Directions for Administration* |
| EL | English learner |
| ELD | English language development |
| ELP | English Language Proficiency |
| ELPAC | English Language Proficiency Assessments for California |
| PLD | performance level descriptor |
| SBE | State Board of Education |
| SCOE | Sacramento County Office of Education |
| SEJ | standard error of judgment |
| TAG | Technical Advisory Group |

## Chapter 1: Introduction

The Alternate English Language Proficiency Assessments for California (ELPAC) consists of two assessments: the Initial Alternate ELPAC and the Summative Alternate ELPAC.

The assessments are administered to eligible students, as determined by each student’s individualized education program team. The purpose of the Alternate ELPAC is twofold:

1. The Initial Alternate ELPAC will provide information to determine a student’s initial classification as an English learner (EL) or as initial fluent English proficient, for students with the most significant cognitive disabilities. Identifying students who need help learning English is important so students can get the support they need to do well in school while receiving instruction in all school subjects.
2. The Summative Alternate ELPAC will provide information on annual student progress toward English language proficiency (ELP) and support decisions on whether or not a student is ready to be reclassified as fluent English proficient, for students with the most significant cognitive disabilities.

The Alternate ELPAC aligns with English language development (ELD) standards set out in the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) via the ELD Connectors. The ELD Connectors offer a reduction in the depth, breadth, and complexity of the standards, as appropriate for students with the most significant cognitive disabilities. In addition to being directed by the 2012 ELD Standards, the ELD Connectors were substantially informed by the ELP level descriptors in the Council of Chief State School Officers’ *English Language Proficiency (ELP) Standards for English Learners with Significant Cognitive Disabilities* (2019).

The Initial Alternate ELPAC will be administered once to eligible students entering the California school system for the first time, in the following grade levels and grade spans:

* Kindergarten
* One
* Two
* Three through five
* Six through eight
* Nine through twelve

The Summative Alternate ELPAC will be administered annually to eligible English learner students in the spring in the following grade levels and grade spans:

* Kindergarten
* One
* Two
* Three through five
* Six through eight
* Nine and ten
* Eleven and twelve

The Alternate ELPAC will report three performance levels—Level 1 through Level 3. These are described by performance level descriptors (PLDs) of four types: general PLDs, range PLDs, threshold PLDs, and reporting PLDs. Prior to the standard setting, Alternate ELPAC general PLDs were presented and approved at the May 2019 State Board of Education (SBE) meeting, as listed in the <https://www.cde.ca.gov/ta/tg/ep/documents/proposedhltdaltelpac.pdf>. These general, or policy, PLDs convey the degree of a student’s ELP using three levels: Novice English Learner (Level 1), Intermediate English Learner (Level 2), and Fluent English Proficient (Level 3).

In fall 2020, range PLDs were reviewed by California educators and approved by the California Department of Education (CDE). Range PLDs further describe what students at each level know and can do. As one of the most critical parts of the standard setting process, participants referred to the range PLDs to define the threshold PLDs, which is the set of knowledge, skills, and abilities expected of threshold students who are at the entry point of each performance level for each grade level or grade span assessed.

The reporting PLDs provide descriptions of the performance levels on the student score reports; they are developed based on both the general and range PLDs. Reporting PLDs will be finalized in January 2022.

In February 2020, to develop threshold-score recommendations, ETS conducted a standard setting workshop for the Alternate ELPAC using Zoom and other remote applications. All items in the Alternate ELPAC item pool were considered in the process of standard setting. The Modified Angoff and Extended Angoff standard setting methods were applied, as appropriate. A standard setting plan was approved by the CDE on July 26, 2019, in preparation for the meetings.

For each grade level or grade span, the standard setting panel recommended threshold scores to indicate the score that must be earned for a student to reach the beginning (i.e., threshold) of two of the three performance levels (Level 2 and Level 3) for the Alternate ELPAC total score. California educators utilized the ELD Connectors, the Alternate ELPAC General PLDs (CDE, 2019), and the range PLDs.

This document provides the following information:

* The purpose of the standard setting workshop and a discussion of the work conducted prior to the workshop
* An overview of the standard setting methods implemented, including discussions of the Modified and Extended Angoff methods used to develop the overall score thresholds
* A description of the panels and materials used in the approach, an overview of the process before and during the workshop, and a discussion of the training
* The results, including summary data from the panel judgments and evaluations by the panelists

### Purpose and General Description of the Standard Setting Workshop

The purpose of standard setting for the Alternate ELPAC was to collect recommendations for the Alternate ELPAC threshold scores, which are the minimum scores at Level 2 and Level 3 for each grade level and grade span. The CDE and the Technical Advisory Group (TAG) reviewed these recommendations, along with additional data; these and the State Superintendent of Public Instruction recommendations were presented to the SBE in May 2022.

A review of the standard setting literature supports the need for attention to best practices (Brandon, 2004; Hambleton & Pitoniak, 2006; Tannenbaum & Katz, 2013), which include the following:

* A careful selection of panel members
* A sufficient number of panel members to represent varying perspectives
* Sufficient time devoted to developing a common understanding of the assessment domain
* Adequate training of panel members
* Development of a description of each performance level
* Multiple rounds of judgments
* Inclusion of data, where appropriate, to inform judgments

The approach used in this study adheres to the guidelines and best practices; specifically, the Modified Angoff and Extended Angoff standard setting methods. These methods allowed for the collection of panelist judgments for each item administered in 2021–22, thereby providing flexibility in the development of threshold scores for reporting the overall score for the Initial and Summative Alternate ELPAC.

## Chapter 2: Method

Chapter 2 includes the following:

* Descriptions of the Modified and Extended Angoff Methods of standard setting
* Descriptions of the standard setting panels

### Modified and Extended Angoff Methods

The Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006) is a probability-based standard setting method. For 1-point items, each panelist judged the item on the likelihood that the threshold student would answer the item correctly. Panelists made judgments using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. The lower the value, the less likely it is that the threshold student would answer the item correctly because the item is difficult for the threshold student. The higher the value, the more likely it is that the threshold student would answer the item correctly.

An Extended Angoffmethod (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used for the 2-point items. For these items, the task was to decide on the assigned score value that would most likely be earned by the threshold student for each constructed-response (CR) item. Panelists were asked to first review the definition of the threshold student and then to review the item and its scoring rubric. The rubric for an extended-response item defines, holistically, the quality of the evidence that would merit a response earning a particular score. The scoring rules for 2-point composite items describe what responses are required to achieve one point and what responses are required to achieve two points.

In standard setting, the critical components involve having a standard setting panel of experts who can provide appropriate consideration and judgments. The panel begins by becoming familiar with the test and considering the content assessed and the relative difficulty of the items. The test-familiarization stage also allows the panelists to experience the test in a manner that is similar to an operational test administration, which allows the panelists to get a sense of the test taker’s experience. After independently reviewing the assessment, the panelists discuss the content measured and the relative difficulty of the items.

Following a discussion about the test content and the students who would take the test, the panelists consider the expectations in the three PLDs to develop threshold student definitions. The panelists work together in small and large groups to draft and reach consensus on the Level 3 threshold student definition followed by the threshold student definition for Level 2. These definitions are the operational description of the threshold scores and are used by the panelists as they make two rounds of judgments.

Prior to making judgments, panelists are trained and have an opportunity to practice using training materials. Once the training is completed and all panelists have indicated on the training evaluations a readiness to proceed, the first round of independent judgments takes place without discussion. Before the Round 2 judgments take place, panelists are presented with feedback data on the panel judgments. Panelists discuss the rationales behind the item-level judgments as they make Round 2 judgments. Presenting more information prior to the second round of judgments allows the panelists to become more informed judges; they are not required to make any changes when entering Round 2 judgments. Any adjustments the panelists make to the judgments are refinements informed by new information, including the rationales of colleagues.

At the conclusion of the standard setting, a final evaluation was administered to obtain feedback concerning the panelists’ perspectives on the standard setting procedures, instructions, and materials. In addition to procedural feedback, the panelists also provide their opinions of the final recommended scores.

Once judgments had been made for all six grade-level and grade-span assessments, a final review across all threshold scores and assessments was held. Representatives from each of the six panels participated in the Vertical Articulation Meeting, which is described in the [*Vertical Articulation*](#_Vertical_Articulation) section of this document.

### Standard Setting Panels

A diverse group, representative of California educators familiar with instructing this student population, were recruited to participate as panelists in the standard setting sessions. In recruiting panelists, the goal was to include a range of California educators who were familiar with the ELD Connectors and who were engaged in the daily instruction of EL students with the most significant cognitive disabilities, and who understood and represented the diverse group of students eligible to take the Alternate ELPAC. It was important to include teachers working with these students, as these teachers provided a perspective on the knowledge, skills, and learning goals for the students taking the Alternate ELPAC. Finally, the educators who were selected indicated that they would administer the Alternate ELPAC operational field test between November 2021 and February 2022. Additional efforts were made to include educators who had other relevant experience, such as speech language pathologists or ELD teachers; however, to participate, they were required to take the training modules for administering the test.

Panelists were assigned to one of six panels of educators; each panel focused on one grade level or grade span, as shown in table 1. The targeted number of panelists from this population of educators was 8 to 10 per panel.

The decision on the panelists’ selection and panel assignment was made by the CDE. Panelists were notified after they were approved, and staff from the Sacramento County Office of Education (SCOE) assisted with any logistical arrangements. Panelists were required to sign a security agreement that notified them of the confidentiality of the materials used in the standard setting. On the last day of the meeting, panelists were asked to destroy all confidential materials.

From February 22 through February 25, three panels were assembled to work on the kindergarten, grade one, and grade two standard setting. From February 28 through March 3, three panels were assembled to work on the three grade-span assessments: grade spans three through five, six through eight, and nine through twelve. Some of the invited educators participated on more than one panel. The number of panelists, by grade-level or grade-span panel, is displayed in table 1.

Table 1. Number of Panelists

| Grade Level or Grade Span | N |
| --- | --- |
| Kindergarten | 8 |
| 1 | 8 |
| 2 | 6 |
| 3–5 | 9 |
| 6–8 | 11 |
| 9–12 | 11 |
| **Total** | **53** |

Because standard setting is based on expert judgment—informed by student performance data—it is important that panelists collectively reflect the diversity of the educators working with students who take the assessment. Special efforts were made to assemble panels that were representative of the geographic and socioeconomic diversity of California in general and the Alternate ELPAC educator population in particular. Panelists’ self-reported data was collected using the standard setting survey submitted at the end of the workshop (refer to [Appendix 1, Attachment C: Final Evaluation Form](#_Attachment_C:_Final)). The data provided is panel-specific and can be found in table 2 through table 9. Nine educators participated on more than one grade-level panel. Table 2 provides the distribution of the panel by gender.

Table 2. Panelist Gender

| Gender | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Female | 8 | 7 | 5 | 9 | 8 | 10 | 47 |
| Male | 0 | 1 | 1 | 0 | 3 | 1 | 6 |

Table 3 provides the educators’ responses regarding personal ethnic or racial background. Panelists were asked if they identified as Hispanic or Latino and then were asked to choose one or more race. The largest group represented was White (N = 39). Panelists were able to select more than one response to the question, “What is your primary ethnicity/race?” Five panelists declined to respond.

