California Department of Education

Executive Office

SBE-002 (REV. 11/2017)

memo-pptb-adad-feb18item01

# MEMORANDUM

**DATE:** February 1, 2018

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

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

**SUBJECT:** Update on the Initial Assessment Standard Setting Process and Preliminary Review of the Revised Test Blueprints for the Initial English Language Proficiency Assessments for California.

## Summary of Key Issues

### Initial Standard Setting for the English Language Proficiency Assessments for California (ELPAC)

In October 2017, the California Department of Education (CDE) presented an Information Memorandum to the State Board of Education (SBE) detailing the standard setting process for setting threshold scores for the summative ELPAC. In November 2017, the SBE adopted preliminary threshold scores for the summative ELPAC which resulted from the October standard setting conducted by Educational Testing Service (ETS).

In February 2018, ETS will convene standard setting workshops for the ELPAC initial assessment. The workshops will once again be comprised of California educators representing all regions of the state who have extensive experience in working with students learning English. The standard setting panel will recommend threshold scores resulting from the standard setting methods that are similar to those described for the SBE in the October 2017 Information Memorandum. The methods for the February workshops are briefly described below. For detailed information on the standard setting plan, see Attachment 1.

* In the Bookmark Method, an item mapping procedure is used in which participants express their professional judgments by placing markers (or bookmarks) in an ordered item booklet, consisting of a set of ELPAC items ordered by difficulty (i.e., items ordered from easiest to hardest based on data from the fall 2017 field test administration). Panelists will make judgments on the Oral Language composite (Listening and Speaking domains) and Written Language (Reading and Writing domains) composite using this method.
* In the Integrated Judgments Method, which allows participants to consider both the performance on each domain and the overall performance across domains, the overall score is calculated utilizing the score reporting hierarchy approved by the SBE in September 2017.
* The last step in the standard setting will be a cross-grade articulation meeting in which a subset of participants from the six panel rooms, two representatives from each panel, will be recruited to attend. The participants will consider the overall score recommendations and discuss the transitions across adjacent grades, as well as considering continuity across all six sets of threshold score recommendations. The panel will be asked if any of the performance levels might need to be changed to provide better continuity across grades. Panelists will refer to the performance level descriptors, will be provided the recommended bookmark placements indicated, and will review the impact data for all six sets of threshold scores.

Following the panel meeting, ETS will report the panel-recommended threshold scores to the CDE. A review of the standard setting panel’s recommendations will be conducted by psychometricians from the CDE and select ELPAC Technical Advisory Group members, which will inform the creation of the recommendation from the State Superintendent of Public Instruction (SSPI). The SSPI’s recommended preliminary threshold scores, as well as the recommendation for weights of the oral language and written language composite scores used to calculate the overall scale score, will be presented to the SBE for adoption in May 2018.

#### Proposed Revised Test Blueprints for the Initial ELPAC

In September 2017, the SBE approved revised ELPAC test blueprints for the summative assessment. The revised draft test blueprints for the ELPAC initial assessment are included herein for the SBE to perform a preliminary review (see Attachment 2). The proposed test blueprints will be presented to the SBE again in March 2018 for recommended approval. A guide to the definitions of the task types in the test blueprints may be found in Attachment 3. As with the summative test blueprints, these proposed revisions are based on results of the December 2015 ELPAC pilot, the fall 2017 field test, psychometric analyses, and stakeholder and educator input, which is continuing until the test blueprints are presented to the SBE in March.

## Attachment(s)

* Attachment 1: English Language Proficiency Assessments for California (ELPAC) Initial Assessment Standard Setting Plan (19 Pages)
* Attachment 2: Proposed Test Blueprints for the Initial English Language Proficiency Assessments for California (15 Pages)
* Attachment 3: Definitions of Initial Assessment Task Types for the English Language Proficiency Assessments for California (11 Pages)

****

# ATTACHMENT 1:

**English Language Proficiency Assessments for California (ELPAC) Initial Assessment Standard Setting Plan**

**Version 4**

**January 19, 2018**

Prepared by:



Educational Testing Service

660 Rosedale Road

Princeton, NJ 08541

Contract #CN140284

## Table of Contents

Background 3

Purpose and General Description of the Standard Setting Process 5

Panelists 6

Standard Setting Materials 7

Standard Setting Process 8

Test Familiarization 9

Defining the Borderline Student 10

Standard Setting Methodology 11

Bookmark Standard Setting 12

Feedback and Discussion: Round 2 for Each Composite 13

Round 3 Holistic Judgments: Standard Setting for the Overall Score 13

Round 4: Cross-Grade Articulation for the Overall Score 14

Recommendations and Technical Report 14

Staffing, Logistics, and Security of Panel Meetings 15

Appendix A. Sample Rating Forms 16

Appendix B. Sample Agenda 16

Day 1 16

Day 2 18

Day 3 18

Day 4 19

References 19

## List of Tables and Figure

Table 1. ELPAC Method of Administration by Domain and Grade or Grade Span 3

Table 2. Panel Configuration 6

## Background

The English Language Proficiency Assessments for California (ELPAC), aligned with the 2012 California English Language Development (ELD) Standards (California Department of Education [CDE], 2014), is comprised of two separate English Language Proficiency (ELP) assessments: one initial assessment to identify students as English learners, and a second annual summative assessment to both measure a student’s progress in learning English and identify the student’s level of ELP.

The plan presented in this document is for the ELPAC Initial Assessment (IA) standard setting scheduled for February 2018. Much of the process planned for the IA is similar to what was implemented for the Summative Assessment (SA) standard setting that occurred in October 2017 and will therefore be abbreviated where possible. Key differences will be articulated.

Field testing for the ELPAC IA began in fall 2017, and the first operational administration is scheduled to occur in late summer and fall 2018. Standard setting is required so that threshold scores and performance levels will be available at the time of the operational administration. The assessments, given in paper and pencil, will be administered at six grades or grade spans (kindergarten [K], one, two, three through five, six through eight, and nine through twelve) and will assess four domains (Listening, Speaking, Reading, and Writing).

Table 1 below outlines the method of administration for the ELPAC assessment by domain and grade or grade span. The Listening domain is read aloud by the Test Examiner to students in kindergarten and grades one and two, and is administered through streamed recorded audio for grades three through twelve. The Speaking domain is administered by a Test Examiner in a one-on-one setting, and all responses are scored at the time of administration, using task-specific rubrics. The Listening and Reading domains consist entirely of multiple-choice (MC) items, while the Writing and Speaking domains contain only constructed-response (CR) items and no MC items.

**Table 1. ELPAC Method of Administration by Domain, Grade, Grade Span**

| **Domain** | **K** | **1** | **2** | **3–5** | **6–8** | **9–12** |
| --- | --- | --- | --- | --- | --- | --- |
| **Listening** | Read- Aloud MC | Read- Aloud MC | Read- Aloud MC | Recorded Audio MC | Recorded Audio MC | Recorded Audio MC |
| **Speaking** | One-on-one CR | One-on-one CR | One-on-one CR | One-on-one CR | One-on-one CR | One-on-one CR |
| **Reading** | MC | MC | MC | MC | MC | MC |
| **Writing** | CR | CR | CR | CR | CR | CR |

The ELPAC IA will report three performance levels—Levels 1 through 3. Prior to the standard setting, the ELPAC IA general performance level descriptors (PLDs) will be presented for approval at the January 2018 State Board of Education (SBE) meeting, and the grade- and grade span- domain specific PLDs will be finalized by California educators during PLD workshops in late January 2018. The PLDs describe the expectations at each level. Standard setting panelists will utilize the ELPAC PLDs and the 2012 *California English Language Development Standards: Kindergarten Through Grade Twelve* (2012 ELD Standards). Standard setting will be conducted for each grade or grade span; threshold scores will be developed to allow performance levels to be reported for the Overall Score and for Written Language and Oral Language. Figure 1 provides the SBE approved score reporting hierarchy, which applies to the ELPAC IA, for kindergarten through grade twelve. (CDE, 2012)



**Figure 1. ELPAC Initial Assessment Reporting Hierarchy, Kindergarten through Grade Twelve**

The process to develop recommendations for threshold scores for the Oral Language and Written Language composites and the Overall Score will include discussions of all four domains. For each domain and grade or grade span, the standard setting panel will recommend threshold scores that indicate the score that must be earned for a student to reach the beginning (i.e., threshold) of two performance levels—Level 2 and Level 3. One key aspect of the IA related to the standard setting is that the test is designed to facilitate identification of students as English learners, and therefore, the threshold score for Level 3 is a main focus of the process.

