



**California Department of Education  
Assessment Development &  
Administration Division**



**Summative English Language  
Proficiency Assessments for California  
Technical Report**

**2018–2019 Administration**

**Final Submitted July 20, 2020  
Educational Testing Service**



**Contract No. CN140284**

# Table of Contents

<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1. Background .....	1
1.2. Test Purpose .....	1
1.3. Test Content .....	1
1.4. Testing Window .....	2
1.5. Intended Population .....	2
1.6. Intended Use and Purpose of Test Scores .....	2
1.7. Limitations of the Assessment .....	3
1.8. Groups and Organizations Involved with the ELPAC System .....	3
1.8.1. State Board of Education .....	3
1.8.2. California Department of Education .....	3
1.8.3. California Educators .....	4
1.8.4. Contractors .....	4
1.9. Overview of the Technical Report .....	5
References .....	6
<b>Chapter 2: Overview of ELPAC Development Processes</b> .....	<b>7</b>
2.1. Item Development .....	7
2.1.1. Item Format .....	8
2.1.2. Item Writing Guidelines .....	8
2.1.3. Item Banking .....	8
2.2. Test Assembly .....	9
2.2.1. Test Design .....	9
2.2.2. Test Blueprints .....	9
2.2.3. Test Length .....	9
2.2.4. Psychometric Review .....	10
2.2.5. CDE Review .....	10
2.3. Test Administration .....	11
2.3.1. Test Security and Confidentiality .....	11
2.3.2. Procedures to Maintain Standardization .....	11
2.4. Test-Taking Population .....	12
2.5. Accessibility .....	12
2.5.1. Resources for Selection of Accessibility Resources .....	12
2.5.2. Delivery of Accessibility Resources .....	13
2.5.3. Unlisted Resources .....	13
References .....	14
<b>Chapter 3: Item Development and Review</b> .....	<b>15</b>
3.1. Item Development .....	15
3.1.1. Overview .....	15
3.1.2. Item Development Specifications .....	15
3.1.3. Test Blueprints .....	15
3.1.4. Task Types .....	16
3.1.5. Updates to Item Writing Guidelines .....	19
3.2. Item Review Process .....	19
3.2.1. Overview .....	19

3.2.2. Composition of the ELPAC Item Writer and Item Review Meetings and Participant Qualifications.....	19
3.2.3. Selection of Item Writers .....	20
3.2.4. Item Writing by Educators .....	21
3.2.5. Item Writing by Contractors .....	23
3.2.6. Item Review Panels.....	23
3.3. Item Banking.....	25
References .....	26
<b>Chapter 4: Test Development.....</b>	<b>27</b>
4.1. Test Design .....	27
4.1.1. Revision of the Test Blueprints .....	27
4.2. Item Selection.....	28
4.2.1. Test Development Specifications .....	28
4.2.2. Content Criteria .....	29
4.2.3. Statistical and Psychometric Criteria .....	29
4.2.4. CDE Review of Item Selection .....	30
4.3. Forms Development .....	30
4.3.1. Developing Paper-based Test Materials.....	30
4.3.2. Developing Special Version Forms.....	31
4.3.3. CDE Review of Assembled Forms .....	33
References .....	34
<b>Chapter 5: Test Administration .....</b>	<b>35</b>
5.1. Procedures to Maintain Standardization.....	35
5.1.1. LEA ELPAC Coordinator .....	35
5.1.2. Site ELPAC Coordinator.....	36
5.1.3. ELPAC Test Examiner.....	37
5.1.4. Instructions for Test Administration .....	37
5.2. Training .....	38
5.2.1. General Test Administration .....	38
5.2.2. Scoring Training of Trainers Workshops .....	39
5.2.3. Scoring Rubrics .....	40
5.3. Testing Students with Disabilities .....	42
5.3.1. Locally Determined Alternate Assessments .....	43
5.4. Test Security and Confidentiality .....	43
5.4.1. ETS' Office of Testing Integrity.....	43
5.4.2. Test Delivery .....	44
5.4.3. Security of Electronic Files Using a Firewall.....	44
5.4.4. Transfer of Scores via Secure Data Exchange.....	44
5.4.5. Data Management.....	45
5.4.6. Statistical Analysis on Secure Servers .....	45
5.4.7. Student Confidentiality.....	45
5.4.8. Student Test Results .....	46
5.4.9. Security and Test Administration Incident Reporting Process.....	47
References .....	49
<b>Chapter 6: Standard Setting .....</b>	<b>50</b>
6.1. Background .....	50
6.2. Performance Level Descriptors .....	51

6.3. Standard Setting Methodologies.....	51
6.3.1. Bookmark Method (Reading and Listening Domains) .....	51
6.3.2. Performance Profile Method (Speaking and Writing Domains) .....	52
6.4. Standard Setting Procedures.....	53
6.4.1. Panelists.....	53
6.4.2. Materials.....	53
6.4.3. Process .....	54
6.5. Standard Setting Results.....	55
References.....	58
<b>Chapter 7: Scoring and Reporting .....</b>	<b>59</b>
7.1. Procedures for Maintaining and Retrieving Individual Scores .....	59
7.2. Multiple-Choice Scoring.....	59
7.3. Constructed-Response Scoring for Writing .....	59
7.3.1. Scorer Training.....	59
7.3.2. Scoring Practices.....	61
7.3.3. Managing Scoring.....	61
7.3.4. Monitoring Scoring.....	63
7.4. Constructed-Response Scoring for Speaking.....	63
7.4.1. Scorer Training for Speaking.....	63
7.4.2. Scorer Qualifications for Speaking .....	68
7.5. Types of Scores.....	68
7.5.1. Raw Scores .....	68
7.5.2. Scale Scores .....	68
7.5.3. ELPAC Levels .....	69
7.6. Overview of Score Aggregation.....	71
7.6.1. Individual Student Score Distributions and Summary Statistics .....	71
7.7. Reports Produced and Scores for Each Report .....	76
7.7.1. Online Reporting .....	76
7.7.2. Special Cases .....	76
7.7.3. Types of Score Reports.....	77
7.7.4. Score Report Applications .....	78
7.7.5. Criteria for Interpreting Test Scores.....	78
7.7.6. Criteria for Interpreting Score Reports.....	79
References.....	80
<b>Chapter 8: Test Analyses and Results .....</b>	<b>81</b>
8.1. Background .....	81
8.1.1. Summary of the Analyses.....	81
8.1.2. Samples for the Analyses.....	82
8.2. Classical Item Analysis Statistics.....	82
8.2.1. Description of Classical Item Analysis Statistics.....	83
8.2.2. Summary of Classical Item Analysis Flagging Criteria .....	85
8.2.3. Classical Item Analysis Results Summary.....	86
8.2.4. Omit Rates .....	87
8.3. Item Response Theory Analyses.....	88
8.4. Differential Item Functioning (DIF).....	93
8.4.1. Multiple-choice Items.....	94
8.4.2. Constructed-Response Items .....	95