Table 3. Panelist Primary Ethnicity or Race

| Ethnicity or Race | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Hispanic or Latino | 4 | 2 | 1 | 3 | 4 | 3 | 17 |
| American Indian or Alaska Native | 0 | 0 | 2 | 0 | 1 | 2 | 5 |
| Asian | 0 | 1 | 1 | 1 | 0 | 2 | 5 |
| Black or African American | 0 | 2 | 1 | 0 | 1 | 1 | 5 |
| Native Hawaiian or Other Pacific Islander | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| White | 6 | 6 | 3 | 6 | 10 | 8 | 39 |
| Decline to state | 2 | 0 | 1 | 2 | 0 | 0 | 5 |

Table 4 presents the location in which California educators are teaching. A majority of the educators reported working in the southern region of California. All panelists responded.

Table 4. Geographical Region of Panelists

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Region | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| Central | 1 | 0 | 0 | 0 | 0 | 1 | 2 |
| Northern | 1 | 0 | 2 | 2 | 0 | 3 | 8 |
| Southern | 6 | 8 | 4 | 7 | 11 | 7 | 43 |

Table 5 presents the environment in which the educators currently teach; multiple responses were permitted. The table presents the teaching experience of the educators in each panel and across the standard setting workshop by the number of years taught. A majority of educators indicated experience teaching and working with students in special education classrooms at the moderate level (N = 43) and at the severe level (N = 36). A majority also indicated experience working with ELs (N = 39). All panelists responded.

Table 5. Panelist Teaching Experience with Students

| Experience | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Special Education: Moderate | 6 | 8 | 5 | 9 | 7 | 8 | 43 |
| Special Education: Severe | 6 | 4 | 5 | 8 | 6 | 7 | 36 |
| General Education | 2 | 4 | 3 | 3 | 5 | 6 | 23 |
| ELs | 5 | 7 | 5 | 7 | 7 | 8 | 39 |
| Other | 3 | 0 | 1 | 4 | 3 | 0 | 11 |

Table 6 presents the panelists’ years of experience working in special education in each panel and across the standard setting workshop. The responses indicate that all panels were primarily comprised of educators with more than 10 years of experience. Teachers with less experience were also represented in all panels.

Table 6. Panelist Years of Experience in Special Education

| Years of Experience | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1–3 | 0 | 1 | 1 | 0 | 2 | 0 | 4 |
| 4–6 | 1 | 0 | 0 | 0 | 1 | 0 | 2 |
| 7–10 | 0 | 2 | 0 | 2 | 0 | 1 | 5 |
| 10+ | 7 | 5 | 5 | 6 | 7 | 9 | 39 |
| Not applicable | 0 | 0 | 0 | 1 | 1 | 1 | 3 |

Table 7 presents the grade level(s) educators were currently teaching; multiple responses were permitted. The responses show that all panels included a majority of educators who were currently teaching the grade level(s), or the adjacent grade level(s), corresponding to the Alternate ELPAC panel to which they were assigned.

Table 7. Panelist Grade Level(s) Currently Teaching

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level Currently Teaching | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| Kindergarten | 3 | 2 | 2 | 1 | 1 | 3 | 12 |
| Grade 1 | 1 | 2 | 2 | 1 | 0 | 3 | 9 |
| Grade 2 | 1 | 1 | 2 | 1 | 0 | 3 | 8 |
| Grade 3 | 1 | 1 | 4 | 4 | 0 | 2 | 12 |
| Grade 4 | 1 | 2 | 4 | 6 | 0 | 2 | 15 |
| Grade 5 | 1 | 1 | 4 | 5 | 0 | 2 | 13 |
| Grade 6 | 2 | 2 | 3 | 2 | 5 | 2 | 16 |
| Grade 7 | 1 | 2 | 2 | 0 | 5 | 2 | 12 |
| Grade 8 | 0 | 2 | 2 | 0 | 4 | 2 | 10 |
| High school | 0 | 3 | 3 | 2 | 2 | 8 | 18 |

Table 8 presents the educators’ years of experience working with the California 2012 ELD Standards. The responses indicate that all panels included educators who had been working with the standards for more than four years.

Table 8. Panelist Experience Working with the California 2012 ELD Standards

| Years | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Not applicable | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1–3 | 0 | 0 | 1 | 1 | 0 | 1 | 3 |
| 4–6 | 5 | 1 | 0 | 0 | 2 | 2 | 10 |
| 7–10 | 3 | 4 | 2 | 6 | 6 | 5 | 26 |
| More than 10 | 0 | 3 | 3 | 2 | 3 | 3 | 14 |

Table 9 presents the educators’ years of experience working with the ELD Connectors. The responses indicate that all panels included educators who had been working with the connectors for more than two years.

Table 9. Panelist Experience Working with the ELD Connectors

| Years | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–‍5 | Grade Span 6–‍8 | Grade Span 9–‍12 | Total |
| --- | --- | --- | --- | --- | --- | --- | --- |
| None | 0 | 1 | 0 | 2 | 2 | 1 | 6 |
| 1 | 3 | 3 | 1 | 1 | 0 | 4 | 12 |
| 2 | 1 | 0 | 2 | 0 | 1 | 1 | 5 |
| More than 2 | 4 | 4 | 3 | 6 | 8 | 5 | 30 |

## Chapter 3: Materials

Materials were provided to panelists for each assessment, for all standard setting activities. Nonsecure premeeting materials were sent to panelists’ email addresses; other materials were provided during the workshop when needed for each stage of the process. Materials used during the workshop were provided using remote tools, such as the ETS Content Review Tool (CRT) and Zoom web conference software. Materials included the following:

* Workshop agenda
* Alternate ELPAC range PLDs, provided in the premeeting materials as part of preparation for the workshop
* Training materials, provided as prework (i.e., Practice and Training website link, the *Directions for Administration* (*DFAs*), and the scoring rubric for CR items)
* Alternate ELPAC operational field test materials
  + Provided to panelists during the workshop through the use of ETS’ CRT (All unique items administered for the assessment were provided to the panel.)
  + Provided to panelists with an answer key with scoring rules and rubrics, where needed
* Judgment materials
  + Used during the workshop, forms unique to each panelist and provided electronically and remotely (Judgment materials were provided for both training and operational steps in the standard setting process.)
* Standard setting evaluation forms: training evaluation forms and final evaluation form, provided electronically

Panelists developed threshold student definitions in the workshop; refer to [attachment B](#_Attachment_B:_Final) in [appendix 1](#_Appendix_1:_ELA). The judgment training materials and evaluation forms are described more fully in the next subsection.

### Judgment Training Materials

The purpose of the judgment training is to familiarize the panelists with the types of judgments to be made and the online judgment forms. During the standard setting training, the panelists were introduced to the online forms and, as a practice exercise, completed judgments on a sample of items from the Alternate ELPAC. The training process, which includes learning how to make judgments on online forms and for each of the judgment types, is described in [chapter 4](#_Chapter_4:_Process), under [*Preparation and Training*](#_Preparation_and_Training).

### Evaluation Forms

It is important to collect information from the panelists to document procedural validity (Cizek & Bunch, 2007; Hambleton & Pitoniak, 2006). Panelists received training evaluation forms at four points during the process to gauge the panelists’ understanding of the procedures. A final evaluation was included to gather other information (refer to [attachment C](#_Attachment_D:_Evaluation) in [appendix 1](#_Appendix_1:_ELA) for the evaluation forms). Final evaluations included questions about the following:

* Training
* Understanding the tasks
* The influence of different aspects of the standard setting process

Because ETS was interested in knowing as soon as possible if panelists were not satisfied with the level of training provided, four training evaluation forms were given to the panelists at the end of each step in the training to gauge the panelists’ current understanding of the process and comfort level with the tasks to be performed. These training evaluation forms were delivered as Zoom polls and analyzed immediately; responses were reviewed by the panel facilitator and lead facilitator so the facilitators could review any tasks or materials that appeared to be unclear. The results obtained through the training polls is included in the [*Evaluation Results from the Training Evaluation of the Standard Setting Process*](#_Evaluation_Results_from) section of this report.

## Chapter 4: Process

This section of the report describes what occurred prior to, and during, the standard setting workshop.

### Preparation and Training

Prior to the standard setting, a preworkshop assignment, consisting of three parts, was given to the panelists two weeks before the workshop (refer to [attachment A](#_Attachment_A:_Panelist)). For the first part, panelists were provided with a link to the Alternate ELPAC practice test, the *DFA*, and the answer key and were asked to take a training test to become familiar with the task types on the Alternate ELPAC. The assignment included directions for finding the answer key to the training test items. The second part of the assignment included reading the rangePLDs for the panelists’ assigned grade level or grade span and take notes. Specifically, panelists were asked to consider the expected performance of a student in each of the performance levels, take notes about the knowledge and skills of students *at the beginning* of Level 3, and bring those notes to the standard setting workshop. The third part of the assignment was to follow directions to register for the Zoom meeting.

### General Overview of Training and Process

On February 22, the panelists assigned to the kindergarten, grade one, and grade two grade-level-specific panels joined the general session of the workshop. Similarly, in week two, panelists assigned to work on the three grade-span assessments (i.e., grade spans three through five, six through eight, and nine through twelve) joined the general session on February 28. In this first session, Dr. Patricia Baron, ETS standard setting director, provided a general introduction to the purpose of standard setting and to the assessment development process for the Alternate ELPAC. The general session included an explanation of the need for threshold scores and an explanation of the roles and responsibilities of the staff from the CDE, ETS, and SCOE.

Dr. Baron introduced the Modified Angoff and Extended Angoff Methods for setting threshold scores and presented the agenda and expectations for panel members’ participation, as well as the initial training on the methods. Panelists next moved into test-specific groups in Zoom web conference panel rooms (hereafter referred to as panel rooms) where the panel facilitators continued with training and guided the panelists through the rest of the standard setting activities. Panelists were encouraged to ask questions during the workshop and were reminded of the confidentiality and security of materials. Staff from the CDE were available throughout the process to answer questions about the policies surrounding the test; staff from ETS were available to answer questions about the test and the standard setting procedures.

Each week, individual panels worked in the panel rooms, with a panel facilitator, on one grade-level or grade-span assessment. All facilitators had multiple years of experience working with educators in standard setting. Facilitators provided in-depth training and practice on the method, for each step of the process. Panelists completed two rounds of judgments; feedback and discussion took place after each round of judgment. Each panel completed the standard setting process on one grade-level or grade-span assessment. At the end of the workshop, panelists completed a final evaluation survey. On March 3, the last day of the workshop in week two, representatives from each panel reviewed the process and recommendations across all six panels during a two-hour vertical articulation meeting (refer to the [*Vertical Articulation*](#_Vertical_Articulation) section of this report).

### Test Familiarization

Each panel worked independently on their assigned test. The first activity was for the panelists to “take the test.” An ETS assessment specialist acted as a test administrator by presenting the test items to all panelists in a secure environment provided through the Zoom registration. The panelists accessed all materials in the same way students taking the Alternate ELPAC would. The purpose of this activity is to allow panelists to become familiar with the Alternative ELPAC items and discuss the test content before making any judgments. Panelists were reminded that this activity is not an item review; rather, the purpose is to share thoughts about what the Alternate ELPAC is measuring, what panelists think may be particularly challenging for students, and what might be less difficult. The goal of the discussion was for panelists to share with each other perceptions of the general difficulty of the tested material for students.