## Purpose and General Description of the Standard Setting Process

The purpose of standard setting for the IA is to collect recommendations for the ELPAC IA threshold scores. These recommendations will be reviewed by the CDE, along with additional data, and final determination will be made by the SBE in May 2018. The purpose, general process, logistics, security, and staffing for the ELPAC IA standard setting follows the plan for the ELPAC SA [see C-18 ELPAC SA Standard Setting Plan]. The approach used in this study adheres to the guidelines and best practices recommended in the standard setting literature.

The overall approach for setting standards for the ELPAC is aligned with the 2012 ELD Standards, which reflect the interdependence of the four language domains. By design, the ELPAC and standard setting methodology explicitly support a treatment of skills in combination, such as speaking and listening, rather than as isolated skills. In addition, based on the results of the ELPAC dimensionality study and subsequent approval of the score reporting hierarchy, educators working in standard setting panels will consider the skills that are expected in Listening, Speaking, Reading, and Writing, in order to make threshold score recommendations for the Overall Score scale and the Oral Language and Written Language score scales, by considering the interdependence of these skills.

Specifically, the Bookmark standard setting method (Lewis, et al., 1996; Mitzel, et al., 2001) will be applied to the two composites for Oral Language skills and Written Language skills. The calibration of reading and writing items will provide the necessary data for the Written Language composite and the calibration of the speaking and listening items will provide the necessary data for the Oral Language composite. Calibrations will be performed using the one-parameter item response theory (IRT) model (Hambleton, Swaminathan & Rogers, 1991).

Panelists will make two rounds of judgments on the Oral and Written Language composites, and consider, in the third round, threshold scores for the Overall Score. Panelists will be asked to think holistically about the overall threshold score recommendations and will consider impact data in the third round. Impact data provides panelists with an estimate of the percentage of students who would be classified into each of three performance levels. A subset of panelists will assemble for a fourth round; representatives from each panel will be asked to join a meeting to consider the cross-grade articulation (K through high school) of the threshold scores, taking into account the impact data and test material used in the standard setting, and will make recommendations to accept or modify recommendations for all grades and grade spans.

The IA standard setting workshop will be held over a two-week period in February 2018: February 6–9 and February 12–15—at the Sacramento County Office of Education (SCOE) in Mather, California. A walk-through of the process will be conducted for the CDE prior to the workshop by Dr. Patricia Baron, the standard setting director at Educational Testing Service (ETS).

### Panelists

As was done for the SA standard setting, a diverse sample, representative of educators of English learners in California, will be recruited to participate as panelists in the standard setting sessions. In recruiting panelists, the goal is to include a representative group of California educators who are familiar with the 2012 ELD Standards and who have experience in the education of students who will take the ELPAC. It is also of interest to include subject-area teachers working with these students in grades six and above; these teachers will provide a perspective on content-specific learning goals for the students taking the ELPAC, which may be an important consideration in the identification of students who are English learners compared to students who will be identified as initial fluent English proficient (IFEP).

For the ELPAC IA, there will be six panels of educators: three panels—K and grades one and two will meet in the first week of the workshop, and three panels—grade spans three through five, six through eight, and nine through twelve will meet in the second week (**Table 2**). The targeted number of panelists from this population of educators is 12 per panel, or a total of 72 educators.

**Table 2. Panel Configuration**

| **Panel** | **Grade or Grade Span** | **Meeting Dates** |
| --- | --- | --- |
| A | K | February 6–9, 2017 |
| B | 1 | February 6–9, 2017 |
| C | 2 | February 6–9, 2017 |
| D | 3–5 | February 12–15, 2017 |
| E | 6–8 | February 12–15, 2017 |
| F | 9–12 | February 12–15, 2017 |

As with the SA standard setting, the IA panels will be assembled into grade and grade span-specific panel rooms for much of the standard setting work. Panelists will sit at two tables, with six educators at each table. ETS recommends that the composition of each panel include:

1. Educators who are working with English learners, in the grade level(s) assigned to the panel
2. English-language specialists
3. Educators teaching the subject areas of mathematics, science, and/or social studies

Number 3 above was recommended by the ELPAC Technical Advisory Group, that is, to recruit subject-area teachers who are familiar with English learners, particularly students in the upper grades. The rationale for this goal is that these teachers will have important input as to the English language skills that English learners need in an English-medium classroom.

The final decision on the panelists selected for the workshops will be made by the CDE. After the final list of panelists is approved, panelists will be notified and travel arrangements will be made. Panelists will be required to sign a security agreement notifying them of the confidentiality of the materials used in the standard setting and prohibiting the removal of the materials from the meeting area.

## Standard Setting Materials

All materials and security considerations for the IA workshop will be similar to the SA workshop, with the following exceptions. The pre-workshop assignment provided to the panelists will include the ELPAC Domain- and Grade/Grade Span-Specific **IA** PLDs, for **three** levels. Panelists will be asked to consider the expectations of a student in each of the three performance levels described in the PLDs, and as with the SA assignment, panelists will be instructed to take some notes and bring them to the standard setting workshop.

For each ELPAC IA grade or grade span, the following list of materials will be provided. Specific descriptions are included where materials differ from what was used for the SA standard setting.

* Familiarization materials: ELPAC Examiner’s Manuals, Test Books, Answer Books, Listening audio files, and videos of students responding to ELPAC Speaking items
* Keys and rubrics: Listening and Reading answer keys; Writing and Speaking rubrics for constructed response items (note: rubrics are provided within Examiners’ Manuals)
* Student responses for Speaking and Writing sections
* Ordered item books (OIBs) for Oral Language and Written Language and corresponding item maps
* Ancillary materials for use with OIBs: Listening scripts and Reading passage books
* Judgment forms for the Bookmark method
* Consequence data
* Training evaluation forms
* Workshop agenda

**Familiarization materials:** Panelists will use these materials to become familiar with the test content (i.e., panelists “take the test” without the key and self-score). Operational test forms will be used for all grades and grade-spans. Additionally, to demonstrate the administration of the Speaking section, videos (MP4) produced by SCOE will be used to provide sample student responses for each task; a range of scores for each task will be shown.

**Student responses for Speaking and Writing:** Student responses will be used in the creation of the OIBs. Samples will be selected to represent each score point on the rubric. The source for the samples is the exemplars used in training item raters for scoring speaking, and writing benchmark responses used in scoring writing.

**OIBs:** Both OIBs will contain one item per page. Items will be ordered from least difficult to most difficult based on a response probability of 0.67 employed in the IRT model (Mitzel, Lewis, Patz & Green, 2001). Multiple-choice (MC) items will appear once; one exemplar sample will be used to represent each score on the rubric. Scores of zero will not appear in the OIB. The Written Language OIB will contain reading items and writing responses, calibrated together; reading and writing items will be interspersed, based on their item difficulty values. Similarly, the Oral Language OIB will include listening items and transcriptions of speaking responses representing each score.

**Item maps, consequence data, and evaluation forms:** These materials are the same as were used for the SA standard setting.