8.4.3. Classification .....	96
8.5. Reliability Analyses.....	99
8.5.1. Internal Consistency Reliability.....	99
8.5.2. Standard Error of Measurement (SEM).....	100
8.5.3. Conditional Standard Error of Measurement (CSEM).....	102
8.5.4. Writing Score Reliability.....	104
8.5.5. Decision Classification Analyses .....	105
References .....	107
Accessibility Information .....	108
Alternative Text for Equation 8.1 .....	108
Alternative Text for Equation 8.2 .....	108
Alternative Text for Equation 8.3 .....	108
Alternative Text for Equation 8.4 .....	108
Alternative Text for Equation 8.5 .....	108
Alternative Text for Equation 8.6 .....	108
Alternative Text for Equation 8.7 .....	108
Alternative Text for Equation 8.8 .....	108
Alternative Text for Equation 8.9 .....	109
Alternative Text for Equation 8.10 .....	109
Alternative Text for Equation 8.11 .....	109
Alternative Text for Equation 8.12 .....	109
Alternative Text for Equation 8.13 .....	109
Alternative Text for Equation 8.14 .....	109
Alternative Text for Equation 8.15 .....	109
Alternative Text for Equation 8.16 .....	109
Accessibility Reference: Table Data for Test Characteristic Curves.....	110
<b>Chapter 9: Validity .....</b>	<b>153</b>
9.1. Validity of the ELPAC Test Design .....	153
9.1.1. Purpose of the ELPAC .....	153
9.1.2. The Constructs to Be Measured.....	153
9.1.3. The Interpretations and Uses of the Scores .....	154
9.1.4. The Intended Population .....	154
9.2. Validity of the ELPAC Test Content.....	155
9.2.1. Description of the State Standards.....	155
9.2.2. Item Writing Guidelines .....	155
9.2.3. Test Blueprints .....	155
9.2.4. Form Assembly Process.....	155
9.3. Validity of the ELPAC's Internal Structure .....	155
9.3.1. Fairness and Differential Item Functioning .....	156
9.3.2. Reliability.....	156
9.3.3. Other Validity Evidence .....	156
References .....	159
<b>Chapter 10: Quality Control Procedures .....</b>	<b>160</b>
10.1. Quality Control of Item and Test Development.....	160
10.1.1. Quality Control of Item Writing.....	160
10.1.2. Quality Control of Item Selection .....	160
10.2. Quality Control of Test Materials .....	160

10.2.1. Test Administration Manuals .....	160
10.2.2. Processing Test Materials .....	161
10.3. Quality Control of Test Delivery .....	161
10.4. Quality Control of Test Assignment .....	161
10.5. Quality Control of Test Administration .....	161
10.6. Quality Control of Machine-Scoring Procedures.....	162
10.7. Quality Control of Hand Scoring Procedures for Writing.....	162
10.7.1. Second Scores .....	162
10.7.2. Read-Behinds.....	163
10.7.3. Validity Responses .....	163
10.8. Quality Control of Psychometric Processes.....	163
10.8.1. Development of Scoring Specifications .....	163
10.8.2. Development of Scoring Procedures .....	163
10.9. Quality Control of Reporting .....	164
10.9.1. Student Scores from Summary Reports .....	165
10.9.2. End-to-End Testing for Operational Administration .....	165
Reference.....	166
<b>Chapter 11: Historical Results.....</b>	<b>167</b>
11.1. Base-Year Comparison .....	167
11.1.1. Examinee Characteristics.....	167
11.1.2. Scale Score Distributions .....	169
11.2. Test-Form Characteristics .....	170
<b>Chapter 12: Continuous Program Improvement.....</b>	<b>171</b>
12.1. Test Development .....	171
12.2. Test Delivery.....	171
12.2.1. Move to Online Testing.....	171
12.2.2. Postadministration Survey.....	171
12.2.3. Training and Communication.....	172
12.3. Constructed-Response Scoring.....	172

## List of Appendices

### Chapter 2 Appendices

Appendix 2.A: ELPAC Test Takers

Appendix 2.B: Accessibility

### Chapter 7 Appendices

Appendix 7 A: Raw Score and Level Distributions by Domain

Appendix 7.B: Raw Score, Scale Score, and Level Distributions for Each Composite  
Language Skill

Appendix 7.C: Overall Scale Scores and Performance Distribution

Appendix 7.D. Student Performance for Overall and Composite Language Skill Scores

Appendix 7.E: Student Performance for Domains

Appendix 7.F: Demographic Summaries

### Chapter 8 Appendices

Appendix 8.A. Classical Item Analyses

Appendix 8.B: Item Response Theory Parameter Estimates

- [Appendix 8.C. Reliability Estimates](#)
- [Appendix 8.D. Raw-to-Scale-Score Conversions](#)
- [Appendix 8.E. Interrater Agreement for Writing Items](#)
- [Appendix 8.F. Classification Consistency and Accuracy](#)

**Chapter 9 Appendix**

- [Appendix 9.A: Correlation Coefficients](#)

**Chapter 11 Appendices**

- [Appendix 11.A: ELPAC Summative Test Takers Demographic Summary Across Administrations](#)
- [Appendix 11.B: ELPAC Summative Scale Score Distribution Across Administrations](#)
- [Appendix 11.C: Test-Form Statistical Comparison](#)

**List of Tables**

Acronyms and Initialisms Used in the *Summative English Language Proficiency Assessments for California Technical Report*.....ix

Table 3.1 ELPAC Item Writer Workshop (IWW) and Item Review Meeting (IRM) Qualifications, by Meeting Type and Total..... 20

Table 3.2 Status of Stimuli and Items After the 2016 Item Review Panel Meetings ..... 24

Table 3.3 Status of Stimuli and Items After the 2018 Item Review Panel Meetings ..... 25

Table 5.1 2018 AST Training..... 39

Table 5.2 ELPAC Speaking Range Finding (SpRF) and Writing Range Finding (WrRF) Meetings Qualifications, by Meeting Type and Total ..... 42

Table 6.1 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Kindergarten ..... 56

Table 6.2 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade One ..... 56

Table 6.3 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Two ..... 56

Table 6.4 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Three Through Five ..... 56

Table 6.5 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Six Through Eight..... 57

Table 6.6 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Nine and Ten..... 57

Table 6.7 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Eleven and Twelve ..... 57

Table 7.1 Summary of Characteristics of ELPAC Raters ..... 62

Table 7.2 Rater Qualification Standard for Agreement with Correct Scores ..... 63