Once the panelists were familiar with the assessment, the group began with a draft threshold definition for two connectors.

### Threshold Student Definitions

Developing definitions of threshold students is a critical component of any standard setting workshop. The process to arrive at threshold student definitions involved whole-panel training on the process, modeled by the facilitator; this was followed by small-group discussions and the development of draft definitions. The process was completed in a whole-panel discussion of the draft definitions, to reach a panel consensus on what is expected.

For the Alternate ELPAC, two definitions were needed for two thresholds—the Level 2 and Level 3 threshold student definitions. Panels worked first on the Level 3 threshold, because this is the point at which students are classified as being eligible for classification as fluent English proficient.

The first step was a facilitated discussion to begin to articulate the knowledge and skills necessary for students to reach performance Level 3, using the connectors, the range PLDs, and samples of threshold student definitions for two connectors. The panel facilitators guided the discussions, and the panel worked to edit and refine the threshold definitions for Level 3 and Level 2 using the two drafts provided. This modeling process represents the initial training for panelists to define the threshold students. They reached consensus as a whole panel on those two connectors that were included as drafts and moved to small-group discussions via Zoom breakout rooms to finish developing the Level 2 and Level 3 threshold student definitions. The definitions were finalized by a whole-panel, consensus-building process.

The threshold student definitions are provided in [attachment B](#_Attachment_B:_Final) in [appendix 1](#_Appendix_1:_ELA).

### Panelist Judgments

Prior to the start of actual standard setting rounds of judgments on operational items, as described in the [*Modified and Extended Angoff Judgments*](#_Modified_and_Extended) subsection, panelists were trained and then practiced making judgments on three 1-point and three 2-point item types. The practice round included a summary of the judgments and a discussion of rationales. After training, panelists were asked to complete a training evaluation survey confirming an understanding of the procedures and a readiness to proceed.

For each of the performance levels, panelists made two rounds of judgments on operational items. Round 1 judgments were made independently, without discussion. Panelists’ judgments were collected, analyzed, summarized, and shared with the panel; panelists reviewed the panel summary data and participated in a discussion. Before proceeding with the next step, facilitators provided training on how to include feedback and make Round 2 judgments on the online judgment forms; after all educators on the panel indicated a readiness to proceed, feedback from Round 1 was provided. As part of the post-Round 1 feedback, panelists next reviewed the individual judgments in the context of the range of judgments across the panel, and the facilitator shared feedback on the similarity and differences of the panel judgments on the Alternate ELPAC items. Panelists discussed with the other panelists the rationales for the independent judgments.

The feedback and discussion from the Round 1 judgment data then informed Round 2 judgments. Panelists were invited to make any changes to their item-level judgments and enter changes into the Round 2 data collection form; they were informed that they did not have to make any changes to their Round 1 judgments, but had this opportunity to do so, after hearing the panel discussion. After this round, panelists’ judgments were collected, analyzed, summarized, and shared with the panel as the final panel recommendation.

### Item Scoring, Judgments, and Rating Scales

There are multiple item types on the Alternate ELPAC, scored as 1-point or 2-point items. One important goal in standard setting is to reduce the cognitive complexity of making judgments; instructions to the panelists need to be clear and understandable. The more difficult the judgment task, the less accurate (and meaningful) is the panelist’s decision. Instructions and judgments are more intuitive for the panelists when the ratings are aligned to the scoring rules. This was considered in discussions with the CDE and their TAG and incorporated in the selection of standard setting methods.

ETS implemented the standard setting using two standard setting methods:

1. Modified Angoff for 1-point items
2. Extended Angoff for 2-point items

Using these two methods allowed panelists to make judgments aligned with the scoring of the items. Scoring rubrics were provided to panelists, along with the answer key.

### Modified and Extended Angoff Judgments

For items scored as 1-point items, the Modified Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006) was used; for 2-point items, the Extended Angoff method (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used. The 1-point items include discrete items types such as multiple-choice selected response. The 2-point items include extended response with scoring rules that indicate how a student will obtain a score of 0, 1, or 2.

The Modified Angoff method is a probability-based standard setting method. For 1-point items, each panelist was asked to judge the item on the likelihood that the threshold student would answer the item correctly. Panelists made a judgment using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. On this scale, the lower the value, the less likely it is that the threshold student would answer the item correctly because the item is difficult for the threshold student. The higher the value, the more likely it is that the threshold student would answer the item correctly.

The facilitator suggested to the panelists that the judgment process be approached in two stages. The first stage involved reviewing both the description of the threshold student and the item and then considering the probability that the threshold student would answer the question correctly. The facilitator encouraged the panelists to use the following guidelines to inform this decision:

* Items in the 0 to .30 range are those that the threshold student would have a low chance of answering correctly.
* Items in the .40 to .60 range are those that the threshold student would have a moderate chance of answering correctly.
* Items in the .70 to 1 range are those that the threshold student would have a high chance of answering correctly.

In the second stage, the task was to refine the judgment within the range. For example, if a panelist thought that there is a high chance that the threshold student would answer the question correctly, the initial decision would be in the .70 to 1 range. The second decision for the panelist was to judge if the likelihood of a threshold student answering it correctly is .70, .80, .90, .95, or 1.

Panelists were asked to make two judgments for each item. The overall instructions included a reminder that, when making Level 2 (L2) and Level 3 (L3) judgments, it was expected that each judgment value must be at least the same as the value of the level below, for each item. For example, if the threshold L2 judgment was .30, then the threshold L3 judgment had to be .30 or higher. For Extended Angoff judgments, the same applied: the threshold L3 judgment had to be the same, or higher, than the L2 judgments. Note that the judgments were made in an excel tool, formatted so that L3 judgments would be the same or higher than L2 judgments.

An Extended Angoffmethod (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used for the 2-point items. For these items, a panelist decided on the assigned score value that would most likely be earned by the threshold student for each CR item. Panelists were asked to first review the definition of the threshold student and then to review the item and its scoring rubric. The rubric for an extended-response item indicates how a student will obtain a score of 0, 1, or 2.

During this review, each panelist independently considered the level of knowledge and skill required to respond to the item, as well as the features of a response that would earn a particular score as defined by the scoring rubric. Each panelist decided on the score most likely to be earned by each threshold student from the possible values a student can earn. Panelists were reminded to refer to the knowledge and skills of the threshold student definition and the scoring rules and not to expect the three performance levels to match to the three possible scores.

### Feedback and Discussion

The purposes of feedback and discussion were to allow panelists to hear the rationales of the other panelists and arrive at a mutual understanding of the expectations of threshold students’ performance on this test. The process of judgment, feedback, and discussion was repeated over the entire standard setting workshop until all threshold score recommendations were collected.

Panelists were provided with the judgments for all items and received feedback after Round 1 judgments were collected and summarized. The mean, minimum, maximum, and range of the panel judgments (from low to high) were projected to the room. Panelists reviewed item-level judgment information for each item; the facilitator projected, for the room, a presentation that identified where panelists were closer to consensus in the judgments for some items and where judgments were more diverse for other items. Panelists were encouraged to discuss the judgments and rationales. Panelists made notes on the judgment forms and entered independent Round 2 judgments for any items they wanted to change.

After making Round 2 independent judgments, results were again projected in each panel room, including summary statistics of the panel’s threshold scores and the panel’s range of judgments. The panelists discussed the Round 2 data and compared it to the mean and range from Round 1 summary statistics. After the discussion, panelists responded to a final, confidential evaluation form. Notes from the Round 2 discussion were included in the Vertical Articulation Meeting, as described in the next section.

### Vertical Articulation

Panelists from each of the six grade-level and grade-span tests were asked to participate in a two-hour meeting to evaluate the process and recommendations across all six pairs of threshold score judgments. Twenty-five panelists participated in a discussion, facilitated by the standard setting director on the last day of the standard setting, March 3, 2022. Representatives from all panels participated in the discussion (refer to table 10). Panel representatives described the process that transpired, including reaching consensus on threshold student definitions, discussions about differences in judgments, and rationales. Panelists discussed the similarities and differences in the Level 3 threshold student definitions starting with kindergarten and proceeding through to grade span nine through twelve. Specifically, representatives in adjacent grades described the progression of expectations as the grades increased (e.g., panelists from grades one and two; panelists from grade two and grade span three through six). There was general agreement that the threshold student definitions for entering Level 3 appropriately described the expectations as the grades increased.

Table 10 presents the number of panelists in the Vertical Articulation Meeting.

Table 10.  Number of Panelists in the Vertical Articulation Meeting

| Grade Level or Grade Span | N |
| --- | --- |
| Kindergarten | 4 |
| 1 | 4 |
| 2 | 4 |
| 3–5 | 3 |
| 6–8 | 6 |
| 9–12 | 4 |
| **Total** | **25** |

Once these discussions concluded, the facilitator described the impact data available, with appropriate caveats. The percentages of students classified across kindergarten, grade one, and grade two were first displayed and discussed. The interpretation of the data was straightforward, since there was only one form administered to these grades during the Alternate ELPAC operational field test administration; all students took the same set of items in the administration of each of these three grade-level tests. By applying the panel raw score threshold scores to the total score based on the items administered to all students in each of the three lower grade levels, it was possible to interpret the impact data directly. Panelists discussed the data and noted that the percentages across the three levels were reasonable; they noted that the percentage of students in Level 1 was similar in kindergarten and grade two, whereas the percentage in Level 1 for grade one was lower (refer to table 29 through table 31).

The facilitator next presented the impact data for the three upper grade assessments. However, because students were administered two field test forms in those grade spans, impact data could not be compared directly, as this required additional analyses to allow comparisons across all students in those grades. Because of the timing of the field test assessment and the standard setting, the calibration analysis and scaling were conducted simultaneously with the standard setting workshop. The facilitator provided the impact data for each form based on raw scores, for each of the three grade-span assessments, and explained that these results must be reviewed with caution. There was some variation across the two field test forms in each grade-span assessment. The facilitator explained that once the scaling process is conducted, the percentages in each level will vary, but the direction of the variation cannot be predicted; the presentation of data concluded, and panelists continued to discuss their expectation of student performance more generally. Panelists in the Vertical Articulation Meeting noted that, along with their review of the results for kindergarten through grade two, it seemed reasonable that the impact data for the upper grades would show higher percentages of students in Level 3, compared to the lower grade levels, because of more instructional time. *The impact data presented in table 29 through table 34 are results after the calibration and scaling analyses were conducted.*

In concluding comments, panelists in the vertical articulation panel stated that the comparisons of threshold student definitions across the three grade-level and grade-span assessments showed the appropriate expectation of learning progressions. They further articulated an appreciation of the standard setting process, and many commented that participating in this process helped them to understand more about the development and use of the Alternate ELPAC. The CDE observed the process, including the Vertical Articulation Meeting. The feedback from the panelists, which indicates that the process was conducted similarly across panels, provides additional input when considering final recommendations for the threshold scores for this assessment.