## Standard Setting Process

Panelists will attend a general session that will include an overview of ELPAC IA and the bookmark standard setting procedure, as well as the process for setting threshold scores on the Overall score. Panelists will receive training on the method and complete two rounds of judgments for each composite (Oral Language and Written Language). Feedback and discussion will take place after each round of judgment (see the Feedback and Discussion section). After two rounds of judgments are made for each of the composites, panelists will receive training on the holistic judgment process; Round 3 holistic judgments will be made to develop recommendations for the Overall score.

Table leaders will be identified by panel facilitators during the first day of discussion. The responsibility of the table leader is to help keep discussions on track at the table, report interim discussions to the room, and collect materials at the table. Table leaders will be advised of their role during the first day and will join the lead facilitator for table-leader training prior to bookmark judgments.[[1]](#footnote-1)

### Test Familiarization

As with the SA standard setting, immediately following the general training session, panelists will break into their assigned groups associated with the test for which they will be setting standards. For the IA, there are six rather than seven panels; there is one IA assessment for grades nine through twelve, rather than separate nine through ten and eleven through twelve assessments as found on the SA. The test familiarization process is the same as was conducted for the SA standard setting. Panelists will record their responses to the items and check their responses against the answer key, and discuss what they think might be particularly challenging for students and what might be less difficult. The goal of this activity is for panelists to begin to think about and articulate their perception of the general difficulty of the tested content for students.

The roles and responsibilities will be explained to the group, as with the SA standard setting, as follows, and the panel facilitator will respond to any process questions. An ETS content expert will be available to respond to questions about items, and a CDE representative will be available to respond to any policy-related questions, as appropriate. Once the panelists are familiar with the content of the assessment, they will begin the discussion of the pre-workshop assignment, including articulation of the knowledge and skills necessary to reach proficiency Level 2 and Level 3. The focus in each room will be on the assessment level assigned; however, the PLDs for all grades/grade spans will be available to all panelists. These materials are provided to allow the panels to have a clear understanding of the progression of expectations across grades. These materials, as well as impact data, will be provided on day four, during the vertical articulation phase.

### Defining the Borderline Student

Developing definitions of the borderline students is a critical component of any standard setting workshop. For each grade or grade-span, panelists will work in small groups to define the borderline students for the Written Language composite and the Oral Language composite. This process will differ from the SA, in that the panelists will consider reading and writing skills in combination to develop the borderline student definitions for Written Language skills, rather than individual definitions for Reading and Writing. They will follow the same procedure, to develop borderline student definitions for Oral Language domains, considering listening and speaking skills. The process to arrive at borderline student definitions will be the same as was used in the SA standard setting: small group discussions and development of draft borderline student definitions, followed by whole-panel discussion of the small-group definitions in order to reach a panel consensus of what is expected. For the IA, there are two definitions needed, for two thresholds, i.e., Level 2 borderline and Level 3 borderline. Panels will work first on the Level 3 borderline, because this is the point at which a student would be classified as IFEP, differentiating them from students classified as English learners.

For the ELPAC IA, panelists will refer to the specific PLDs that describe the full range for each of three levels. For each of the two composites (Written and Oral Language), borderline student definitions will be written for the student just entering Level 2, the *Borderline Level 2 Student*, and the student just entering Level 3, the *Borderline Level 3 Student.* The definitions are developed by considering the specific PLD description for the appropriate domains, e.g., the PLDs for the Reading and Writing domains are considered when developing the Written Language borderline student definitions. ETS facilitators will instruct panelists to limit the definitions of their borderline students to a sufficient, but not all-encompassing, description.

For the three performance levels on the ELPAC IA, two borderline students will be identified at the entry points into Levels 2 and 3, as indicated in Figure 2.

Level 1

Level 2

Level 3

Borderline Level 2 Student

Borderline Level 3 Student

**Figure 2. Borderline Students for Levels 2 and 3**

## Standard Setting Methodology

Panelists will be trained and have an opportunity to practice prior to the start of actual standard setting, as described below. After training, panelists will be asked to sign a training evaluation form confirming their understanding and readiness to proceed. Panelists will make two rounds of judgments for each of the two composites (Written and Oral Language). The first round (Round 1) of judgments is made independently, without discussion; however, feedback and discussion are important once the Round 1 judgments are collected. Round 2 judgments are also made independently. Panelists’ considerations in the third and final round (Round 3) of judgments are made holistically, and will be described below. In Round 3, panelists will consider both composites to recommend threshold scores based on the Overall Score. After each round, panelists’ judgments are collected, analyzed, and summarized. Feedback and discussion is similar across methods and is described below.

Each test-specific panel is seated in two small groups to facilitate discussion. This table format provides an environment more conducive to panelists sharing their opinions and rationales, as some panelists may be less inclined to speak or have less opportunity to be heard in a large group. The table format also increases the likelihood of alignment among panelists of the performance expectations; each table of experts reviews the results of the table members’ recommendations, and then discuss with the other table. This process invites a discussion of differences in rationales and understanding. Table recommendations are reviewed, discussed and then aggregated across the tables. This also allows analysis of the variability across tables and can be considered a type of replication.

### Bookmark Standard Setting

This method is the same as was used in the SA standard setting for Reading and Listening. For the IA, there are two differences: judgments are made for two composites (Written and Oral Language), and there are only two threshold scores needed for each composite, not three as was needed for the SA.

To make judgments and place bookmarks in the OIB, panelists review each item in the OIB in sequence and consider whether the student at the beginning of Level 2, known as the borderline Level 2 student, would most likely be able to answer the multiple-choice (MC) item correctly or to receive the score based on the rubric for a constructed response (CR) item. A panelist places the Level 2 bookmark on the first item encountered in the OIB that he or she believes the borderline Level 2 student would most likely not be able to address, which indicates that the items beyond that point are too difficult for that borderline student. The panelist continues from that point in the OIB and then stops at the item that the borderline Level 3 student would most likely not be able to address (i.e., the item that likely exceeds the content understanding of the borderline Level 3 student). In the Bookmark method, the definition of “most likely” is related to the IRT model used to order the items. That is, panelists are instructed to think of “most likely” as having a two-thirds likelihood of answering a MC item correctly the score based on the rubric for a CR item. As mentioned earlier, the item ordering in the OIB for ELPAC standard setting is based on a response probability of 0.67 (RP67) as recommended by most research (e.g., see Cizek, 2012, page 135). Using RP67 for item ordering and instructing panelists to think about a two-thirds likelihood, which is easily understood, provides an alignment between the instructions and the analytical model.

Panelists record the bookmark page, or OIB number, for each threshold score. Judgments are summarized and discussed prior to the next round of judgments (see below). As in the SA standard setting, panelists will be given an opportunity to practice making bookmark judgments prior to the start of the actual standard setting. After panelists have received training and responded on the training evaluation form that they are ready to proceed, they will be asked to place their first judgment independently.

The instructions to the panelists for the operational judgments are as follows:

1. Focus on Level 2 first.
2. Review the borderline student definition and refer to the PLDs as needed.
3. Review the first item and identify the knowledge and competencies required to respond successfully to the item. Continue to the next item.
4. Repeat steps 2–3 for Level 3, starting with the next item.

### Feedback and Discussion: Round 2 for Each Composite

The process described here for Round 2 is the same as was described in the SA standard setting, however the work for IA is not at the domain level; the OIBs for the IA are for the two composites, as has been mentioned previously.

The purpose of feedback and discussion is to allow panelists to hear rationales of the other panelists, to receive empirical information about item performance and student performance, and to arrive at a mutual understanding of the expectations of the borderline students on this test. The process of judgment, feedback, and discussion is repeated over the four-day period until all threshold scores are set.