Table 7.3 Available Scoring Training Videos ..... 66

Table 7.4 Available Training Presentations for Speaking ..... 66

Table 7.5 Training and Calibration Quizzes by Task Type ..... 67

Table 7.6 Composite Language Skills and Overall Reporting Scale Score Ranges for Each Reporting Level by Grade Level or Grade Span ..... 70

Table 7.7 Mean and Standard Deviation of Overall, Written, and Oral Language Scale Scores ..... 72

Table 7.8 Percentage of Students in Each Proficiency Level—Overall ..... 73

Table 7.9 Percentage of Students in Each Proficiency Level—Oral Language Composite .. 74

Table 7.10 Percentage of Students in Each Proficiency Level—Written Language Composite .....	75
Table 7.11 Demographic Student Groups Reported.....	76
Table 8.1 Number of Students Tested by Include Indicator and Grade Level .....	82
Table 8.2 Mean <i>p</i> -values .....	84
Table 8.3 Mean Point-Biserial Correlation .....	85
Table 8.4 Number of Items with Classical Item Analysis Flag by Domain .....	86
Table 8.5 Classical Item Statistics for Each Domain .....	86
Table 8.6 Mean Omit Rates.....	88
Table 8.7 IRT <i>a</i> -values (Discrimination Parameter) for 2018–2019 Oral Language Tests by Grade Level or Grade Span.....	89
Table 8.8 IRT <i>a</i> -values (Discrimination Parameter) for 2018–2019 Written Language Tests by Grade Level or Grade Span.....	90
Table 8.9 IRT <i>b</i> -values (Item Difficulty Parameter) for 2018–2019 Oral Language Tests by Grade Level or Grade Span.....	90
Table 8.10 IRT <i>b</i> -values (Item Difficulty Parameter) for 2018–2019 Written Language Tests by Grade Level or Grade Span.....	91
Table 8.11 MH Data Structure.....	94
Table 8.12 DIF Categories for MC.....	95
Table 8.13 DIF Categories for CR Items .....	96
Table 8.14 Gender DIF Classification.....	96
Table 8.15 Ethnicity DIF Classification .....	98
Table 8.16 Reliability Coefficient of Domains and Composite Scores .....	100
Table 8.17 SEM based on Classical Test Theory.....	101
Table 8.18 SEM Based on Scale Score .....	102
Table 8.19 Summary Rater Agreement by Possible Maximum Score Points and Grade Level or Grade Span .....	104
Table 8.20 Classification Accuracy and Consistency for Reported Composite Scores .....	106
Table 8.21 TCC Data for the Oral Language Composite (Figure 8.1) .....	110
Table 8.22 TCC Data for the Written Language Composite (Figure 8.2).....	132
Table 9.1 Correlation Among Two Composites and the Overall Score for the Summative ELPAC.....	157
Table 9.2 Correlation Among Composites and the Overall Score for the Initial ELPAC .....	157
Table 9.3 Correlation of Summative ELPAC Overall and Smarter Balanced for ELA Scores .....	158
Table 11.1 Number of Students, by Grade, Tested Across Years .....	167
Table 11.2 Overall Scale Score Mean and Standard Deviation Across Administrations ....	168
Table 11.3 Oral Scale Score Mean and Standard Deviation Across Administrations .....	168
Table 11.4 Written Scale Score Mean and Standard Deviation Across Administrations ....	169

## List of Figures

Figure 6.1 Summative ELPAC Score Reporting Hierarchy.....	50
Figure 7.1 Percentage of students at each overall proficiency level.....	73
Figure 7.2 Percentage of students at each oral language proficiency level.....	74
Figure 7.3 Percentage of students at each written language proficiency level .....	75
Figure 8.1 Test characteristic curve for the oral language composite.....	92
Figure 8.2 Test characteristic curve for the written language composite .....	93



**Acronyms and Initialisms Used in the *Summative English Language Proficiency Assessments for California Technical Report***

Term	Definition
AAF	Accessibility and Alternate Formats
AERA	American Educational Research Association
AIS	average item score
APA	American Psychological Association
AST	Administration and Scoring Training
CAASPP	California Assessment of Student Performance and Progress
CCR	<i>California Code of Regulations</i>
CDE	California Department of Education
CDS	county/district/school
CELDT	California English Language Development Test
COE	county office of education
CR	constructed response
CSEM	conditional standard error of measurement
DIF	differential item functioning
EC	<i>Education Code</i>
ECD	evidence-centered design
EL	English learner
ELA	English language arts/literacy
ELAS	English language acquisition status
ELD	English language development
ELD Standards	English Language Development Standards
ELP	English language proficiency
ELPAC	English Language Proficiency Assessments for California
eSKM	Enterprise Score Key Management
ESSA	Every Student Succeeds Act
ETS	Educational Testing Service
GIS	Group Information Sheet
GPC	generalized partial credit
IBIS	Item Banking Information System
IEP	individualized education program
IRT	item response theory
K	kindergarten
LEA	local educational agency
MC	multiple choice
MH	Mantel-Haenszel
MH DIF	Mantel-Haenszel differential item functioning
MSICL	multiple-selection inline choice list
NCME	National Council on Measurement in Education

Table of Acronyms and Initialisms (*continued*)

<b>Term</b>	<b>Definition</b>
NS	No Score
OIB	ordered item booklet
ONE	Online Network for Evaluation
OTI	Office of Testing Integrity
PAR	Psychometric Analysis & Research
PIN	problem item notification
PLD	performance level descriptor
Pre-ID	pre-identification
QC	quality control
RFCA	request for CDE approval
RFEP	reclassified fluent English proficient
RFR	request for review
SBE	State Board of Education
SCOE	Sacramento County Office of Education
SD	standard deviation
SEM	standard error of measurement
SFTP	secure file transfer protocol
SMD	standardized mean difference
SSR	Student Score Report
TCC	test characteristic curve
TIPS	Technology and Information Processing Services
TOMS	Test Operations Management System
USC	United States Code

# Chapter 1: Introduction

---

## 1.1. Background

The English Language Proficiency Assessments for California (ELPAC) “is the required state test for English language proficiency (ELP) that must be given to students whose primary language is a language other than English. State and federal laws require that local educational agencies (LEAs) administer a state test of ELP to eligible students in kindergarten through grade twelve” (California Department of Education [CDE], 2019). California *Education Code (EC)* Section 313(a) requires that the assessment of ELP be done upon initial enrollment and annually thereafter until the LEA reclassifies the student as English proficient.

In November 2012, the California State Board of Education (SBE) adopted the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards). The CDE began transitioning from the paper-pencil California English Language Development Test (CELDT) to the paper-pencil ELPAC in 2017–2018. The Summative ELPAC was used as the annual ELP assessment in spring 2018 and spring 2019. The Initial ELPAC was used for initial identification beginning in July 2018.

This technical report describes the development, administration, and results of the 2018–2019 administration of the Summative ELPAC.