## Chapter 5: Results

This section describes the results from the workshop, which include the item judgments and total score recommendations, and an evaluation of the process based on questionnaires completed by the panelists. Also presented are the projected distributions based on Round 2 recommendations and the range of scores using conditional standard errors of measurement (CSEM) using the standard setting scale metric. Because of the timing of the field test assessment and the standard setting, no impact data was presented to panelists prior to the final panel judgments. Data types 4 and 5, listed in the next section, were created after the calibration and scaling, which allowed data across field test forms to be combined and total scores to be reported for all students on the scale score metric. This data was presented to the CDE after standard setting for consideration.

Data for each panel is presented in this chapter. Five types of tables are presented; a general description of the five types follows.

### Five Types of Data Tables

1. **Mean threshold scores, by round:** Mean raw score threshold scores are presented in table 11 through table 16. The range of possible scores is equal to the number of possible points in the pool of items administered in 2021–22 to students in each grade level or grade span; the test includes 1-point and 2-point items.
2. **Standard errors of judgment (SEJs), by round:** SEJs are presented in the raw score metric, based on the panelists’ judgments, in table 17 through table 22.
3. **Round 2 raw score judgments +/-1 SEJ and +/-2 SEJs:** The range around the final recommended threshold score is presented in table 23 through table 28.
4. **Projected distribution of 2021–22 Alternate ELPAC students, shown as the percent, at each level based on the Round 2 recommended threshold scores on the standard setting scale score metric:** Refer to table 29 through table 34.
5. **Projected percentage of 2021–22 Alternate ELPAC students at or above the Round 2 recommended threshold score, +/-1 CSEM, and +/-2 CSEM scores on the standard setting scale score metric:** Refer to table 35 through table 40.

### Data Presentation

Panel threshold score recommendations were presented to panelists as a raw score. The raw score is based on judgments on all items administered in the 2021–22 Alternate ELPAC operational administration. The test administration model for the Alternate ELPAC included field test items, which meant that in those grades tested with more than one form of the test, not all students were administered the same set of items. In the lower grade levels (kindergarten, grade one, and grade two), only one form was administered.

When panelists considered the impact data, this feedback was provided on a scale that was based on the underlying theta distribution of the total score for all students who took the Alternate ELPAC. ETS transformed the raw scale to a scale score unique to each grade level or grade span, with a range of approximately 100 points, via a linear translation of the theta scale. All scale score information included in this technical report is based on the working scale—the Standard Setting Scale.

### Alternate ELPAC Threshold Score Results

Table 11 through table 16 display the mean threshold scores after each round for each test. These raw scores were based on the complete item pool and are rounded up to the nearest whole number. The tables show how panelists moved the judgments across rounds. Lower numbers indicate a lower threshold score. Higher numbers translate to a higher threshold score; a higher threshold score means that more is required for a student to be included in the level.

For five of the six Alternate ELPAC tests, the mean raw threshold score remained the same or decreased from Round 1 to Round 2 for both levels. In most cases, the decrease was one point. For kindergarten at Level 3 and for grade one Level 2, the Round 2 threshold score decreased by two points. For grade span three through five, the Level 2 threshold score increased by one point, and the Level 3 threshold score decreased by one point.

Table 11. Mean Raw Score Threshold Scores at the End of Each Round: Kindergarten

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 16 | 15 |
| Level 3 | 33 | 31 |

Table 12. Mean Raw Score Threshold Scores at the End of Each Round: Grade One

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 18 | 16 |
| Level 3 | 32 | 32 |

Table 13. Mean Raw Score Threshold Scores at the End of Each Round: Grade Two

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 21 | 20 |
| Level 3 | 33 | 32 |

Table 14. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Three Through Five

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 21 | 22 |
| Level 3 | 44 | 43 |

Table 15. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Six Through Eight

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 27 | 26 |
| Level 3 | 46 | 45 |

Table 16. Mean Raw Score Threshold Scores at the End of Each Round: Grade Span Nine Through Twelve

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 34 | 34 |
| Level 3 | 50 | 50 |

Table 17 through table 22 provide estimates of the SEJ after each round by panel. The SEJ is one way of estimating the reliability or consistency of a panel’s standard setting judgments. Lower numbers from Round 1 to Round 2 indicate the convergence of panelists’ judgments over rounds during the process. Ideally, the SEJ should decrease across rounds; however, the diversity of educators working with students taking the Alternate ELPAC is a consideration. An SEJ assumes that panelists are randomly selected and that standard setting judgments are independent. It is seldom the case that panelists are randomly sampled, and only the first round of judgments may be considered independent. The SEJ, therefore, likely underestimates the uncertainty of passing scores (Tannenbaum & Katz, 2013).

In all six panels, for Level 3, and in five panels for Level 2, the SEJs decreased over two rounds. In kindergarten, for Level 2, the SEJ increased by .02 in Round 2.

Table 17. Raw Score SEJs by Round: Kindergarten

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 0.61 | 0.63 |
| Level 3 | 0.85 | 0.30 |

Table 18. Raw Score SEJs by Round: Grade One

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 2.79 | 1.71 |
| Level 3 | 1.50 | 1.02 |

Table 19. Raw Score SEJs by Round: Grade Two

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 1.78 | 1.76 |
| Level 3 | 2.32 | 1.76 |

Table 20. Raw Score SEJs by Round: Grade Span Three Through Five

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 1.72 | 1.51 |
| Level 3 | 1.05 | 1.00 |

Table 21. Raw Score SEJs by Round: Grade Span Six Through Eight

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 2.77 | 1.98 |
| Level 3 | 2.03 | 1.55 |

Table 22. Raw Score SEJs by Round: Grade Span Nine Through Twelve

| Level | Round 1 | Round 2 |
| --- | --- | --- |
| Level 2 | 2.07 | 1.55 |
| Level 3 | 1.14 | 1.13 |

Table 23 through table 28 provide the panel-recommended threshold score +/- 1 SEJ and +/- 2 SEJs by panel. Ranges around the panel-recommended threshold score are rounded to one decimal.

Table 23. Kindergarten Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 13.7 | 30.4 |
| -1 SEJ | 14.4 | 30.7 |
| Panel recommended | 15.0 | 31.0 |
| +1 SEJ | 15.6 | 31.3 |
| +2 SEJ | 16.3 | 31.6 |

Table 24. Grade One Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 12.6 | 30.0 |
| -1 SEJ | 14.3 | 31.0 |
| Panel recommended | 16.0 | 32.0 |
| +1 SEJ | 17.7 | 33.0 |
| +2 SEJ | 19.4 | 34.0 |

Table 25. Grade Two Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 16.5 | 28.5 |
| -1 SEJ | 18.2 | 30.2 |
| Panel recommended | 20.0 | 32.0 |
| +1 SEJ | 21.8 | 33.8 |
| +2 SEJ | 23.5 | 35.5 |

Table 26. Grade Span Three Through Five Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 19.0 | 41.0 |
| -1 SEJ | 20.5 | 42.0 |
| Panel recommended | 22.0 | 43.0 |
| +1 SEJ | 23.5 | 44.0 |
| +2 SEJ | 25.0 | 45.0 |

Table 27. Grade Span Six Through Eight Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 22.0 | 41.9 |
| -1 SEJ | 24.0 | 43.5 |
| Panel recommended | 26.0 | 45.0 |
| +1 SEJ | 28.0 | 46.6 |
| +2 SEJ | 30.0 | 48.1 |

Table 28. Grade Span Nine Through Twelve Round 2 Recommended Threshold Scores +/- 1 SEJ and +/- 2 SEJs

| Threshold | Level 2 Raw Score | Level 3 Raw Score |
| --- | --- | --- |
| -2 SEJ | 30.9 | 47.7 |
| -1 SEJ | 32.5 | 48.9 |
| Panel recommended | 34.0 | 50.0 |
| +1 SEJ | 35.6 | 51.1 |
| +2 SEJ | 37.1 | 52.3 |

Table 29 through table 34 provide the percentage of students projected in each performance level based on the 2021–22 distribution applied to the panel-recommended threshold scores.

Table 29. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Kindergarten

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 39.6 |
| Level 2 | 254 | 44.0 |
| Level 3 | 273 | 16.4 |

Table 30. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade One

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 30.1 |
| Level 2 | 249 | 49.9 |
| Level 3 | 269 | 19.9 |

Table 31. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Two

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 43.3 |
| Level 2 | 253 | 36.4 |
| Level 3 | 270 | 20.3 |

Table 32. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Three Through Five

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 21.4 |
| Level 2 | 240 | 40.5 |
| Level 3 | 260 | 38.0 |

Table 33. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Six Through Eight

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 18.1 |
| Level 2 | 234 | 28.6 |
| Level 3 | 253 | 53.4 |

Table 34. Projected Distribution of 2021–22 Students Based on Round 2 Recommendation: Grade Span Nine Through Twelve

| Performance Level | Threshold Score | Percentage |
| --- | --- | --- |
| Level 1 | N/A | 30.0 |
| Level 2 | 240 | 24.6 |
| Level 3 | 258 | 45.5 |

Table 35 through table 40 provide the percentage of students projected at or above Level 2 and Level 3 based on the 2021–22 distribution applied to the panel-recommended threshold scores, +/- 1 and +/- 2 CSEM by panel.

Table 35. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Kindergarten

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 245 | 69.6 | 264 | 35.9 |
| -1 CSEM | 250 | 64.1 | 269 | 23.0 |
| Panel recommended | 254 | 60.4 | 273 | 16.4 |
| +1 CSEM | 258 | 48.8 | 277 | 10.6 |
| +2 CSEM | 263 | 39.5 | 282 | 6.5 |

Table 36. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade One

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 240 | 80.3 | 260 | 41.5 |
| -1 CSEM | 245 | 75.2 | 264 | 30.7 |
| Panel recommended | 249 | 69.9 | 269 | 19.9 |
| +1 CSEM | 253 | 61.5 | 274 | 14.6 |
| +2 CSEM | 258 | 48.7 | 278 | 11.3 |

Table 37. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Two

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 244 | 75.5 | 260 | 40.4 |
| -1 CSEM | 249 | 65.7 | 265 | 28.1 |
| Panel recommended | 253 | 56.7 | 270 | 20.3 |
| +1 CSEM | 257 | 47.0 | 275 | 12.9 |
| +2 CSEM | 262 | 34.2 | 280 | 8.7 |

Table 38. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Three Through Five

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 233 | 87.9 | 252 | 55.1 |
| -1 CSEM | 236 | 84.5 | 256 | 46.2 |
| Panel recommended | 240 | 78.6 | 260 | 38.0 |
| +1 CSEM | 244 | 72.4 | 264 | 33.1 |
| +2 CSEM | 247 | 64.8 | 268 | 27.7 |

Table 39. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Six Through Eight

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 227 | 87.9 | 245 | 64.6 |
| -1 CSEM | 230 | 86.0 | 249 | 58.1 |
| Panel recommended | 234 | 81.9 | 253 | 53.4 |
| +1 CSEM | 238 | 74.9 | 257 | 47.3 |
| +2 CSEM | 241 | 70.4 | 261 | 40.6 |

Table 40. Projected Percentage of 2021–22 Students at or Above the Recommended Threshold Score, +/-1 CSEM and +/-2 CSEM for Grade Span Nine Through Twelve

| Threshold | Level 2 Scale Score | Level 2 Percent at or Above | Level 3 Scale Score | Level 3 Percent at or Above |
| --- | --- | --- | --- | --- |
| -2 CSEM | 233 | 79.7 | 249 | 57.0 |
| -1 CSEM | 236 | 75.7 | 253 | 51.9 |
| Panel recommended | 240 | 70.0 | 258 | 45.5 |
| +1 CSEM | 244 | 66.1 | 263 | 38.9 |
| +2 CSEM | 247 | 61.8 | 267 | 34.3 |

### Incorporating Additional Considerations in Setting Threshold Scores

In standard setting, policymakers sometimes wish to reduce the number of examinees who fall below the panel-recommended threshold scores because of random error. In addition to measurement error metrics (e.g., CSEM, SEJ), policymakers should consider the likelihood of classification error; that is, when adjusting a threshold score, policymakers should consider whether it is more important to minimize a false-positive decision or to minimize a false-negative decision.