Feedback will be given to the panelists after Round 1 judgments are collected and summarized. The table-level feedback provides an opportunity for the panelists to discuss in a small group setting the range of judgments and rationales for why they made the judgments they did. The panelists see median and range of the panel judgments for the table, and they discuss in table-level groups. They next review the same data for the entire panel, and the facilitator invites a room-level discussion. Results will be projected in each panel room, including summary statistics of the panel’s threshold scores: the panel average (median), minimum, maximum, and the range of judgments. Each table leader provides a summary of the comments and questions from the table-level discussion. The last feedback provided is empirical data showing the impact or consequences of the Round 1 judgments on the distribution of students. This empirical feedback shows “what percentage of students will fall into each category based on these decisions.”

After this discussion, panelists are asked to make an independent Round 2 judgment on the composite for all levels. Feedback from the Round 2 composite judgments is provided at the start of the Round 3 Overall Score process (see below).

### **Round 3 Holistic Judgments: Standard Setting for the Overall Score**

The process described here for Round 3 is similar to the process implemented for the SA standard setting, however the work for IA may be considered more straightforward, since the panelists have used only the Bookmark method for judgments at this point in the workshop. Also the calculation of the Overall Score by the ETS psychometrics staff is more straightforward because they can use the scoring tables directly and provide feedback based on the dummy scale (see below).

After Round 2 judgments have been completed for both composites, feedback will be provided to the panel for each composite and for the Overall Score. The Overall Score after Round 2 will be based on the judgments so far (Round 2), and will be provided using a dummy scale, or “standard setting scale” created specifically for standard setting so that panelists will see scores in a more familiar metric. Panelists will be given an opportunity to consider whether they would revise any of the threshold scores on any of the composite score threshold scores by considering the resulting overall threshold scores, comparing their own individual judgments to the panel’s median judgment. Panelists will be encouraged to ask questions about the data and to discuss rationales for this judgment.

Panelists will also consider the resulting impact data for the Overall Score—the percentage of students that will be placed into each category based on the Overall Score recommendations. For K, the Overall Score will be comprised of a weighted average of the Written Language (30 percent) and Oral Language (70 percent) composites. For all other ELPAC IA grades and grade spans, Written Language and Oral Language composites will be weighted equally to obtain the Overall Score. The facilitator will ask panelists to share their rationales; all comments and questions will be encouraged. Panelists will be reminded to refer to the PLDs, as well as all information received, in their considerations, and make the final Round 3 judgments.

### Round 4: Cross-Grade Articulation for the Overall Score

The last step in the standard setting will involve a subset of participants from the six panel rooms—two representatives from each panel will be recruited to attend the cross-grade articulation meeting. Panelists will be selected to participate prior to the standard setting workshop, to ensure availability and encourage participation. The Round 4 meeting will take place at the end of week two, on February 15, 2018.

The goal of Round 4 is to consider the Overall Score recommendations and to reach a consensus decision for all six sets of threshold score recommendations. The median recommendation will be used if consensus cannot be reached. The panel facilitator will ask panelists to share their rationales; all comments and questions will be encouraged. Panelists will be reminded to refer to the PLDs, will be provided the OIBs, which will have the recommended bookmark placements indicated, and will review the impact data for all six sets of threshold scores. ETS facilitators will be present to guide the discussion and collect the recommendations as the discussion takes place. Additional ETS staff will assist in the process and will document the process.

## Recommendations and Technical Report

ETS will deliver the final recommendations resulting from the first week of standard setting (K and grades one and two) to the CDE on Monday, February 12, 2018. Recommended threshold scores and the data files containing score distributions for these grades will be included.

After the second week of workshops, ETS will deliver the recommended threshold scores and the data files containing score distributions for the balance (grade spans three through five, six through eight, and nine through twelve) to the CDE on Friday, February 23, 2018. Any tables in addition to the recommended threshold score tables, typically designed for presentation to the SBE, may be developed; further discussion between ETS and the CDE to define the composition of the tables and a timeline for delivery will be required prior to the standard setting.

ETS will produce and deliver the final technical report for the standard setting by April 15, 2018. The technical report will contain a description of the process used to set standards, a description of the panelists’ qualifications, results presented during the standard setting process, and statistical information related to the threshold score judgments: two standard errors of judgment, and two standard errors of measurement above and below the panel-recommended threshold score.

## Staffing, Logistics, and Security of Panel Meetings

To allow the standard setting meetings to run smoothly, all groups will be led by trained, experienced standard setting facilitators who will conduct the training, facilitate the process, and keep the discussions on track. Dr. Patricia Baron will lead the introductory training session and the table-leader training and will oversee the workshop process. In addition, ETS will provide one assessment development content specialist, a data analyst, and two psychometricians experienced in standard setting, Dr. Kyunghee Suh and lead psychometrician Dr. Joyce Wang, for the duration of the workshop. ETS Program Manager Sara Querubin will also attend the session and be available to the CDE as needed. All logistics and panelists’ travel concerns will be addressed by SCOE. ETS understands that CDE staff will be present during the standard setting sessions to hear discussion, observe the process, and address any policy-level issues, as appropriate.

Groups will be provided with materials on the first day of each week at the time of registration and other materials as needed during the four-day process. At the end of the process each week, ETS staff will collect and destroy all confidential material.

## Appendix A. Sample Rating Forms

**Bookmark Recording Form**

**Panel Member ID\_\_\_\_\_\_\_\_\_\_\_\_ Table \_\_\_\_\_\_\_\_\_\_\_**

**Test—Circle one: Grade or Grade span—Circle one:**

**Written or Oral—K | 1 | 2 | 3–5 | 6–8 | 9–12**

Please record the number of the item (*item* number, *not* page number) on which you placed your bookmark. This should be the first item in the ordered item booklet (OIB) where the borderline student is not likely to be able to answer the item correctly.

| **Performance Level** | **Bookmark Item # Round 1** | **Bookmark Item #****Round 2** |
| --- | --- | --- |
| **Level 2** | \_\_\_\_\_\_ | \_\_\_\_\_\_ |
| **Level 3** | \_\_\_\_\_\_ | \_\_\_\_\_\_ |
| **Panelist Initials** | \_\_\_\_\_\_ | \_\_\_\_\_\_ |

Please initial the bottom of each column to certify that these are your final judgments.

## Appendix B. Sample Agenda

### Day 1 Standard Setting Sample Agenda

| **Time** | **Activity** |
| --- | --- |
| **7:30 a.m.** | RegistrationWelcome (Board Room) |
| **8 a.m.** | Welcome and introductions |
|  | Overview of standard setting in contextReview process, complete training, and practice |
|  | Begin work in breakout rooms by grade/grade span |
| **Noon** | Lunch break |
| **1 p.m.** | Test familiarization—Reading and WritingReview of the content- and grade-specific performance level descriptorsBegin development of borderline student definitions for Written Language Skills |
| **5 p.m.** | End of Day 1 |

### Day 2 Standard Setting Sample Agenda

| **Time** | **Activity** |
| --- | --- |
| **7:30 a.m.** | Sign in and receive materialsWelcome (Board Room) |
| **8 a.m.** | Assemble in breakout roomsComplete development of borderline student definitionsTraining and practice for bookmark standard setting processBegin standard setting judgments for Written Language Skills |
| **Noon** | Lunch break |
| **1 p.m.** | Finish Round 2 judgments for Written Language SkillsTest familiarization—Listening and SpeakingReview of the content- and grade-specific performance level descriptors for Oral Language Skills |
| **5 p.m.** | End of Day 2 |

### Day 3 Standard Setting Sample Agenda

| **Time** | **Activity** |
| --- | --- |
| **7:30 a.m.** | Sign in and receive materialsWelcome (Board Room) |
| **8 a.m.** | Development borderline student definitions for Oral Language Skills Standard setting process judgments for Oral Language Skills |
| **Noon** | Lunch break |
| **1 p.m.** | Standard setting process judgments for Oral Language Skills |
| **5 p.m.** | End of Day 3 |

### Day 4 Standard Setting Sample Agenda

| **Time** | **Activity** |
| --- | --- |
| **7:30 a.m.** | Sign in and receive materialsWelcome (Board Room) |
| **8 a.m.** | Standard setting process, includes weighting of Written and Oral Judgments for Overall ScoreFinal evaluation for grade and grade span judgment process |
| **Noon** | Lunch break—wrap-up for most participants (Workshop ends with final evaluations on week one; Round 4 takes place after lunch on week two only.) |
| **1:30 p.m.** | Training for Round 4: Cross-grade articulation for ELPAC Overall Score |
| **2 p.m.** | Cross-grade articulation judgments |
| **4 p.m.** | Wrap-up and end of workshop |

***Thank you for your time and contributions!***

## References

California Department of Education. (2012). *California English Language Development Standards (Electronic Edition): Kindergarten Through Grade 12.* Retrieved from <https://www.cde.ca.gov/sp/el/er/documents/eldstndspublication14.pdf>

Cizek, G. J. (Ed.). (2012). *Setting performance standards: foundations, methods, and innovations.* Routledge.