## 1.2. Test Purpose

The ELPAC consists of two assessments: the Initial ELPAC and the Summative ELPAC. The Initial ELPAC identifies whether a student is initial fluent English proficient or an English learner (EL) who would benefit from additional instructional supports.

Students identified as ELs on the Initial ELPAC or previous state ELP assessments go on to take the Summative ELPAC, which is administered annually to students in kindergarten through grade twelve who have been identified as ELs. The Summative ELPAC has two purposes:

1. Determine the level of ELP of EL students
2. Assess the progress of EL students in acquiring the skills of listening, speaking, reading, and writing in English

## 1.3. Test Content

Under the Every Student Succeeds Act (ESSA) and California state regulations, students who are identified as ELs are required to take the ELPAC in the domains of Listening, Speaking, Reading, and Writing.

The content of the Summative ELPAC is aligned with the 2012 ELD Standards. The test content corresponds to the *Common Core State Standards: English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects* (Smarter Balanced, 2015a). Items on the Summative ELPAC also correspond to the *California Common Core State Standards for Mathematics* (Smarter Balanced, 2015b) as well as the *Next Generation Science Standards for California Public Schools, Kindergarten through Grade Twelve* (CDE, 2019).

## 1.4. Testing Window

*California Code of Regulations (CCR)*, Title 5, Section 11518(d), establishes the Summative ELPAC testing window from February 1 through May 31 annually, which was the testing window in 2018–2019. During this time period, any student identified as an EL must be administered the Summative ELPAC.

CCR Section 11518(m) establishes the Initial ELPAC testing window from July 1 through June 30 of each school year.

## 1.5. Intended Population

All students who previously took the CELDT or ELPAC, who were identified as ELs, and who were enrolled between February 1, 2019, and May 31, 2019, were required to take the Summative ELPAC. All students classified as ELs must be tested annually during the Summative ELPAC window until they are reclassified as fluent English proficient (RFEP) based on the CDE’s established guidelines for reclassification established by the SBE (EC 313[f]).

Students with disabilities whose individualized education program (IEP) or Section 504 plan specifies they cannot take one or more domains of the ELPAC with allowed universal tools, designated supports, or accommodations are eligible for a domain exemption(s). Students with the most significant cognitive disabilities who cannot access the ELPAC with approved accessibility resources are eligible to take a locally determined alternate assessment, as noted in their IEP.

## 1.6. Intended Use and Purpose of Test Scores

The SBE approved the reporting hierarchy of the Summative ELPAC in September 2017. Individual student scores for the Summative ELPAC for all grade levels (i.e., kindergarten through grade twelve) included

- an overall score based on a continuous scale,
- an oral language subscore that reflects performance on the Listening and Speaking domains based on a continuous scale,
- a written language subscore that reflects performance on the Reading and Writing domains based on a continuous scale, and
- the student’s proficiency within each domain (i.e., Listening, Speaking, Reading, and Writing) based on three reporting levels.

Each student who took the 2018–2019 paper-pencil Summative ELPAC received an overall score, an oral language subscore, and a written language subscore, which placed the student within one of the four ELPAC proficiency levels:

1. Beginning stage
2. Somewhat developed
3. Moderately developed
4. Well developed

The three scale scores—overall score, oral language subscore, and written language subscore—were all linked to the four ELPAC proficiency levels.

Similar scale scores across adjacent grade levels or grade spans and adjacent editions indicated a comparable degree of ELP. For example, similar scale scores on the grade two and grade span three through five assessments, or the grade span three through five and grade span six through eight assessments, indicate similar degrees of ELP. Further, similar scale scores from the 2018–2019 and 2019–2020 administrations will indicate similar degrees of ELP.

The purpose of the scale scores is to track student progress in ELP from year to year. For example, the scale scores will be used to track student progress in ELP from 2018–2019 to 2019–2020, once students who continue to be designated as ELs take the 2019–2020 Summative ELPAC.

## 1.7. Limitations of the Assessment

Students who are identified as ELs must be tested annually during the annual assessment window—February 1 through May 31—until they are reclassified as RFEP. Because the Summative ELPAC is the ELP assessment developed pursuant to *EC* Section 60810, scores from the Summative ELPAC are one set of criteria used to determine whether individual students qualify for RFEP on the basis of the reclassification process developed by the LEA.

ELPAC results are not used to measure academic achievement. Instead, results from the Summative ELPAC may be used to plan for instruction and to assist LEAs and schools in the ongoing process of program monitoring and evaluation.

## 1.8. Groups and Organizations Involved with the ELPAC System

### 1.8.1. State Board of Education

The SBE is the state agency that establishes educational policy for kindergarten through grade twelve in the areas of standards, instructional materials, assessment, and accountability. The SBE adopts textbooks for kindergarten through grade eight, adopts regulations to implement legislation, and has the authority to grant waivers of the *EC*.

In addition to adopting the rules and regulations for itself, its appointees, and California’s public schools, the SBE also is the state educational agency responsible for overseeing California’s compliance of the ESSA and the state’s Public School Accountability Act, which measures the academic performance and progress of schools on a variety of academic metrics (CDE, 2020a).

### 1.8.2. California Department of Education

The CDE oversees California’s public school system, which is responsible for the education of more than 6,180,000 children and young adults in more than 10,500<sup>1</sup> schools. California aims to provide a world-class education for all students, from early childhood to adulthood. The CDE serves the state by innovating and collaborating as a team with educators, school staff, parents/guardians, and community partners to prepare students to live, work, and thrive in a highly connected world.

---

<sup>1</sup> Retrieved from the CDE Fingertip Facts on Education in California - *CalEdFacts* web page at <https://www.cde.ca.gov/ds/sd/cb/ceffingertipfacts.asp>

Within the CDE, the Instruction & Measurement Branch oversees programs promoting innovation and improving student proficiency. Programs include oversight of statewide assessments and the collection and reporting of educational data (CDE, 2020b).

### 1.8.3. California Educators

A variety of California educators, including school administrators and teachers experienced in teaching ELs, were selected based on their qualifications, experiences, demographics, and geographic locations and were invited to participate in the ELPAC development process. In this process, California educators participated in tasks that included defining the purpose and scope of the assessment, assessment design, item development, standard setting, score reporting, and scoring constructed-response items.

Refer to [3.2.2 Composition of the ELPAC Item Writer and Item Review Meetings and Participant](#) Qualifications for the composition of the item writer training and item review meetings. Refer to [5.2.3.2.1 Composition of ELPAC Speaking Range Finding and Writing Range Finding Meetings and Participant Qualifications](#) for the composition of the Speaking range finding and the Writing range finding meetings.