A false-positive decision occurs when a test taker’s score suggests one level of knowledge and skills, but the student’s actual level is lower (i.e., the student does not possess the required skills). A false-negative decision occurs when a test taker’s score suggests that the student does not possess the required skills, but that student nevertheless actually does possess those skills.

To reduce the number of false negatives, policymakers will decide to lower the threshold score(s). On the other hand, they may desire to reduce the number of test takers who attain a score above the recommended threshold score because of random error at each level to reduce the number of false positives and thus raise the threshold score(s).

Raising threshold scores reduces false positives but increases false negatives; the reverse occurs when threshold scores are lowered. Policymakers need to consider which decision error to minimize; it is not possible to eliminate both types of decision errors simultaneously.

### Evaluation Results from the Training Evaluation of the Standard Setting Process

Each panelist was asked at four points over the course of the workshop to respond to training polls prior to completing steps in the process. The questions asked about the panelists’ understanding of the following topics:

1. Threshold student definitions and how to use them in making judgments
2. Steps to complete standard setting judgments for the 1-point items
3. Summary data from Round 1 and how to make Round 2 judgments
4. Steps to complete standard setting judgments for the 2-point items

Panelists’ ratings were collected using Zoom polls. The purpose of the training polls, completed prior to each step in the process, was to provide an early check on the level of panelists’ understanding of the task and to identify any areas of confusion. Assessing the level of clarity prior to beginning the judgment process is essential to validating the overall standard setting process. Generally, panelists indicated understanding of each step following training across the six panels. One panelist in the grade span nine through twelve panel indicated on how to complete the 1-point item judgments; additional training was provided, after which the panelists indicated they were ready to proceed. In both the grade span six through eight and grade span nine through twelve panels, one panelist had questions on the process to enter Round 2 judgments into the tool and, after discussion with the panel facilitator, indicated a readiness to proceed.

### Evaluation Results from the Alternate ELPAC Standard Setting Final Evaluations

At the end of the workshop, a final evaluation form contained additional questions used to analyze the whole process, including the following topics:

* Training
* Usefulness of materials and procedures
* Influence of policy documents and work products
* Individual and group perceptions
* Student performance data
* Discussion

Results from the final evaluation forms are panel-based and are specific to each panel. There was no cross-panel discussion during the process of the standard setting workshop; therefore, any comparisons across panels should acknowledge the independence of the panels.

The evaluation forms are in [appendix 1](#_Appendix_1:_ELA): [attachment C](#_Attachment_D:_Evaluation).

Table 41 through table 58 provide the results of final evaluations. The results provide information about panelists’ thoughts as to the usefulness and influence of materials and other aspects of the three-day process. They also provide insight into each panelist’s stated belief as to the appropriateness of the threshold-score recommendations and whether the panelist could support them.

In the final evaluation, the majority of panelists indicated having a clear understanding of the standard setting process and indicated that the materials and processes were somewhat or very useful. Overall, panelists indicated that most of the process materials, data, and discussion were somewhat or very influential. Five panelists in the grade five panel and one panelist in the high school panel indicated that “Completing the pre-workshop assignment” was not at all useful (refer to [attachment A](#_Attachment_A:_Panelist)). In some panels, one panelist indicated that one aspect was not influential (e.g., “the percent of students in each performance level was not at all influential”).

The majority of panelists indicated that the amount of time for different components of the process was about right. However, panelists’ responses to the questions about the appropriate amount of time allowed for each step varied somewhat. In all three panels, one or more of the panelists indicated that there was either too little or too much time allotted to some aspect of the process. For example, in the grade five panel, two panelists indicated there was too little time for group discussion, and two panelists indicated there was too much time for group discussion. Experience indicates that variability in panelists’ sense of the training and process is expected and dependent on the characteristics and interactions of the panel.

In the vertical articulation panel, panelists provided comments and discussed their level of comfort with the threshold-score recommendations. They confirmed that the final recommendations for their own panel was appropriate, and they supported the panel recommendations.

Table 41. Final Evaluation on the Usefulness of Materials, Kindergarten

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 0 | 0 | 3 | 38 | 5 | 63 |
| Taking the test before making judgments | 0 | 0 | 4 | 50 | 4 | 50 |
| Defining the threshold students | 0 | 0 | 0 | 0 | 8 | 100 |
| Practicing the procedure | 0 | 0 | 0 | 0 | 8 | 100 |
| Group discussions | 0 | 0 | 0 | 0 | 8 | 100 |

Table 42. Final Evaluation on the Influence of Process Components, Kindergarten

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 1 | 13 | 7 | 88 |
| Threshold student definitions | 0 | 0 | 0 | 0 | 8 | 100 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 2 | 25 | 6 | 75 |
| My experience with the students | 0 | 0 | 0 | 0 | 8 | 100 |
| Group discussions | 0 | 0 | 0 | 0 | 8 | 100 |
| Judgments and rationales of other panelists | 0 | 0 | 1 | 13 | 7 | 88 |
| My sense of what students need to know to be proficient | 0 | 0 | 1 | 13 | 7 | 88 |

Table 43. Final Evaluation on Timing, Kindergarten

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 6 | 75 | 2 | 25 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 8 | 100 | 0 | 0 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 8 | 100 | 0 | 0 |
| Group discussion | 0 | 0 | 7 | 88 | 1 | 13 |

Table 44. Final Evaluation on the Usefulness of Materials, Grade One

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 0 | 0 | 4 | 50 | 4 | 50 |
| Taking the test before making judgments | 0 | 0 | 2 | 25 | 6 | 75 |
| Defining the threshold students | 0 | 0 | 1 | 13 | 7 | 88 |
| Practicing the procedure | 0 | 0 | 0 | 0 | 8 | 100 |
| Group discussions | 0 | 0 | 1 | 13 | 7 | 88 |

Table 45. Final Evaluation on the Influence of Process Components, Grade One

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 1 | 13 | 7 | 88 |
| Threshold student definitions | 0 | 0 | 1 | 13 | 7 | 88 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 3 | 38 | 5 | 63 |
| My experience with the students | 0 | 0 | 1 | 13 | 7 | 88 |
| Group discussions | 0 | 0 | 2 | 25 | 6 | 75 |
| Judgments and rationales of other panelists | 0 | 0 | 1 | 13 | 7 | 88 |
| My sense of what students need to know to be proficient | 0 | 0 | 3 | 38 | 5 | 63 |

Table 46. Final Evaluation on Timing, Grade One

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 5 | 63 | 3 | 38 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 6 | 75 | 2 | 25 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 6 | 75 | 2 | 25 |
| Group discussion | 0 | 0 | 4 | 50 | 4 | 50 |

Table 47. Final Evaluation on the Usefulness of Materials, Grade Two

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 1 | 17 | 1 | 17 | 4 | 67 |
| Taking the test before making judgments | 1 | 17 | 0 | 0 | 5 | 83 |
| Defining the threshold students | 1 | 17 | 0 | 0 | 5 | 83 |
| Practicing the procedure | 1 | 17 | 0 | 0 | 5 | 83 |
| Group discussions | 1 | 17 | 0 | 0 | 5 | 83 |

Table 48. Final Evaluation on the Influence of Process Components, Grade Two

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 1 | 17 | 5 | 83 |
| Threshold student definitions | 0 | 0 | 1 | 17 | 5 | 83 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 0 | 0 | 6 | 100 |
| My experience with the students | 0 | 0 | 0 | 0 | 6 | 100 |
| Group discussions | 0 | 0 | 0 | 0 | 6 | 100 |
| Judgments and rationales of other panelists | 0 | 0 | 1 | 17 | 5 | 83 |
| My sense of what students need to know to be proficient | 0 | 0 | 1 | 17 | 5 | 83 |

Table 49. Final Evaluation on Timing, Grade Two

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 6 | 100 | 0 | 0 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 6 | 100 | 0 | 0 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 5 | 83 | 1 | 17 |
| Group discussion | 0 | 0 | 5 | 83 | 1 | 17 |

Table 50. Final Evaluation on the Usefulness of Materials, Grade Span Three Through Five

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 2 | 22 | 1 | 11 | 6 | 67 |
| Taking the test before making judgments | 0 | 0 | 1 | 11 | 8 | 89 |
| Defining the threshold students | 0 | 0 | 0 | 0 | 9 | 100 |
| Practicing the procedure | 0 | 0 | 0 | 0 | 9 | 100 |
| Group discussions | 0 | 0 | 0 | 0 | 9 | 100 |

Table 51. Final Evaluation on the Influence of Process Components, Grade Span Three Through Five

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 2 | 22 | 7 | 78 |
| Threshold student definitions | 0 | 0 | 0 | 0 | 9 | 100 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 2 | 22 | 7 | 78 |
| My experience with the students | 0 | 0 | 1 | 11 | 8 | 89 |
| Group discussions | 0 | 0 | 1 | 11 | 8 | 89 |
| Judgments and rationales of other panelists | 0 | 0 | 3 | 33 | 6 | 67 |
| My sense of what students need to know to be proficient | 0 | 0 | 2 | 22 | 7 | 78 |

Table 52. Final Evaluation on Timing, Grade Span Three Through Five

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 9 | 100 | 0 | 0 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 9 | 100 | 0 | 0 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 9 | 100 | 0 | 0 |
| Group discussion | 0 | 0 | 8 | 89 | 1 | 11 |

Table 53. Final Evaluation on the Usefulness of Materials, Grade Span Six Through Eight

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 0 | 0 | 3 | 27 | 8 | 73 |
| Taking the test before making judgments | 0 | 0 | 0 | 0 | 11 | 100 |
| Defining the threshold students | 0 | 0 | 0 | 0 | 11 | 100 |
| Practicing the procedure | 1 | 9 | 0 | 0 | 10 | 91 |
| Group discussions | 0 | 0 | 1 | 9 | 10 | 91 |

Table 54. Final Evaluation on the Influence of Process Components, Grade Span Six Through Eight