Hambleton, R. K., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.). *Educational Measurement* (4th ed., pp. 433–470). Westport, CT: Praeger.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Newbury Park, CA: Sage.

Karatonis, A., & Sireci, S. G. (2006). The bookmark standard-setting method: a literature review. *Educational Measurement: Issues and Practice*, *25*(1), 4–12.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In Cizek, G. J. (Ed.). *Setting performance standards: concepts, methods, and perspectives* (pp. 249–281). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.



# **ATTACHMENT 2:**

**Proposed Test Blueprints for the Initial English Language Proficiency Assessments for California**

**January 16, 2018**

Prepared by:

****

Educational Testing Service

660 Rosedale Road

Princeton, NJ 08541

Contract #CN140284

**Table of Contents**

Background and Overview 3

Table 1: Proposed Initial Assessment Listening Blueprint: Items and Points by Task Type and Grade 5

Table 2: Proposed Initial Assessment Speaking Blueprint: Items and Points by Task Type and Grade 7

Table 3: Proposed Initial Assessment Reading Blueprint: Items and Points by Task Type and Grade 9

Table 4: Proposed Initial Assessment Writing Blueprint: Items and Points by Task Type and Grade 11

Table 5: Overview of Initial Assessment Items and Points by Domain and Grade 15

## Background and Overview

The English Language Proficiency Assessments for California (ELPAC) is an English language development (ELD) assessment system for students in kindergarten through grade twelve (K–12) that will replace the California English Language Development Test (CELDT). The ELPAC must comply with California *Education Code* (*EC*) sections 60810 et seq. by which the Legislature required the State Superintendent of Public Instruction and the State Board of Education (SBE) to select or develop a test that assesses the ELD of students whose primary language is a language other than English. Beginning with the 2000–01 school year, the new law required the assessment of ELD to be done upon initial enrollment and annually thereafter until the local educational agency (LEA) reclassified the student. State law required the state test of ELD to be aligned with the state adopted ELD Standards (California *EC* Section 60810[c][7]). *EC* Section 60811 (as amended by Assembly Bill [AB] 899 in 2013) requires the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards), to be linked with academic content standards for mathematics and science in order to meet state law and federal accountability requirements.

The ELPAC assessment system consists of two separate assessments: the initial assessment for initial identification and the annual summative assessment. The ELPAC initial assessment is a paper-based assessment that is administered to six grades/grade spans: kindergarten (K), one (1), two (2), three through five (3–5), six through eight (6–‍8), and nine through twelve (9–12). The ELPAC is aligned with the 2012 ELD Standards adopted by the SBE in November 2012. Items also correspond to the Common Core State Standards (CCSS) Mathematical Practices and the Science and Engineering Practices in the California Next Generation of Science Standards (CA‍ NGSS). The initial assessment has a single test at grades nine through twelve (9–‍12) because the 2012 ELD Standards are very similar at grades nine and ten (9–10) and eleven and twelve (11–12) and because students take the initial assessment one time only. That is, there is no need to create separate initial assessments for grades nine and ten (9–10) and grades eleven and twelve (11–12) to limit a student’s exposure to the same items.

The purpose of the initial assessment is to collect information that contributes to the decision as to whether a student should be classified as an English learner or as initial fluent English proficient (IFEP). A goal of the initial assessment is to collect enough evidence to make this decision while keeping the test as short as possible to support efficient administration and scoring. For this reason, the initial assessment contains fewer items, and fewer task types, than the summative assessment.

The task types used on the initial assessment are a subset of task types appearing on the summative assessment. The following task types appear in the summative assessment but do not appear in the initial assessment:

* Speaking—Present and Discuss Information (Speaking with Reading)
* Reading—Read a Student Essay
* Writing—Write About Academic Information (Writing with Reading)

In November 2015, the SBE approved the *Proposed Test Blueprints for the ELPAC*, which included some task types adapted from CELDT items determined to be aligned with the 2012 ELD Standards. After the SBE approval of the *Proposed Test Blueprints for the ELPAC*, the first pilot of ELPAC items, the standalone sample field test of the summative assessment, and the standalone field test of the initial assessment were administered. Analysis of the pilot and the standalone sample field test results led to modifications of the ELPAC test blueprints. The names of some of the task types were changed, some of the task types were removed, and one task type was added to the test blueprints. In addition, the ELPAC test blueprints for the initial assessment (which are in this document) were separated from the ELPAC test blueprints for the summative assessment (which the SBE approved in September 2017). The result of this process are the ELPAC test blueprints for the initial assessment, which appear in Tables 1–4 on the following pages. Table 5 provides an overview of items and points on the ELPAC initial assessment by domain and grade.

Because SBE members reviewed a previous version of this document in November 2015, the following information appears in brackets for the convenience of SBE reviewers. The bracketed information will be removed when the test blueprints are posted to the ELPAC Web site for public use. The brackets make note of:

* An added task type
* Numbers for items and points that appeared in the November 2015 test blueprints
* Task types removed from the test blueprints after the first pilot of ELPAC items (These task types were removed because the pilot evaluation indicated that they were not efficient at gathering information about student English language proficiency.)
* Standards removed that correspond to removed task types

## Table 1: Proposed Initial Assessment Listening Blueprint: Items and Points by Task Type and Grade

| **Listening Task Type** | **Aligned Primary ELD Standard(s)[[2]](#footnote-2)1** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Listen to a Short Exchange [New task type][[3]](#footnote-3) | Part (P)I.A.1,PI.B.5,PII.A.2 | Discrete, 1 point | 3[0][[4]](#footnote-4) | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] | 3[0] |
| Listen to a Classroom Conversation  | PI.A.1,PI.A.3,PI.B.5 | Set of 3 items, 3 points per set[Discrete, 1 point] | 0[4] | 0[4] | 0[4] | 0[4] | 0[4] | 0[4] | 3[4] | 3[4] | 3[4] | 3[4] | 3 | 3 |
| Choose a Reply [Removed][[5]](#footnote-5) | [PI.A.1] | [Discrete, 1 point] | 0[4] | 0[4] | 0[4] | 0[4] | 0[4] | 0[4] | 0[3] | 0[3] | 0[3] | 0[3] | 0[3] | 0[3] |
| Listen to a Story | PI.B.5,PII.A.1 | Set of 3 items, 3 points per set | 6[3] | 6[3] | 6[3] | 6[3] | 6[3] | 6[3] | 3 | 3 | 0[3] | 0[3] | 0 | 0 |
| Listen to an Oral Presentation | Grades K–12PI.B.5Grades 6–12PI.B.7,PI.B.8,PII.A.1 | Set of 3–4 items, 3–4 points per set | 3 | 3 | 3 | 3 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 |