### 1.8.4. Contractors

#### 1.8.4.1. Primary Contractor—Educational Testing Service

The CDE and the SBE contract with ETS to develop and administer both the Summative ELPAC and the Initial ELPAC. As the prime contractor, ETS has the overall responsibility for working with the CDE to implement and maintain an effective assessment system and to coordinate the work of ETS with its subcontractors. Activities directly conducted by ETS include, but are not limited to, the following:

- Providing management of the program activities
- Providing tiered help desk support to LEAs
- Developing high-quality items that are aligned to the 2012 ELD Standards
- Constructing, producing, and controlling the quality of ELPAC test forms and related testing materials, including grade- and content-specific *Examiner's Manuals*
- Hosting and maintaining a website with resources for the ELPAC
- Developing, hosting, and providing support for the Test Operations Management System
- Producing and distributing Student Score Reports
- Developing a public score reporting website
- Completing all psychometric procedures

#### 1.8.4.2. Subcontractor—Sacramento County Office of Education

ETS contracted with the Sacramento County Office of Education to manage all activities associated with training and outreach, including the following:

- Supporting and training county offices of education, LEAs, and charter schools
- Developing informational materials
- Recruiting and logistics for educator trainings
- Producing training videos

## 1.9. Overview of the Technical Report

This technical report addresses the characteristics of the ELPAC administered in spring of the 2018–2019 school year and contains 11 additional chapters, as follows:

- [Chapter 2](#) presents an overview of the processes involved in a testing cycle for the Summative ELPAC. This includes item development, test construction, test administration, test participation, and accessibility.
- [Chapter 3](#) describes the procedures followed during item development, various reviews (e.g., item content and bias and sensitivity reviews), and the process of item review.
- [Chapter 4](#) describes the process of test assembly, including the content being measured, as well as the content and psychometric criteria. Also discussed is materials development.
- [Chapter 5](#) details the processes involved in the actual 2018–2019 administration. It also describes the procedures followed to maintain test security throughout the test administration process.
- [Chapter 6](#) summarizes the standard setting process that established the base year (i.e., 2017–2018) ELPAC scores. Details include the performance level descriptors, an overview of the standard setting methodology, and the process to establish the threshold scores that define the score ranges for each Summative ELPAC level. These standard setting processes were based on student testing results from the spring 2017 stand-alone field test, which occurred between March and April 2017.
- [Chapter 7](#) provides information on the scoring processes and summarizes the types of scores and score reports.
- [Chapter 8](#) summarizes the statistical procedures and results for 2018–2019, including
  - classical item analysis;
  - differential item functioning analysis;
  - item response theory calibration, linking, and scaling;
  - reliability analyses; and
  - analyses of the consistency and accuracy of the performance level classifications.
- [Chapter 9](#) discusses the procedures designed to ensure the validity of score uses and interpretations.
- [Chapter 10](#) highlights the quality-control processes used at various stages of the 2018–2019 Summative ELPAC administration, including item development, test assignment, test administration, scoring procedures, psychometric analysis processes, and score reporting.
- [Chapter 11](#) presents year-to-year comparisons for the Summative ELPAC operational administration between 2017–2018 and 2018–2019.
- [Chapter 12](#) details the ongoing means of program improvement.

## References

- California Department of Education. (2014). *California 2012 English Language Development Standards: Kindergarten Through Grade 12*. Adopted by the State Board of Education November 2012. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/sp/el/er/documents/eldstndpublication14.pdf>
- California Department of Education. (2019) *NGSS for California public schools, K-12*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/pd/ca/sc/ngssstandards.asp>
- California Department of Education. (2020b, October). *State Board of Education responsibilities*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/be/ms/po/index.asp>.
- California Department of Education. (2020a, July). Sacramento, CA: California Department of Education. *Organization*. Retrieved from <https://www.cde.ca.gov/re/di/or/>
- Smarter Balanced Assessment Consortium. (2015a). *Content specifications for the summative assessment of the Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Los Angeles, CA: Smarter Balanced Assessment Consortium.
- Smarter Balanced Assessment Consortium. (2015b). *Content specifications for the summative assessment of the Common Core State Standards for mathematics*. Los Angeles, CA: Smarter Balanced Assessment Consortium.



## Chapter 2: Overview of ELPAC Development Processes

---

This section describes the processes used to develop a high-quality pool of items, assemble tests, administer tests, and provide accommodations as required by *Education Code* Section 60810 for the 2018–2019 Summative English Language Proficiency Assessments for California (ELPAC).

### 2.1. Item Development

To construct test forms for the 2018–2019 Summative ELPAC, an appropriate pool of items needed to be developed. The item development process started with the creation of item development specifications, which described the quantity of items to be created and the process to be followed. After the item development specifications were reviewed and approved by the California Department of Education (CDE), Educational Testing Service (ETS) assessment specialists worked with two groups to draft test items: California educators and ETS contractors. In February 2016, California educators attended an item writer training workshop where they received training, generated lists of topics for items, and drafted items. The educators focused on the development of Speaking and Writing items as well as shorter Listening and Reading items.

At the same time, the ETS contractors developed topics for longer Listening sets and Reading sets. ETS then compiled the topics from both groups and submitted them to the CDE for review. Once approved, the topics for the longer sets were sent to five ETS contractors with prior experience in developing Listening and Reading sets. The contractors then submitted their draft items to ETS for review.

All items drafted by California educators and ETS contractors went through internal ETS reviews, including two content reviews, a fairness review, and an editorial review. The items were then submitted to the CDE for review and approval.

Each item was then reviewed during two educator meetings: a Content Review Panel meeting and a Bias and Sensitivity Review Panel meeting.

During the Content Review Panel meeting, California educators considered whether each item would appropriately measure the aligned standard(s), whether each item was appropriate for the designated grade level or grade span, and whether each item was presented clearly and effectively. Multiple-choice (MC) items were also reviewed to ensure that each one had a single best key and distractors that were all plausible yet wrong. In addition, constructed-response (CR) items were reviewed to make sure that each prompt would elicit a response that allowed students to demonstrate their language abilities, as described by the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014).

During the Bias and Sensitivity Review Panel meeting, educators considered whether each item was free of content that was potentially biased against, or offensive to, any identified group, such as students from other countries or students who are deaf or hard of hearing. If an item contained potentially biased or offensive content, the educators considered whether the item could be revised to remove the potentially biased or offensive content.

Educators at both the Content Review Panel meeting and the Bias and Sensitivity Review Panel meeting had the option of making one of three decisions regarding each item: approve the item as is, approve the item with revisions, or reject the item. Whenever an item was approved with revisions, educators specified the revisions needed to text or images and the reasons for the proposed revisions.

After the educator meetings, CDE staff reviewed the proposed revisions and made final decisions as to whether each educator's proposed revisions should be implemented. ETS assessment specialists then applied the CDE-approved revisions.

After the items were revised, CDE staff confirmed that revisions were entered correctly. After ETS implemented any necessary final revisions, the CDE approved the items for use as field test items.