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 2 | 18 | 9 | 82 |
| Threshold student definitions | 0 | 0 | 1 | 9 | 10 | 91 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 4 | 36 | 7 | 64 |
| My experience with the students | 0 | 0 | 3 | 27 | 8 | 73 |
| Group discussions | 0 | 0 | 3 | 27 | 8 | 73 |
| Judgments and rationales of other panelists | 1 | 9 | 3 | 27 | 7 | 64 |
| My sense of what students need to know to be proficient | 0 | 0 | 1 | 9 | 10 | 91 |

Table 55. Final Evaluation on Timing, Grade Span Six Through Eight

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 10 | 91 | 1 | 9 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 8 | 73 | 3 | 27 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 8 | 73 | 3 | 27 |
| Group discussion | 0 | 0 | 9 | 82 | 2 | 18 |

Table 56. Final Evaluation on the Usefulness of Materials, Grade Span Nine Through Twelve

| How *useful* was each of the following materials or procedures in completing the standard setting process? | Not at All Useful N | Not at All Useful % | Somewhat Useful N | Somewhat Useful % | Very Useful N | Very Useful % |
| --- | --- | --- | --- | --- | --- | --- |
| Completing the pre-workshop assignment | 0 | 0 | 5 | 45 | 6 | 55 |
| Taking the test before making judgments | 0 | 0 | 1 | 9 | 10 | 91 |
| Defining the threshold students | 1 | 9 | 0 | 0 | 10 | 91 |
| Practicing the procedure | 1 | 9 | 1 | 9 | 9 | 82 |
| Group discussions | 0 | 0 | 1 | 9 | 10 | 91 |

Table 57. Final Evaluation on the Influence of Process Components, Grade Span Nine Through Twelve

| How *influential* was each of the following in making your judgments? | Not at All Influential N | Not at All Influential % | Somewhat Influential N | Somewhat Influential % | Very Influential N | Very Influential % |
| --- | --- | --- | --- | --- | --- | --- |
| Performance level descriptors | 0 | 0 | 3 | 27 | 8 | 73 |
| Threshold student definitions | 0 | 0 | 1 | 9 | 10 | 91 |
| My perception of the difficulty of the items and tasks | 0 | 0 | 4 | 36 | 7 | 64 |
| My experience with the students | 0 | 0 | 1 | 9 | 10 | 91 |
| Group discussions | 0 | 0 | 5 | 45 | 6 | 55 |
| Judgments and rationales of other panelists | 0 | 0 | 6 | 55 | 5 | 45 |
| My sense of what students need to know to be proficient | 0 | 0 | 4 | 36 | 7 | 64 |

Table 58. Final Evaluation on Timing, Grade Span Nine Through Twelve

| How appropriate was the *amount of time* you were given to complete the different components of the process? | Too Little Time N | Too Little Time % | About Right N | About Right % | Too Much Time N | Too Much Time % |
| --- | --- | --- | --- | --- | --- | --- |
| Test familiarization ("Taking the test") | 0 | 0 | 6 | 55 | 5 | 45 |
| Training and practice in the procedure (Angoff) | 0 | 0 | 9 | 82 | 2 | 18 |
| Training and practice in the procedure (Extended Angoff) | 0 | 0 | 9 | 82 | 2 | 18 |
| Group discussion | 0 | 0 | 8 | 73 | 3 | 27 |

## Chapter 6: Post Standard Setting Results

The 2021–22 administration of the Alternate ELPAC included operational field test items, which were analyzed using classical and item response theory item analyses. These analyses were completed after the standard setting workshop. All items in the prebuilt forms for the Initial Alternate ELPAC functioned as expected, and sufficient well-performing items were available to build the Summative Alternate ELPAC for reporting. The impact data results presented in table 29 through table 40 were calculated after the calibration and scaling analyses were conducted and are presented in the standard setting scale score metric. Although these final results were not available during the workshop, the data was considered by the CDE and the TAG prior to the presentation of the State Superintendent recommendations.

### Conclusion

At the request of the CDE, ETS conducted a standard setting workshop for the Alternate ELPAC for kindergarten and grades one and two from February 22 to 25, 2022, and for grade spans three through five, six through eight, and nine through twelve from February 28 to March 3, 2022. The Modified Angoff and Extended Angoff Methods were applied. The process was implemented as planned. Two rounds of judgments with feedback and discussion were completed, and evidence of internal procedural validity was collected via the panelists’ training polls and final evaluations. A vertical articulation meeting was held on March 3, 2022, in which a group of 25 educators representing all six panels reviewed the standard setting process and panel recommendations across all grade levels and grade spans.

The results of the evaluations indicated that the panelists understood the process and the tasks they were asked to complete, found the instructions easy to follow and the training and materials sufficient and clear, and had adequate time to complete the various tasks. In the Vertical Articulation Meeting, there was general agreement that the standard setting process was implemented in the same manner across all six panels and the expectations defining the threshold of Level 3, the threshold student definitions, were appropriate. Panelists in the vertical articulation panel stated that the comparisons of threshold student definitions across the six grade-level and grade-span assessments showed the appropriate expectation of learning progressions.

Immediately following the workshop, preliminary results were provided to the CDE in the form of raw score recommended threshold scores for each performance level for the total score for all six grade levels and grade spans. Data files were provided to the CDE on April 5, 2022. The final standard setting report presented here provides details about panelists, materials, and processes that were not included in the preliminary results table.

## References

Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard setting topics. *Applied Measurement in Education, 17*(1), 59–88.

California Department of Education. (2019). *NGSS for California public schools, K-12.* Sacramento, CA*:* California Department of Education*.* https://www.cde.ca.gov/pd/ca/‌sc/ngssstandards.asp

Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests.* Thousand Oaks, CA: Sage Publications.

Council of Chief State School Officers. (2019). *English language proficiency (ELP) standards for English learners with significant cognitive disabilities.* Washington, DC: Council of Chief State School Officers.

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.). *Educational Measurement* (4th ed., pp. 433–70). Westport, CT: Praeger.

Hambleton, R. K., & Plake, B. S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education, 8*(1), 45–55.

Tannenbaum, R. J. & Katz, I. R. (2013). Standard setting. In K. F.Geisinger (Ed.), *APA handbook of testing and assessment in psychology,* (Vol.3, pp. 455–77). Washington, DC: American Psychological Association.

## Appendix 1: Attachments

### Attachment A: Panelist Pre-Workshop Assignment, Grades Nine Through Twelve Sample

##### Panelist Invitation to Participate

Dear Standard Setting Panelist:

Thank you for agreeing to serve as a member of a standard setting panel for the Alternate English Language Proficiency Assessments for California (ELPAC). You have been selected to participate in this workshop because of your expertise. You will be working with other California special education and English language educators who know the 2012 California English Language Development (ELD) Standards, as measured by the ELD connectors, are familiar with the Alternate ELPAC, or are working with students who take this assessment.

Your grade-level assignment is located on the top of your notetaking form. You have been assigned to the Grades Nine through Twelve panel.

**What to Expect**

During the standard setting workshop, you will be trained on the process of developing threshold scores that define three Alternate ELPAC performance levels. Standard setting facilitators and assessment specialists from ETS will guide you through the process, and the California Department of Education (CDE) will be present to answer any policy questions you may have. ETS will present the results of the process to the CDE when the workshop is complete. After the 2021–22 administration of the Alternate ELPAC, the CDE will present the State Superintendent of Public Instruction’s threshold score recommendations to the California State Board of Education for approval.

An important part of your work as a panelist will be to define the knowledge and skills at the “entry point” of the performance levels. To help you become familiar with the entry points and the expectations for the Alternate ELPAC, we have attached an excerpt from the range performance level descriptors (PLDs). The first column of the range PLD document includes ELD connectors, which are derived from the 2012 California English Language Development Standards.

**Preparing for Workshop**

Please complete steps 1–3 as follows:

**Step 1: Take the Alternate ELPAC training test.** The training test provides examples of the item types found on the operational Alternate ELPAC. If you have not administered the Alternate ELPAC FT, we recommend that you become familiar with the *Directions for Administration* and items types by reviewing the Training Test. During the standard setting, we will go through operational test items, answer keys, and scoring rules.

To access the training test, proceed through the following steps:

1. Go to the **Practice and Training Test Resources** web page.
2. Select the link in the second bullet under the “Practice and Training Tests” section that says, “Students should access the Practice and Training Tests for the online tests.”
3. Log on as a guest by selecting the [**Sign In**] button (at the bottom of the web page).
4. Select your assigned grade level (kindergarten through grade twelve) from the *Grade* drop-down list. You may notice that the training test includes a range, either kindergarten through grade five or grades six through twelve. The items will allow you to see the item types on the assessments for this grade level range.
5. On the *Your Tests* web page that appears, move down the page and then select the [**Start Alternate ELPAC *Grades X* Training Test**] button.
6. Select the [**Select**] button without making any changes.
7. Select [**Begin Test Now**] and take the training test.
8. It will be helpful for you to become familiar with the training tests *Directions for Administration (DFA)*, the *Alternate ELPAC Training Test Kindergarten Through Grade Five DFA* and the *Alternate ELPAC Training Test Grades Six Through Twelve DFA.*
9. It is not necessary prior to the standard setting for you to access the rest of the test materials; we will go through the operational test materials together during the workshop.

**Step 2: Review the excerpt of the range PLDs for your assigned grade level.** Use the attached **notetaking form** to help you structure your thoughts as you review the PLDs provided. Please focus on your assigned grade level when using your notetaking form and have these notes available during the standard setting workshop. We will have the complete set of PLDs as well as other materials for your reference at the workshop.

**Step 3: Zoom Registration**

Please register for the Zoom meeting by noon on Friday, February 18, 2022, so we can verify that you were able to access the necessary materials. Take the following steps to register for the Zoom meeting:

1. Log on to the Content Review Tool (CRT). You previously received two emails from noreply@ets.org, providing access to CRT.
2. Select the standard setting session and scroll to the bottom of the page to the “Resources” section.
3. Open the PDF titled “Zoom Registration Links for Alternate ELPAC.” Follow the directions described in the PDF document.

**Completing these preworkshop tasks will help you to feel more prepared at the workshop.** If you have any questions or concerns regarding standard setting, please contact me by email at [PBaron@ets.org](mailto:PBaron@ets.org). Thank you in advance for your involvement in this very important work, and we look forward to working with you soon.

Sincerely,

Patricia Baron, Ed.D.  
Standard Setting Director  
ETS

##### Note-taking Task

The Alternate ELPAC performance level descriptors (PLDs) reflect expected performance for a range of students at each performance level. Figure 1 represents students ordered according to their knowledge of the English Language Development (ELD) Standards as measured through the ELD Connectors.

Three performance levels are indicated. In Levels 2 and 3, the student at the beginning of each level is the **threshold student**. The Level 2 (solid pink) threshold student has slightly more knowledge than the highest-performing student in Level 1 (peach checkered pattern). There is no threshold student for Level 1 because any student whose performance is lower than the Level 2 threshold student is automatically in Level 1.

In this task, you will focus only on the PLDs for grades 9–12. Your task is to think about the Level 3 threshold student. The task on the following pages will allow you to become familiar with the PLDs and with the type of comparisons we will be making at the standard setting workshop.