## Table 1: Proposed Initial Assessment Listening Blueprint: Items and Points by Task Type and Grade

| **Listening Task Type** | **Aligned Primary ELD Standard(s)[[6]](#footnote-6)** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Listen to a Speaker Support an Opinion [Listen to Speakers Support Opinions] | PI.A.3,PI.B.5,PI.B.7,PI.B.8,PII.A.1 | Set of 4 items, 4 points per set | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4[0] | 4[0] | 4 | 4 |
| - | - | Totals | 12[14] | 12[14] | 12[14] | 12[14] | 12[14] | 12[14] | 13[14] | 13[14] | 14 | 14 | 14 | 14 |

## Table 2: Proposed Initial Assessment Speaking Blueprint: Items and Points by Task Type and Grade

| **Speaking Task Type** | **Aligned Primary ELD Standard(s)[[7]](#footnote-7)6** | **Aligned Secondary ELD Standard(s)[[8]](#footnote-8)7** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Talk About a Scene | Part (P)I.A.1 | PII.B.3, PII.B.4, PII.B.5 | Set of 6 items, 9 points per set[Set of 3 items, 6 points per set] | 6[3][[9]](#footnote-9)8 | 9[6] | 6[3] | 9[6] | 6[3] | 9[6] | 6[3] | 9[6] | 6[3] | 9[6] | 6[3] | 9[6] |
| Speech Functions | PI.A.4 | PII.B.3, PII.B.4, PII.B.5 | Discrete, 2 points | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 4 | 2 | 4 | 2 | 4 |
| Support an Opinion | PI.C.11 | PII.B.3, PII.B.4, PII.B.5, PII.C.6 | Discrete, 2 points | 1[0] | 2[0] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## Table 2: Proposed Initial Assessment Speaking Blueprint: Items and Points by Task Type and Grade

| **Speaking Task Type** | **Aligned Primary ELD Standard(s)[[10]](#footnote-10)9** | **Aligned Secondary ELD Standard(s)[[11]](#footnote-11)10** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Retell a Narrative(Speaking with Listening)[4-Picture Narrative] | PI.C.9 | PI.B.5, PI.C.12, PII.A.1, PII.A.2, PII.B.3, PII.B.4, PII.B.5, PII.C.6 | Discrete, 4 points | 1 | 4 | 1 | 4 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| Summarize an Academic Presentation (Speaking with Listening) | PI.C.9 | PI.B.5,PII.A.2,PII.B.3,PII.B.4,PII.B.5,PII.C.6,PII.C.7 | Discrete, 4 points | 0[1] | 0[4] | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 | 1 | 4 |
| - | - | - | Totals | 8[5] | 15[14] | 8[5] | 17[14] | 8[5] | 17[14] | 9[6] | 17[14] | 9[6] | 17[14] | 9[6] | 17[14] |

## Table 3: Proposed Initial Assessment Reading Blueprint: Items and Points by Task Type and Grade

| **Reading Task Type** | **Aligned Primary ELD Standard(s)[[12]](#footnote-12)11** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Read-Along Word with Scaffolding | Part (P)III[[13]](#footnote-13)12PI.B.6 | Set of 2 items,3 points per set | 4[6][[14]](#footnote-14)13 | 6[3] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read-Along Story with Scaffolding | PIIIPI.B.6  | Set of 4 items, 5 points per set | 4[5] | 5 | 4[5] | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read-Along Sentence[Removed][[15]](#footnote-15)14 | [PI.B.6] | [Discrete, 1 point] | 0[2] | 0[2] | 0[2] | 0[2] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read-Along Information | PI.B.6 | Set of 3 items, 3 points per set | 0 | 0 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read and Choose a Word | PI.B.6 | Discrete, 1 point | 0 | 0 | 2[0] | 2[0] | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read and Choose a Sentence | PI.B.6 | Discrete, 1 point | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |

## Table 3: Proposed Initial Assessment Reading Blueprint: Items and Points by Task Type and Grade

| **Reading Task Type** | **Aligned Primary ELD Standard(s)[[16]](#footnote-16)15** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Read a Short Informational Passage | Grades 2–12PI.B.6Grades 3–12PI.B.7,PI.B.8,PII.A.1,PII.A.2 | Set of 2–3 items, 1 point per item | 0 | 0 | 0 | 0 | 3 | 3 | 2–3 | 2–3 | 2–3 | 2–3 | 2–3 | 2–3 |
| Read a Literary Passage | PI.B.6,PI.B.7,PI.B.8, PII.A.1, PII.A.2 | Set of 3 items, 1 point per item | 0 | 0 | 0 | 0 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| Read an Informational Passage | PI.B.6,PI.B.7,PI.B.8, PII.A.1, PII.A.2 | Grades 3–12: Set of 5–6 items,1 point per item | 0 | 0 | 0 | 0 | 0 | 0 | 5–6 | 5–6 | 5–6 | 5–6 | 5–6 | 5–6 |
| - | - | Totals | 8[13] | 11[10] | 9[10] | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |

## Table 4: Proposed Initial Assessment Blueprint Writing: Items and Points by Task Type and Grade

| **Writing Task Type** | **Aligned Primary ELD Standard(s)[[17]](#footnote-17)16** | **Aligned Secondary ELD Standard(s)[[18]](#footnote-18)17** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Label a Picture—Word, with Scaffolding | Part (P)I.C.10 | – | Set of 4 items, 6 points per set | 4 | 6 | 4 | 6 | 0[4][[19]](#footnote-19)18 | 0[6] | 0 | 0 | 0 | 0 | 0 | 0 |
| Write a Story Together with Scaffolding | Grades K–2PI.A.2Grades 1–2PI.C.10 | – | Grade K:Set of 4 items, 6 points per set[7 points per set]; Grades 1, 2: Set of 4 items, 7 points per set | 4 | 6[7] | 4 | 7 | 4 | 7 | 0 | 0 | 0 | 0 | 0 | 0 |
| Write an Informational Text Together[Removed][[20]](#footnote-20)19 | [PI.A.2PI.C.10] | – | [Set of 2 items, 5 points per set] | 0 | 0 | 0 | 0 | 0 | 0 | 0[2] | 0[5] | 0 | 0 | 0 | 0 |

## Table 4: Proposed Initial Assessment Blueprint Writing: Items and Points by Task Type and Grade

| **Writing Task Type** | **Aligned Primary ELD Standard(s)[[21]](#footnote-21)20** | **Aligned Secondary ELD Standard(s)[[22]](#footnote-22)21** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Write and Support an Opinion[Removed][[23]](#footnote-23)22 | [PI.C.11] | [PII.B.3, PII.B.4, PII.B.5, PII.C.6] | [Discrete,Grades K–1:2 points;Grades 2–5:3 points] | 0 | 0 | 0[1] | 0[2] | 0[1] | 0[3] | 0[1] | 0[3] | 0 | 0 | 0 | 0 |
| Describe a Picture(Writing with Reading)[Label a Picture—Sentence] | Grade 2PI.C.10Grades 3–5PI.A.2,PII.C.6 | PII.B.3,PII.B.4, PII.B.5,PII.C.7 | Grade 2: Discrete, 3 points;Grades 3–5:2 sets of 2 items, 4 points per set[Discrete, 3 points] | 0 | 0 | 0 | 0 | 2 | 6 | 4[2] | 8[6] | 0 | 0 | 0 | 0 |

## Table 4: Proposed Initial Assessment Blueprint Writing: Items and Points by Task Type and Grade

| **Writing Task Type** | **Aligned Primary ELD Standard(s)[[24]](#footnote-24)23** | **Aligned Secondary ELD Standard(s)[[25]](#footnote-25)24** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Read and Respond to a Message[Removed][[26]](#footnote-26)25 | [PI.C.10] | [PI.B.6, PII.B.3, PII.B.4, PII.B.5, PII.C.6] | [Discrete,3 points] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0[1] | 0[3] | 0 | 0 |
| Write About an Experience | PI.C.10 | PII.B.3, PII.B.4, PII.B.5, PII.C.6 | Discrete, 4 points | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 4 | 1[0] | 4[0] |
| Justify an Opinion | PI.C.11 | PI.C.12,PII.A.1,PII.B.3,PII.B.4,PII.B.5,PII.C.6 | Discrete,4 points | 0 | 0 | 0 | 0 | 0 | 0 | 1[0] | 4[0] | 1[0] | 4[0] | 1 | 4 |