All items that were used in the 2018–2019 Summative ELPAC were administered in the 2016–2017 stand-alone field test and approved for operational use as described in subsection [2.2.4 Psychometric Review](#).

### 2.1.1. Item Format

The 2018–2019 paper-pencil Summative ELPAC contained three item formats: (1) single-selection MC items, (2) multiple-selection inline choice list (MSICL) items, and (3) CR items.

1. **MC items** contained a question that was followed by three or four options as answer choices, one of which was the correct option.
2. **MSICL items**, which were found in the kindergarten (K) Reading test, contained a series of questions. After the test examiner assessed each of the student's responses to the series of MSICL questions as correct, incorrect, or no response, scoring rules were used to assign the student with full, partial, or no credit. This item format was treated the same as the CR item format for statistical analysis.
3. **CR items** consisted of a prompt that elicited either a spoken response or a written response. A rubric was used to assess the quality of the response on a scale of 0–1, 0–2, 0–3, or 0–4. The rubrics described typical characteristics of a response at each score point based on criteria that were derived from the 2012 ELD Standards (CDE, 2014).

### 2.1.2. Item Writing Guidelines

Item writing guidelines were developed to define the task types and content of the items to provide guidance to item writers and drive consistency and efficiency in item development. The guidelines were used to facilitate the development of comparable items that measure appropriate skills and content aligned with the 2012 ELD Standards (CDE, 2014).

### 2.1.3. Item Banking

After items were drafted, they were entered in the ETS Item Banking Information System (IBIS). IBIS contains fields for entering item content and information about items for MC and CR items. IBIS was used to store item text, graphics, scripts for audio recordings, scoring information, statistical information, and metadata. After ETS staff drafted and reviewed items in IBIS, CDE staff used IBIS to review items in preparation for educator reviews and to ensure that ETS had revised items accurately after the educator reviews.

## 2.2. Test Assembly

ETS assessment specialists assembled the Summative ELPAC tests, which were reviewed and approved by the CDE. This process began with the creation of test development specifications, which described the content characteristics, psychometric characteristics, and quantity of items to be used in the 2018–2019 Summative ELPAC. ETS created the test development specifications that the CDE reviewed and approved.

After the test development specifications were approved, ETS assessment specialists assembled the tests in IBIS according to the specifications. IBIS then generated form planners, which are spreadsheets containing essential item information such as the number of items, the alignment of items according to the 2012 ELD Standards, and the keys to MC items. ETS assessment specialists and psychometricians reviewed the form planners for essential item information such as keys, maximum score points, content standards, and alignment with the test blueprint before they were delivered to the CDE for review. The CDE reviewed and approved the form planners after ETS revised the form planners as needed.

### 2.2.1. Test Design

The Summative ELPAC is administered to students in the following grade levels and grade spans: K, grade one, grade two, grades three through five, grades six through eight, grades nine and ten, and grades eleven and twelve.

Four domains of English language proficiency (ELP) were assessed in the 2018–2019 Summative ELPAC: Listening, Speaking, Reading, and Writing. Students in K and grade one were tested one-on-one in all four domains. Students in grade two were tested one-on-one in the Speaking domain. In the Listening, Reading, and Writing domains, grade two students were tested in small groups of up to 10 students.

Students in grade spans three through five, six through eight, nine and ten, and eleven and twelve were tested one-on-one in the Speaking domain and in a group administration in the Listening, Reading, and Writing domains. A proctor assisted the test examiner during test administrations to groups comprised of more than 20 students.

### 2.2.2. Test Blueprints

Test blueprints were developed to describe the content of the Summative ELPAC. The test blueprints contain four tables with information about the task types in each of the four language domains of Listening, Speaking, Reading, and Writing. Task types are individual items or sets of items that require a student to perform an activity to elicit information about the student's ELP.

The test blueprints provide information about the number of items and points that are administered per task type within each grade level and domain. The Summative ELPAC test blueprints also provide two types of alignment between task types and the standards: "primary" and "secondary." Primary alignment indicates there is a close or strong match in terms of the language knowledge, skills, and abilities covered by both the task type and the standard. Secondary alignment indicates that there is a moderate or partial match between the standard and the item in terms of language knowledge, skills, and abilities.

### 2.2.3. Test Length

Because the blueprints identify the numbers of items to be tested within each domain, they govern test length. When the Summative ELPAC test blueprints were developed, the goal was to include sufficient numbers of items to provide valid and reliable assessments of ELP,

while keeping the administration time at a reasonable level. The number of items increases from K through grade span three through five to make the length of the test appropriate for students as they gain the ability to focus for longer periods of time.

The Summative ELPAC is an untimed test. Estimated administration times were provided in the *Examiner's Manuals*, but only as a basis for planning, because students were allowed as much time as they need to complete their responses in each domain. Additionally, the testing schedule might have been altered to give students sufficient breaks to avoid fatigue, and testing may be administered over the course of several days.

Test examiners were trained to administer an entire domain in a single sitting except for the Reading and Writing domains at grades three through twelve, which could have been administered in either one or two sittings.

#### **2.2.4. Psychometric Review**

All operational items in the 2018–2019 Summative ELPAC were field tested in the 2016–2017 stand-alone field test. After the administration of the field test, all items underwent statistical analysis. The ETS Psychometric Analysis & Research group used student responses to compile classical item statistics and flag any items that fell outside of acceptable parameters. Assessment specialists reviewed each flagged item and made one of three recommendations:

- Keep the flagged item as is and classify it as operationally ready
- Revise the flagged item and classify it as field-test-ready for a future form
- Reject the flagged item and discontinue using it

Items that were classified as operationally ready were used to develop the 2018–2019 Summative ELPAC.

After assessment specialists assembled the tests, ETS psychometricians reviewed the statistical characteristics of the tests to ensure that the full range of ELP would be assessed. Tests were revised, if needed, based on the feedback of ETS psychometricians, and then the tests were submitted to CDE psychometricians for review. ETS assessment specialists made any further revisions needed to the tests to obtain approval from the CDE psychometricians.

#### **2.2.5. CDE Review**

The CDE used a two-stage process to review all test materials: (1) request for review (RFR) and (2) request for CDE approval (RFCA). Test materials for review and approval by the CDE included form planners, *Examiner's Manuals*, Test Books, Answer Books, braille versions of *Examiner's Manuals* and forms, and large-print versions of forms. All test materials were approved at RFCA before they were submitted to vendors for reproduction.

For the first stage, ETS initiated the review by submitting an RFR to the CDE. CDE consultants performed the initial review and returned comments and requests for revisions to ETS. ETS staff then revised the documents as requested and returned them to the CDE consultants, who then reviewed the updated materials. If the test materials needed additional revisions, they were returned to ETS for further modifications.