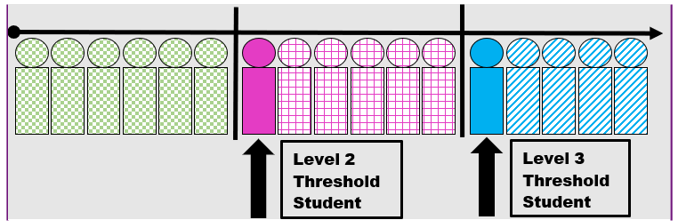


Figure 1. Threshold Student Definitions

##### Task

**Description:** Participants were asked to answer the following questions using brief sentences or bullets. They were asked to have these notes available to guide their thinking during the standard setting workshop. The format of this task has been modified from previously posted standard-setting technical reports to increase accessibility.

**Alternate ELPAC Level 3 threshold student**

For a student who just barely meets the requirements for Level 3, participants were asked to provide responses to the following open-ended questions:

1. What does the student know and what can the student do relative to the performance level descriptors (PLDs)?
2. What might the student not be able to do?
3. How would you distinguish this student from the highest-performing Level 2 student?

### Attachment B: Final Threshold Student Definitions

##### Alternate ELPAC Kindergarten Threshold Student Definitions by Connector

Level 2 Threshold Student can

* Connector 1
  + Occasionally answer simple yes/no questions
  + Attempts to answer simple “wh” questions (who, what)
  + Sometimes identify key pictures or objects
* Connector 2
  + Occasionally identify how someone feels
* Connector 3
  + Sometimes recognize that a story has a beginning
* Connector 4
  + Attempt to identify an event from an experience or a story
* Connector 5
  + Occasionally express a response to simple yes/no questions
  + Sometimes communicate simple feelings
  + Occasionally label a picture
* Connector 6
  + Occasionally express their own feeling or preference
* Connector 7
  + Attempt to identify one element of the story (e.g., character or setting)
  + Sometimes use nouns
  + Occasionally use noun phrases
  + Occasionally use common verbs
* Connector 8
  + Occasionally use the word “and”

Level 3 Threshold Student can

* Connector 1
  + Usually can answer simple yes/no questions
  + Sometimes can answer simple “wh” questions (who, what, where) about key details
  + Usually identify key pictures, objects, or sometimes words (with picture support)
* Connector 2
  + Usually identify the preference of others
* Connector 3
  + Usually recognize that a story starts at the beginning and sometimes has an ending
* Connector 4
  + Usually identify the initial event in a sequence from an experience or story
  + Sometimes identify the connecting word “and”
* Connector 5
  + Express a response to simple yes/no and sometimes “wh” questions (who, what, and occasionally where)
  + Usually communicate simple information or feelings
  + Usually label at least 2 pictures
* Connector 6
  + Usually can express a preference (either one’s own or others’ preference)
* Connector 7
  + Occasionally retell simple information (or 2 elements of a story)
  + Usually use nouns
  + Sometimes use noun phrases
  + Sometimes use common verbs
  + Attempt prepositional phrases
* Connector 8
  + Usually use the word “and” to combine words

##### Alternate ELPAC Grade One Threshold Student Definitions by Connector

Level 2 Threshold Student can

* Connector 1
  + Occasionally answer simple yes/no questions
  + Occasionally identify key pictures, words, or objects
* Connector 2
  + Occasionally identify a preference of others
  + Occasionally identify one reason for the preference of others
* Connector 3
  + Occasionally identify that a simple story has a beginning
  + Occasionally identify the first word in a sentence
* Connector 4
  + Occasionally identify simple information about a familiar story or topic or event
  + Occasionally understand frequently occurring connector words (eg., and, or)
* Connector 5
  + Occasionally express a response to simple yes/no questions
  + Occasionally communicate simple information
  + Occasionally use a few common words or expressions related to the topic
* Connector 6
  + Occasionally express one’s own feeling or preference
* Connector 7
  + Occasionally communicate simple information about a familiar topic
  + Occasionally use a few frequently occurring verbs and nouns
* Connector 8
  + Occasionally use the connector word “and”

Level 3 Threshold Student can

* Connector 1
  + Often answer simple yes/no and simple wh-questions (who, what, where, and sometimes when)
  + Sometimes identify the main topic
* Connector 2
  + Often identify a preference of others
  + Often identify one reason to support the preference of others
* Connector 3
  + Often recognize that simple texts start at the beginning and can sometimes recognize that they have an ending
  + Often identify the first word in a sentence
* Connector 4
  + Often identify simple information about a specific story or topic
  + Often identify two events
  + Often understand a few common connector words (eg., and, or, but)
* Connector 5
  + Often express a response to simple yes/no and wh- questions (who, what, where, and sometimes when)
  + Often communicate simple information
  + Often use a few common words and expressions related to the topic
* Connector 6
  + Often express a preference (own or others)
* Connector 7
  + Often communicate simple information about a familiar event or topic
  + Often retell two events (i.e., steps)
  + Often use common verbs, nouns, and prepositions
* Connector 8
  + Often use a connecting word (e.g., and, or, then) to combine two ideas or simple sentences into one

##### Alternate ELPAC Grade Two Threshold Student Definitions by Connector

Level 2 threshold student can

* Connector 1
  + Sometimes answer at least one yes/no question
  + Sometimes answer simple wh- questions (who, what, where) about key details with prompting
  + Attempts to identify a basic topic
  + Sometimes identify characters in a story
  + Attempt to identify a part of a story
* Connector 2
  + Sometimes identify a preference of others
  + Sometimes identify one reason for the preference of others
* Connector 3
  + Sometimes identify the beginning or end of a story in simple terms
  + Sometimes recognize a first word
  + Attempt to recognize capitalization and end punctuation
* Connector 4
  + Sometimes identify simple information or a detail about a familiar topic
  + Sometimes identify an event in a simple sequence
  + Sometimes understand at least one frequently occurring connector words
* Connector 5
  + Sometimes identify a small number of simple familiar words (three or more) that are used in everyday routines
* Connector 6
  + Sometimes express a response to simple who- or what-questions and attempt to express a response to simple where-questions
  + Sometimes communicate at least one simple detail based on given information
  + Sometimes use simple words or simple expressions somewhat appropriate to the social context
* Connector 7
  + Sometimes express a preference
  + Sometimes provide a reason for a preference
* Connector 8
  + Sometimes communicate simple information presented explicitly in a story about an event
  + Attempt to communicate simple information about a topic in a story
  + Sometimes retell at least one basic detail about an event
  + Sometimes recognize and use common verbs, nouns and attempts to use prepositions
* Connector 9
  + Sometimes use “and” to combine ideas and simple sentences
  + Attempt to use other frequently occurring connecting words to combine ideas and simple sentences (e.g., or, then)

Level 3 threshold student can

* Connector 1
  + Usually answer who- and what-questions about key details and sometimes answer where- and when-questions about key details
  + Sometime identify a part of the main idea
  + Sometime identify a part of a story
* Connector 2
  + Usually identify an opinion or preference of others
  + Sometimes identify one reason for the preference of others
* Connector 3
  + In a story, usually identify beginning or end and sometimes identify the middle
  + Usually identify at least one sentence feature (such as: first word, capitalization, and end punctuation
* Connector 4
  + Usually identify simple information or a detail about a familiar topic
  + Usually identify one event in a sequence, sometimes two or more
  + Sometimes understand at least one temporal word
  + Usually understand frequently occurring connector words
* Connector 5
  + Usually identify some simple familiar words that are used in classroom contexts or everyday routines
* Connector 6
  + Usually express a response to who-, what- and where-questions and sometimes express a response to when-questions
  + Usually communicate information with one or more simple details
  + Usually use simple words and expressions appropriate to the social context and sometimes use simple words and expressions appropriate to an academic context
* Connector 7
  + Usually express own opinion or preference
  + Sometimes support opinion or preference with a reason
  + Usually express the opinion or preference of others and sometimes express the other’s supporting reason
* Connector 8
  + Usually communicate some basic information about an event or topic
  + Sometimes retell a short simple sequence of events, using beginning and end details
  + Usually use common verbs, nouns, and prepositions and sometimes uses verb phrases, noun phrases and prepositional phrases
  + Usually produce simple sentences and attempt to use compound sentences
* Connector 9
  + Usually use frequently occurring connecting words to combine sentences (e.g., “and,” “or,” “then”)
  + Attempt to condense simple sentences or ideas

##### Alternate ELPAC Grades Three Through Five Threshold Student Definitions by Connector

Level 2 threshold student can

* Connector 1
  + Occasionally answer yes/no and simple wh questions about a key detail (who, what, where)
  + Occasionally identify a key detail of the text
  + Occasionally identify the meaning of some familiar words
  + Occasionally retell a part of a story
* Connector 2
  + Occasionally identify a preference of others
  + Occasionally identify one reason for the preference of others
* Connector 3
  + Occasionally identify how a simple text is organized by sequence
  + Occasionally identify one sentence feature: first word, capitalization, or end punctuation
* Connector 4
  + Occasionally identify one piece of information about an event or familiar topic
  + Occasionally identify one part of the sequence of events
  + Occasionally understand some familiar connector words
* Connector 5
  + Occasionally identify a small number of familiar words that are used in classroom and everyday routines
* Connector 6
  + Occasionally express a response to simple yes/no and some wh- questions (who, what, where)
  + Occasionally ask a one-word question
  + Occasionally communicates simple information about a familiar topic by telling or composing one or two words
  + Occasionally use a key word to add detail to a text
  + Occasionally use familiar words appropriate to the social or academic context
* Connector 7
  + Occasionally express their own opinion or preference
  + Occasionally provide a reason for an opinion or preference
* Connector 8
  + Occasionally communicate simple information about an event or familiar topic
  + Occasionally retell a single part of a sequence
  + Occasionally identify and use a small number of familiar verbs and nouns
  + Occasionally identify and attempt to use a familiar preposition
  + Sometimes produce a short phrase
* Connector 9
  + Occasionally use a familiar connecting word to combine simple ideas (ex: and)

Level 3 threshold student can

* Connector 1
  + Often answer a few wh- questions (who, what, where, when)
  + Often identify the main idea and one key detail
  + Often identify the meaning of familiar words and simple phrases (e.g., “four dogs”, “in the morning”)
  + Often retell two parts of a story
* Connector 2
  + Often identify a preference or opinion of others
  + Often identify one reason for the preference or opinion of others
* Connector 3
  + Sometimes identify one common text structure (e.g., sequencing)
  + Often identify how simple texts are organized by sequence (first, next, last)
  + Often identify one sentence feature: first word, capitalization, or end punctuation
* Connector 4
  + Often identify simple information about a familiar topic or event
  + Often identify at least two parts of a sequence
  + Often understand simple temporal, connector, and transition words (e.g., and, too, also, or, because, before, after, next)
* Connector 5
  + Often identify familiar words that are used in classroom contexts and everyday routines
* Connector 6
  + Often express a response to wh- questions (who, what, where, when)
  + Often asks simple questions
  + Often communicate information by telling or composing a simple text
  + Often use key words to add details to a text
  + Often use familiar words appropriate to the social and academic context (e.g., book, pencil, paper, friend, family, play, recess)
* Connector 7
  + Usually express own opinion or preference
  + Often support one’s own opinion with one simple reason
* Connector 8
  + Often communicate information about a familiar topic and attempt to use a detail
  + Often retell two or more events in sequence
  + Sometimes use simple temporal and transition words (e.g., and, first, last, after, next)
  + Often use familiar verbs and nouns
  + Often use familiar prepositions and prepositional phrases (on, in, under, next to)
  + Often produce two simple sentences
* Connector 9
  + Often use familiar connecting words to combine simple ideas