## Table 4: Proposed Initial Assessment Blueprint Writing: Items and Points by Task Type and Grade

| **Writing Task Type** | **Aligned Primary ELD Standard(s)[[27]](#footnote-27)26** | **Aligned Secondary ELD Standard(s)[[28]](#footnote-28)27** | **Discrete/Set, Point Value** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Summarize a Presentation(Writing with Listening)[Removed][[29]](#footnote-29)28 | [PI.C.10] | [PI.B.5,PII.B.3, PII.B.4, PII.B.5, PII.C.6] | [Discrete,4 points] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0[1] | 0[4] |
| - | - | - | Totals | 8 | 12[13] | 8[9] | 13[15] | 6[11] | 13[22] | 5 | 12[14] | 2 | 8[7] | 2 | 8 |

## Table 5: Overview of Initial Assessment Items and Points by Domain and Grade

| **Domain** | **K****Items** | **K****Points** | **1****Items** | **1****Points** | **2****Items** | **2****Points** | **3–5****Items** | **3–5****Points** | **6–8****Items** | **6–8****Points** | **9–12****Items** | **9–12****Points** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Listening | 12[14][[30]](#footnote-30)29 | 12[14] | 12[14] | 12[14] | 12[14] | 12[14] | 13[14] | 13[14] | 14 | 14 | 14 | 14 |
| Speaking | 8[5] | 15[14] | 8[5] | 17[14] | 8[5] | 17[14] | 9[6] | 17[14] | 9[6] | 17[14] | 9[6] | 17[14] |
| Reading | 8[13] | 11[10] | 9[10] | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Writing | 8 | 12[13] | 8[9] | 13[15] | 6[11] | 13[22] | 5 | 12[14] | 2 | 8[7] | 2 | 8 |
| Totals | 36[40] | 50[51] | 37[38] | 52[53] | 36[40] | 52[60] | 37[35] | 52 | 35[32] | 49[45] | 35[32] | 49[46] |

# Attachment 3:

**Definitions of Initial Assessment Task Types for the English Language Proficiency Assessments for California**

**November 6, 2017**

Prepared by:



Educational Testing Service

660 Rosedale Road

Princeton, NJ 08541

Contract #CN140284

This document is intended to provide context for the Proposed Initial Assessment Test Blueprints for the English Language Proficiency Assessments for California (ELPAC). It provides of a definition of each task type in each of the four language domains—listening, speaking, reading, and writing—with the accompanying ELPAC grades and grade spans.

The ELPAC consists of seven grades and grade spans, as referenced below: kindergarten (K); grade one (1); grade two (2); grades three through five (3–5); grades six through eight (6–8); grades nine and ten (9–10); and grades eleven and twelve (11–12).

## Listening

All listening items are multiple choice comprehension questions. At K and grades 1 and 2, the test examiner reads all questions and options aloud. At grades three through twelve (3–12), students listen to an audio recording. The test examiner enters responses for K through grade 1 (K–1) students. Students in grades two through twelve (2–12) mark their own responses in the Answer Book.

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Listen to a Short Exchange**Communicative Context: The test taker shows the ability to listen to a short exchange between two speakers attentively by answering one question.Stimulus: The test taker listens to a short exchange between two speakers in a school context. | All grades/grade spans |
| **Listen to a Classroom Conversation**Communicative Context: The test taker shows the ability to listen to a conversation attentively by answering questions.Stimulus: The test taker listens to a conversation between two students or a student and a teacher. | 3–5, 6–8, 9–10, 11–12 |

## Listening (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Listen to a Story (Similar to California English Language Development Test [CELDT] Listening—Extended Listening Comprehension)**Communicative Context: The test taker demonstrates active listening to a story by answering detailed questions.Stimulus: The test taker listens to a story. The story includes a conversation, which is provided using direct speech and/or indirect speech.  | K, 1, 2, 3–5 |
| **Listen to an Oral Presentation (Similar to CELDT Listening—Extended Listening Comprehension)** Communicative Context: The test taker demonstrates active listening to an oral presentation by answering detailed questions.Stimulus: The test taker listens to a teacher give a presentation. | All grades/grade spans |
| **Listen to a Speaker Support an Opinion**Communicative Context: The test taker answers detailed questions to demonstrate active listening to a speaker who is supporting an opinion.Stimulus: The test taker listens to an extended conversation between two speakers in a school context. In the conversation, one classmate provides support for an opinion. | 6–8, 9–10, 11–12 |

## Speaking

All speaking items are constructed-response items. The test examiner scores each student’s response in real time based on speaking rubrics.

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Talk about a Scene** Communicative Context: The test taker describes a common scene to a teacher.Stimulus: The test taker views a scene from a school or a familiar place that shows a number of people doing common activities.Prompt: The test examiner asks a number of questions about the scene.Response: The test taker responds by answering questions about the scene. | All grades/grade spans |
| **Speech Functions (Same as CELDT Speaking—Speech Functions)**Communicative Context: The test taker uses language to inform, persuade, make a request, etc. in an appropriate manner to a student or a teacher.Stimulus: The test examiner describes a situation.Prompt: The test examiner asks what the test taker would say or ask in the situation.Response: The test taker provides an appropriate response for the situation. | 3–5, 6–8, 9–10, 11–12 |
| **Support an Opinion**Communicative Context: The test taker shares his/her opinion and support for the opinion expressed.Stimulus: A common topic is introduced. The test taker has a choice between two objects, activities, etc. Prompt: The test examiner asks the test taker to provide his/her opinion along with appropriate support. Response: The test taker provides his/her opinion along with support. | K |

## Speaking (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Retell a Narrative (Integrated Skills: Speaking with Listening)**Communicative Context: The test taker retells a story that includes a series of events.Stimulus: The test taker views a series of pictures while listening to the test examiner read a story aloud. Prompt: The test examiner asks the test taker to retell the story using the pictures.Response: The test taker uses the pictures to retell the story. | K, 1, 2 |
| **Summarize an Academic Presentation (Integrated Skills: Speaking with Listening)** Communicative Context: The test taker summarizes a presentation that was given by a teacher. Stimulus: The test taker listens to a presentation while viewing images that go along with the presentation.Prompt: The test taker is prompted to retell the main points of the presentation with the help of the visuals that were provided during the presentation.Response: The test taker summarizes the main points of the presentation. | 1, 2, 3–5, 6–8, 9–10, 11–12 |

## Reading

All reading items are multiple choice comprehension questions; kindergarten also includes foundational literacy items with select task types. The test examiner enters responses for K–1 students. Students in grades 2–12 mark their own responses in the Answer Book.