For RFCA, if the CDE consultants approved the test materials during the RFR stage, then the CDE submitted the test materials to the CDE administrator with an RFCA. Test materials that were approved with revisions were revised by ETS and resubmitted for approval. Test materials that were not approved needed significant revisions and had to be submitted to

the consultants for RFR again before they could be resubmitted for RFCA. Test materials that were approved without edits moved on to the composition phase.

## 2.3. Test Administration

Standardization and security of the ELPAC is of utmost importance in order to maintain the integrity and validity of the assessment. ELPAC test administration manuals provided information to local educational agencies (LEAs) and testing personnel on how to efficiently receive, organize, administer, and return test materials for scoring.

### 2.3.1. Test Security and Confidentiality

All testing materials for the 2018–2019 Summative ELPAC—Test Books, Answer Books, and *Examiner’s Manuals*—were considered secure documents. Every person having access to test materials was required to maintain the security and confidentiality of the test materials. ETS’ Code of Ethics requires that all test information, including tangible materials (e.g., test booklets, test questions, test results), confidential files, processes, and activities are kept secure.

To ensure security for all tests that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI).

In the pursuit of enforcing secure practices, ETS and the OTI strive to safeguard the various processes involved in a test development and administration cycle. For the 2018–2019 Summative ELPAC, those processes included the following:

- Test development
- Item and data review
- Item banking
- Transfer of forms and items to the CDE
- Security of electronic files using a firewall
- Printing and publishing
- Test administration
- Test delivery
- Processing and scoring
- Data management
- Transfer of scores via secure data exchange
- Statistical analysis
- Reporting and posting results
- Student confidentiality
- Student test results

### 2.3.2. Procedures to Maintain Standardization

ELPAC processes were designed so the tests are administered and scored in a standardized manner. ETS took all necessary measures to ensure the standardization of the ELPAC, as described in this subsection.

#### 2.3.2.1. Test Administration

Roles and responsibilities for each person involved in the ELPAC administration were defined in the *Summative ELPAC Test Administration Manual* (CDE, 2019). Providing clear definitions and delineation for each role ensured test security and standardized administration. These processes are discussed in more detail in subsection [5.1 Procedures to Maintain Standardization](#).

### 2.3.2.2. Test Directions

A series of instructions compiled in detailed manuals is provided to testing personnel. For the 2018–2019 Summative ELPAC, such documents included, but were not limited to, the following:

**Examiner’s Manuals**—These were grade-level manuals used by test examiners to administer the ELPAC to students and were to be followed exactly so that all students have an equal opportunity to demonstrate their level of English proficiency. (Refer to [5.1.4.1 Examiner’s Manuals](#) in [chapter 5](#) for more information.)

**Summative ELPAC Test Administration Manual**—This manual contained test administration procedures for LEA ELPAC coordinators and site ELPAC coordinators (CDE, 2019). (Refer to [5.1.4.2 Summative ELPAC Test Administration Manual](#) in [chapter 5](#) for more information.)

**Test Operations Management System (TOMS) Guide for the ELPAC**—This manual provided instructions for LEA ELPAC and site ELPAC coordinators to perform tasks in TOMS in support of the program, including providing organization information, adding and managing users, searching and viewing student information, ordering pre-identification labels that contain student demographic data, ordering test materials, viewing reports, and accessing audio files (CDE, 2018a). (Refer to [5.1.4.3 TOMS Guide for the ELPAC](#) in [chapter 5](#) for more information.)

## 2.4. Test-Taking Population

California *Education Code* Section 313 requires LEAs to administer the Initial ELPAC to all eligible students in K through grade twelve whose primary language is a language other than English. LEAs are required to administer the Summative ELPAC annually to students identified as English learners until they are reclassified as fluent English proficient.

Table 2.A.1 through table 2.A.4 in [appendix 2.A](#) provide the number of test takers and the percent of test takers and select demographic groups for each test during the 2018–2019 administration. Note that the data in the *Number Registered* column includes students who were registered within a grade and eligible for the Summative ELPAC during the 2018–2019 administration. The *Number Tested* columns include students who tested at the current grade level and exclude off-grade testers and students registered who did not test.

## 2.5. Accessibility

To ensure a fair and valid testing experience for all students who took the 2018–2019 Summative ELPAC, ETS provided accessible versions of the test materials for each ELPAC grade level and grade span. Braille and large-print test materials were available for students who have a Section 504 plan or whose individualized education program (IEP) indicated a need for an accommodated version of the ELPAC.

### 2.5.1. Resources for Selection of Accessibility Resources

The CDE developed Matrix Four to list the student accessibility resources available during administration of both the Summative ELPAC and the Initial ELPAC, and then made the document available on the CDE’s website. Matrix Four assists LEAs in understanding the accessibility resources that are available for an ELPAC administration and follows a three-

tiered accessible approach that includes universal tools, designated supports, and accommodations (CDE, 2018b).<sup>2</sup>

The following types of accessibility resources were available to students taking the 2018–2019 Summative ELPAC:

- **Universal tools** were available to *all* students on the basis of student preference and selection.
- **Designated supports** were available to *all* students when determined for use by an educator or team of educators—with parent/guardian and student input, as appropriate—or specified in the student’s IEP or Section 504 plan.
- **Accommodations** were to be permitted on ELPAC tests to all eligible students if specified in the student’s IEP or Section 504 plan.

Accessibility resources allowed all students to show what they know and can do. These resources were not intended to give a testing advantage, but, rather, to allow students the opportunity for a fair and valid testing experience.

### 2.5.2. Delivery of Accessibility Resources

ELPAC test materials were available in braille and large-print for each ELPAC grade level and grade span. Additionally, as noted previously, Matrix Four outlined the accessibility resources that are permitted during the ELPAC administration.

The percentages of accessibility resources used during the 2018–2019 Summative ELPAC field test by grade level and for each domain are presented in [appendix 2.B](#), in table 2.B.1.

### 2.5.3. Unlisted Resources

Unlisted resources are not universal tools, designated supports, or accommodations. Unlisted resources are made available if specified in an eligible student’s IEP or Section 504 plan and only on approval by the CDE. Part 4 of Matrix One contains some unlisted resources for the Initial ELPAC that change the construct being tested (CDE, 2018b).

To request the use of an unlisted resource for the 2018–2019 Summative ELPAC, the LEA ELPAC coordinator submitted a request to the CDE a minimum of 10 business days before the student’s first day of testing. The CDE replied to the request within four business days.

Approval of an unlisted resource that was not previously identified may have been granted by the CDE on the basis of the IEP team’s or Section 504 plan’s designation and if the unlisted resource did not compromise test security. Prior to administration, the CDE determined if the unlisted resource changed the construct being measured. If so, the LEA ELPAC coordinator was instructed to mark *Alternate Assessment* on the Answer Book for all affected domains. The student received the lowest obtainable scale score for any domains in which *Alternate Assessment* was marked.