##### Alternate ELPAC Grades Six Through Eight Threshold Student Definitions by Connector

Level 2 threshold student can

* Connector 1
  + Sometimes answer “yes/no”- and “who”-questions; occasionally answer other wh- questions: ‘what,” “where” and “when”
  + Occasionally identify the central idea or theme
  + Sometimes identify a few key words and phrases and their meaning
  + Attempt to identify a few academic and content-specific words
* Connector 2
  + Sometimes identify an opinion or a preference of someone else about a familiar everyday topic
  + Occasionally identify a reason for a preference
* Connector 3
  + Occasionally identify a common text structure.
  + Occasionally identify how a simple text is organized by sequence
  + Occasionally recall an event or part of an event from texts
* Connector 4
  + Occasionally identify simple information about an event or familiar topic
  + Sometimes identify an event or a step in a process
  + Sometimes understand a frequently occurring connector and attempt to use transition words to link ideas
* Connector 5
  + Sometimes express a response to a “yes/no”- “who”- or “what”-question and attempt a response to other simple wh- questions (“where,” “when”)
  + Occasionally ask simple questions
  + Occasionally express an idea with a detail
  + Occasionally use descriptive language appropriate for a task or audience
  + Occasionally use a key word to convey precise meaning
  + Occasionally use some frequently occurring words
* Connector 6
  + Sometimes express a preference or occasionally an opinion (either ones’ own or others’ preference or opinion)
  + Occasionally provide one reason to support one’s own or others’ opinion or preference
* Connector 7
  + Occasionally communicate simple information about an event or a familiar topic
  + Occasionally retell part of a sequence
  + Occasionally use common vocabulary words and simple sentences
  + Sometimes uses common nouns and verbs
  + Occasionally use common noun phrases, verb phrases, and prepositions
* Connector 8
  + Sometimes use common connecting words to link words; Attempt to use common connecting words to link ideas
  + Occasionally produce simple sentences on familiar topics

Level 3 threshold student can

* Connector 1
  + Often answer simple questions (e.g., “yes/no”) and some wh- questions (who, what, where); Sometimes answer “when”-questions
  + Usually determine central ideas or themes and
  + Sometimes identify supporting details
  + Often identify the meaning of general academic and content-specific words, phrases; Occasionally identify the meaning of expressions
* Connector 2
  + Usually identify a preference of someone else; Sometimes identify the opinion of someone else
  + Usually identify a reason for a preference; Sometimes identify a reason to support an opinion
* Connector 3
  + Usually identify a common text structure
  + Often identify how a simple text is organized by a sequence of events
  + Usually recall an event from a text
* Connector 4
  + Usually identify details about an event or familiar topic
  + Usually, can identify the first and last steps in a process
  + Often understand some frequently occurring connector and transition words to link ideas
* Connector 5
  + Often express responses to simple questions and some “who”- “what”- and “where”-questions. Sometimes express responses to “when”-questions.
  + Usually ask simple questions
  + Often express ideas with a detail
  + Often use descriptive language appropriate for the task or audience and attempts for purpose
  + Often use a key word to convey precise meaning
  + Often use some frequently occurring social, general academic, and content-specific words
* Connector 6
  + Often express own opinions and sometimes provide one reason to support the opinions
  + Often express others’ opinions and sometimes provide one reason to support the opinions
* Connector 7
  + Often communicate simple information about an event or a familiar topic
  + Often retell the sequence by expressing the first and last event or first and last step in a process
  + Often use simple temporal words or common transitional words
  + Often use some common connector words to link ideas
  + Usually use nouns and verbs
  + Sometimes use noun phrases, verb phrases, prepositional phrases
* Connector 8
  + Often use common connecting words to link ideas
  + Usually produce simple sentences on topics
  + Occasionally combine ideas using high-frequency conjunctions

##### Alternate ELPAC Grades Nine Through Twelve Threshold Student Definitions by Connector

Level 2 threshold student can

* Connector 1
  + Sometimes answer explicit who, what, where, and when questions
  + Sometimes identify and summarize the central idea or theme and attempts to include a detail
  + Sometimes identify the meaning of a general academic or content-specific or common word or phrase
  + Attempt to identify the meaning of a few general academic or content-specific common expressions
* Connector 2
  + Sometimes recognize an opinion or preference of others
  + Sometimes recognize at least one reason to support the opinion of others
* Connector 3
  + Recognize a common text structure
  + Recognize how simple written texts and oral presentations are organized
* Connector 4
  + Sometimes identify simple information about ideas or events and how they are connected
  + Sometimes identify a part of the sequence of events or steps in a process
  + Attempt to understand a temporal word or a common transitional word or phrase
  + Sometimes understand a familiar connector word and attempts to link ideas, events, and reasons
* Connector 5
  + Sometimes attempt to ask simple questions to express information and ideas (who, what, where, when)
  + Sometimes attempt to answer simple questions to express information and ideas (who, what, where, when)
  + Sometimes attempt to express an idea with a detail
  + Sometimes attempt to use descriptive language appropriate for a task or audience
  + Sometimes attempt to use some social, general academic, or content-specific words or phrases
* Connector 6
  + Occasionally express an opinion or preference (either one’s own or others’ opinion)
  + Sometimes attempt to provide a reason for one’s own opinion or others’ opinion
* Connector 7
  + Often communicate information about a familiar topic and attempt to use a detail
* Connector 8
  + Sometimes attempt to use common connecting words to link ideas
  + Sometimes attempt to produce simple sentences
  + Sometimes attempt to produce one compound sentence using high-frequency conjunctions

Level 3 threshold student can

* Connector 1
  + Usually answer who, what, where and when questions
  + Sometimes answer ‘why’ questions
  + Usually identify and summarize the central idea by including a supporting detail
  + Usually Identify the meaning of a few general academic and content-specific words, phrases, and expressions
* Connector 2
  + Usually recognize an opinion of others
  + Usually recognize a preference of others
  + Usually recognize at least one reason or fact to support the opinion of others
* Connector 3
  + Usually, identify a common text structure
  + Usually, identify how simple written texts and oral presentations are organized
* Connector 4
  + Usually identify a few pieces of information about ideas or events and how they are connected
  + sometimes identify the sequence of events or steps in a process
  + usually understand a temporal word, or common transitional word, or phrase
  + Sometimes understand a few connector words that link ideas, events, and reasons
* Connector 5
  + Usually ask multiple simple questions to express information and ideas of who, what, where, when
  + Attempt to ask why, how questions to express information and ideas
  + Usually answer multiple simple questions to express information and ideas of who, what, where, when
  + Sometimes attempt to answer why, how questions to express information and ideas
  + Usually express ideas with a detail
  + Usually use descriptive language appropriate for a task or audience
  + Sometimes attempts to use descriptive language appropriate for purpose
  + Usually use some social, general academic and content-specific words or phrases and attempt to create meaningful text
* Connector 6
  + sometimes express own opinion and provide a reason or fact to support their opinion
  + sometimes express others’ opinion and provide one reason or fact to support the opinion
  + attempt to negotiate using simple phrases (e.g., “but I…”)
* Connector 7
  + Usually communicate information about an event or topic sometimes with a detail
  + Frequently retell the sequence of events or steps in a process
  + Usually use some temporal words or common transitional words, and attempts to use a phrase
  + Usually use some common connector words to link ideas, events, or reasons
  + Usually use verbs, verb phrases, nouns, noun phrases, adjectives, and prepositional phrases
  + Attempt to use verb tenses and adverbs
* Connector 8
  + Usually use common connecting words to link ideas
  + Usually produce simple sentences and attempt to connect ideas
  + Usually produce one compound sentence using high-frequency conjunctions.
  + Attempt to use an increasing range of high-frequency conjunctions to combine ideas

### Attachment C: Final Evaluation Form

The purpose of the final evaluation form was to obtain the participants’ feedback about the standard setting process overall. They were informed that their feedback would provide a basis for evaluating the training, methods, and materials in the standard setting process and that their responses would be anonymous; no individuals would be identified. The format of this evaluation has been modified from previously posted standard-setting technical reports to increase accessibility.

1. The following demographic information was collected:
2. Gender
3. Race
4. Grade(s) currently teaching
5. Current teaching experience with students, including:
   1. Special Education: Moderate
   2. Special Education: Severe
   3. General Education
   4. English Learners
   5. Other
6. Years of experience in special education
7. Years of experience working with the 2012 CA ELD standards
8. Years of experience working with the ELD Connectors
9. Each participant was asked how useful each of the following materials or procedures were in completing the standard setting process, by choosing either (1) Not at all useful, (2) Somewhat useful, or (3) Very Useful.
   1. Completing the prework assignment
   2. Taking the test before making judgments
   3. Defining the threshold students
   4. Practicing the procedure
   5. Group discussions
10. Each participant was asked how influential each of the following was in making their judgements, by choosing either (1) Not at all influential, (2) Somewhat influential, or (3) Very influential.
    1. Performance level descriptors
    2. Threshold student definitions
    3. My perception of the difficulty of the items and tasks
    4. My experiences with the students
    5. Group discussions
    6. Judgments and rationales of other panelists
    7. My sense of what students need to know to be proficient
11. Each participant was asked how appropriate the amount of time they were given to complete the different components of the process, by choosing either (1) Not at all useful, (2) Somewhat useful, or (3) Very useful, for the following statements:
    1. Test familiarization (“Taking the test”)
    2. Training and practice in the procedure (Angoff)
    3. Training and practice in the procedure (Extended Angoff)
    4. Group discussion

### Attachment D: Nondisclosure Agreement Form

**Description:** Participants were instructed to read and sign the following nondisclosure agreement form:

Test security for California Assessment of Student Performance and Progress (CAASPP) and the English Language Proficiency Assessments for California (ELPAC) is of the utmost importance, and it is the California Department of Education’s obligation to ensure the security of all test materials. The nature and content of any test, test item, proposed or draft test item, or other secure assessment material, including but not limited to the specific language or the subject of test items or any art such as drawings, graphs, tables and sketches, must not be divulged.

By signing below, you acknowledge and agree that the CAASPP and ELPAC test materials are highly secure and that the unauthorized disclosure of any test materials associated with CAASPP and ELPAC could result in substantial monetary and nonmonetary costs to the State to replace the test and materials. You agree that your access to CAASPP and ELPAC test items, proposed or draft test items, or any other test materials is only for the purpose of review as charged by your role as a member of this panel. You agree not to reproduce the tests or any questions within them, directly or indirectly, and not to reveal the nature or content of the test or test items to any other person other than those participating in this meeting.

You further acknowledge and agree that these materials are being provided only for use at this meeting, and you agree to return these materials to the California Department of Education staff member or Contractor staff member at the conclusion of each day of the meeting.

**Participant (Example)**

**By:** Signature of Participant

**Print Name:** John Doe

**Affiliation/Organization:** Story Unified School District

**Date:** 7/1/2022