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Read-Along Word with Scaffolding**Communicative Context: The test taker and a teacher are reading together.Stimulus: The test taker listens to a word and reads along while looking at three picture options in the Answer Book. This is preceded by a foundational literacy skills item, in which the test examiner supports the test taker in decoding the word.Prompt: The test taker is asked to decode a word. The test taker is then asked which picture matches the word.Response: The test taker provides spoken responses to the first question about the names of the letters in a word, the sound of the initial letter, and the test taker’s ability to read the word. For the second question, the test taker points to the picture that represents the word. | K |
| **Read-Along Story with Scaffolding**Communicative Context: The test taker reads a story together with the teacher.Stimulus: The test taker listens to a story and reads along. The test examiner sweeps his or her finger under the text while reading the story aloud. This is preceded by a foundational literacy item in which the test examiner supports the test taker in demonstrating print concepts.Response: The test taker provides spoken responses to the first question about the pre-reading skills of where to begin reading and the direction of reading. For the remaining three comprehension questions, the test taker chooses the correct answer from a set of three written and spoken or picture options. | K, 1 |

## Reading (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Read-Along Information** Communicative Context: The test taker and a teacher read an informational text together.Stimulus: The test taker listens to informational text and reads along. The test examiner sweeps his or her finger under the text while reading the information aloud.Response: The test taker chooses the correct answer from a set of three written and spoken or picture options.  | 1 |
| **Read and Choose a Word** Communicative Context: The test taker is reading grade-level words independently.Stimulus: The test taker looks at a picture.Prompt: The test taker is asked to choose the word that represents the picture.Response: The test taker reads three words and chooses the word that matches the picture.  | 1, 2 |
| **Read and Choose a Sentence** Communicative Context: The test taker is reading independently.Stimulus: The test taker looks at a picture.Prompt: The test taker is asked to choose the sentence that represents the picture.Response: The test taker reads three sentences and chooses the sentence that describes the picture.  | 2, 3–5, 6–8, 9–10, 11–12 |
| **Read a Short Informational Passage** Communicative Context: The test taker reads a short informational passage about a topic from science or the social sciences.Stimulus: The test taker reads an informational passage. Response: The test taker answers questions about the passage. | 2, 3–5, 6–8, 9–10, 11–12 |

## Reading (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Read a Literary Passage (Similar to CELDT Reading—Reading Comprehension)**Communicative Context: The test taker reads a literary passage that would be presented in an English language arts class.Stimulus: The test taker reads a literary passage.Response: The test taker answers a set of multiple choice questions. Questions include comprehension of main idea and details as well as questions concerning language use and word choice. | 2 |
| **Read an Informational Passage** **(Similar to CELDT Reading—Reading Comprehension)** Communicative Context: The test taker reads an informational passage that would be presented in an English language arts or a history, science, or social studies class.Stimulus: The test taker reads an informational passage.Response: The test taker answers a set of multiple choice questions. Questions include comprehension of main idea and details as well as questions concerning language use and word choice. | 3–5, 6–8, 9–10, 11–12 |

## Writing

All writing items are constructed-response items. After the test administration, raters score student responses based on writing rubrics.

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Label a Picture—Word, with Scaffolding**Communicative Context: The test taker is collaborating with a teacher to write about a picture for a classroom display.Stimulus: The test taker looks at a picture. Prompts: The test taker is prompted to write labels for a picture. The test examiner supports the test taker by prompting for letter-level responses before prompting for full words.Responses: The test taker writes letters and words for items in the picture. | K, 1 |
| **Write a Story Together with Scaffolding**Communicative Context: The test taker is collaborating with a teacher to jointly compose a short literary text. Stimulus: The test taker sees a picture and is provided the initial sentence of the story followed by a sentence frame. The test examiner supports the test taker by prompting for letter level output, then word level, and finally one sentence (at grades 1 and 2).Prompts 1–2: The test taker hears the title and writes the missing (initial) letters.Prompt 3: The test taker hears a sentence and writes the missing word. (At kindergarten, there are two word-level prompts.)Prompt 4: The test taker composes and writes an original sentence to complete the story (at grades 1 and 2).Responses: The test taker writes letters, a word, and a sentence in the blank spaces. | K, 1, 2 |

## Writing (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Describe a Picture (Grade 2)**Communicative Context: The test taker looks at a picture and writes a brief description about what is happening.Stimulus: The stimulus consists of an image. The image shows an easily depicted, common action. Context, contents, and expected vocabulary are grade appropriate.Prompt: The test taker is instructed to write a sentence describing the picture. Response: The test taker writes a sentence(s) to describe the picture. | 2 |
| **Describe a Picture (Integrated Skills: Writing with Reading) (Grades 3–5)**Communicative Context: The test taker is working with a classmate to write a paragraph about a picture.Stimulus: The stimulus consists of an image and a short paragraph about the image. The image shows an easily depicted, common action. Context, contents, and expected vocabulary are grade appropriate. The paragraph may have errors.The test taker answers two of the following prompts:Prompt 1: The test taker is asked to rewrite a sentence with more details.Prompt 2: The test taker is asked to correct two errors in a sentence.Prompt 3: The test taker is asked to combine and condense two sentences.Prompt 4: The test taker is asked to write a new sentence to describe what might happen next. Response: The test taker writes a sentence in response to each prompt.  | 3–5 |

## Writing (continued)

| **Task Type** | **Grades/Grade Spans** |
| --- | --- |
| **Write About an Experience** Communicative Context: The test taker is provided with a common topic, such as a favorite celebration or a memorable trip. The test taker is prompted to write about the topic from his or her own personal experience.Stimulus: The test taker is provided with a common topic, such as a favorite celebration or a memorable trip.Prompt: The test taker is prompted to write about the topic.Response: The test taker writes a paragraph about a personal experience. | 6–8, 9–10, 11–12 |
| **Justify an Opinion** Communicative Context: The test taker writes an essay about a school-related issues as if the essay will be given to the school principal.Stimulus: A common topic (e.g., wearing school uniforms, best type of exercise) is introduced.Prompt: The test examiner asks the test taker to provide his/her opinion along with appropriate support.Response: The test taker writes a paragraph containing his/her opinion along with support. | 3–5, 6–8, 9–10, 11–12 |

1. Leader training will take place prior to the start of panel room work on the second day of the workshop, as was done for the SA standard setting. Table leaders will be provided with instructions regarding their roles and responsibilities. [↑](#footnote-ref-1)
2. 1 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. Additional secondary standards may apply to select item types. [↑](#footnote-ref-2)
3. Listen to a Short Exchange is a new task type that was added to the test blueprint after the first pilot of ELPAC items. [↑](#footnote-ref-3)
4. Numbers in square brackets are numbers that appeared in the version of the test blueprints that the SBE approved in November 2015. [↑](#footnote-ref-4)
5. “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-5)
6. Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. Additional secondary standards may apply to select item types. [↑](#footnote-ref-6)
7. 6 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-7)
8. 7 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-8)
9. 8 Numbers in square brackets are numbers that appeared in the version of the test blueprints that the SBE approved in November 2015. [↑](#footnote-ref-9)
10. 9 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-10)
11. 10 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-11)
12. 11 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. Additional secondary standards may apply to select item types. [↑](#footnote-ref-12)
13. 12 PIII foundational literacy skills may not be expressly considered standards. [↑](#footnote-ref-13)
14. 13 Numbers in square brackets are numbers that appeared in the version of the test blueprints that the SBE approved in November 2015. [↑](#footnote-ref-14)
15. 14 “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-15)
16. 15 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. Additional secondary standards may apply to select item types. [↑](#footnote-ref-16)
17. 16 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-17)
18. 17 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-18)
19. 18 Numbers in square brackets are numbers that appeared in the version of the test blueprints that the SBE approved in November 2015. [↑](#footnote-ref-19)
20. 19 “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-20)
21. 20 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-21)
22. 21 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-22)
23. 22 “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-23)
24. 23 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-24)
25. 24 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-25)
26. 25 “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-26)
27. 26 Multiple primary standards are listed for those task types with items that align with different standards. These may vary by grade. [↑](#footnote-ref-27)
28. 27 Secondary standards vary based on the item types. These secondary standards are used in service of the PI Standards and are implicitly accounted for in the constructed-response rubrics. [↑](#footnote-ref-28)
29. 28 “[Removed]” appears under task types that appeared in the November 2015 version of the test blueprints but were removed after the pilot. [↑](#footnote-ref-29)
30. 29 Numbers in square brackets are numbers that appeared in the version of the test blueprints that the SBE approved in November 2015. [↑](#footnote-ref-30)