---

<sup>2</sup> This technical report is based on the version of Matrix Four that was available during the 2018–2019 Summative ELPAC administration.



## References

- California Department of Education. (2014). *California 2012 English Language Development Standards: Kindergarten Through Grade 12*. Adopted by the State Board of Education November 2012. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndspublication14.pdf>
- California Department of Education. (2018b). *Matrix Four: Universal tools, designated supports, and accommodations for the English Language Proficiency Assessments for California*. Retrieved from <https://www.cde.ca.gov/ta/tg/ca/accessibilityresources.asp>
- California Department of Education. (2018a). *Test Operations Management System Guide for the ELPAC*. Sacramento, CA: California Department of Education. Retrieved from <https://elpac.org/s/pdf/ELPAC.TOMS-users-guide.2018-19.pdf>
- California Department of Education. (2019). *Summative ELPAC Test Administration Manual*. Sacramento, CA: California Department of Education. Retrieved from <https://elpac.org/s/pdf/ELPAC.Test-Admin-Manual-Summative.2018-19.pdf>



## Chapter 3: Item Development and Review

---

### 3.1. Item Development

This section describes the work performed to develop a high-quality pool of items for the English Language Proficiency Assessments for California (ELPAC).

#### 3.1.1. Overview

The ELPAC technical proposal from Educational Testing Service (ETS) stated that ETS would undertake three item development efforts during the first ELPAC contract from 2015 through 2018. The first item development work in 2015–2016 generated a total of 2,008 items. Because of the robust nature of the first item development work, the next two item development efforts, which would have yielded about 100 items each, were combined into a single effort to develop approximately 200 field test ready items that could be used as embedded field test items in the 2018–2019 Summative ELPAC.

All of the operational items in the 2018–2019 Summative ELPAC came from the first item development effort of 2015–2016. The second item development effort of 2017–2018 was the main source of items that were selected for embedded field testing. This section describes the item development specifications, task types that were developed, and the development of the items for both item development efforts.

#### 3.1.2. Item Development Specifications

Item development specifications were created by ETS and then reviewed and approved by the California Department of Education (CDE) prior to both item development efforts. The item development specifications described the plans to develop approximately 2,000 field test-ready items from 2015–2016 and another 200 field test-ready items from 2017–2018. The specifications included details about the types of items to be developed, the people who developed the items, the item development processes, and item security measures. The newly developed items were designed to align with the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014); meet the contents of Summative ELPAC test blueprints (CDE, 2015); and meet the validity, reliability, and high technical quality criteria for a high-stakes standardized state assessment system, as required by California *Education Code* Section 60810.

#### 3.1.3. Test Blueprints

The State Board of Education adopted the *Proposed Test Blueprints for the ELPAC* on November 4, 2015, which was prior to the first piloting of the ELPAC items (CDE, 2015). The pilot results provided crucial input for the refinement and streamlining of the Summative ELPAC test blueprints. As ETS and the Sacramento County Office of Education (SCOE) recommended in the *Report on the First Pilot of Items for the ELPAC* (CDE, 2016a), the ELPAC test blueprints were revised to include those task types that best elicited the types of responses needed to assess students' English language abilities as described by the 2012 ELD Standards (CDE, 2014).

Before the first pilot, the test blueprints contained a total of 32 proposed task types across the domains of Listening, Speaking, Reading, and Writing. During the evaluation of pilot data, six task types were removed and one was added, making a total of 27 task types in the *Proposed Test Blueprints for the ELPAC* (CDE, 2015).

### 3.1.4. Task Types

The 2018–2019 Summative ELPAC contained 27 task types. Each task type required a student to perform an activity to elicit information about the student’s English language proficiency (ELP). Each task type had one or more items that aligned with the 2012 ELD Standards. While the 2012 ELD Standards are organized according to three modes of communication (collaborative, interpretive, and productive communication), federal Title I requirements of the Every Student Succeeds Act of 2015 call for scores to be reported according to the four language domains of Listening, Speaking, Reading, and Writing.

The Listening domain of the Summative ELPAC had five task types, the Speaking domain had six task types, the Reading domain had nine task types, and the Writing domain had seven task types. When a task type required the use of integrated language skills, such as listening and speaking, the task type was classified according to the language skill used to provide the response. For instance, the task type *Summarize an Academic Presentation* required a student to listen to a presentation and then summarize the presentation by speaking to the test examiner. Because the student provided the summary as a spoken response, the task type was classified as a Speaking task type.

The next subsections describe the task types used to assess ELP within each domain of the Summative ELPAC.

#### 3.1.4.1. Listening Task Types

Listening task types assessed the ability of an English learner (EL) to comprehend spoken English (conversations, discussions, and oral presentations) in a range of social and academic contexts. Students listened to a stimulus and then demonstrated their ability to actively listen by answering multiple-choice (MC) questions. For the 2018–2019 Summative ELPAC, test examiners used scripts in the *Examiner’s Manuals* to read Listening stimuli aloud to kindergarten (K) through grade two students. Students at grades three through twelve heard audio recordings of the Listening stimuli. The following are descriptions of the stimuli provided for the five Listening task types:

- **Listen to a Short Exchange, K through grade twelve:** Students heard a two-turn exchange between two speakers and then answered a question about the exchange.
- **Listen to a Classroom Conversation, grades three through twelve:** Students heard a multiple-turn conversation between two speakers and then answered three questions about the conversation.
- **Listen to a Story, K through grade five:** Students heard a multiple-turn conversation between two speakers and then answered three questions about the conversation.
- **Listen to an Oral Presentation, K through grade twelve:** Students heard an oral presentation on an academic topic and then answered three to four questions about the presentation.
- **Listen to a Speaker Support an Opinion, grades six through twelve:** Students heard an extended conversation between two classmates. In the conversation, one classmate made an argument in support of an opinion or academic topic. After listening to the conversation, students answered four questions.

### 3.1.4.2. Speaking Task Types

Speaking task types assessed the ability of an EL to express information and ideas and to participate in grade-level conversations and class discussions. All task types included one or more constructed-response items. Test examiners scored student responses in the moment using scoring rubrics. The following are descriptions of the six Speaking task types:

- **Talk About a Scene, K through grade twelve:** The student was presented with an illustration of a familiar scene. The test examiner first asked three who-, what-, and when-type questions about the scene. The test examiner then administered three items intended to generate longer responses.
- **Speech Functions, grades two through twelve:** Students stated what they would say in a situation described by the test examiner.
- **Support an Opinion, K through grade twelve:** The student listened to a presentation about two activities, events, materials, or objects, and was asked to give an opinion about why one was better than the other. At K, grade one, grade two, and grade span three through five, students viewed a picture of the choices for context and support.
- **Retell a Narrative, K through grade five:** The student listened to a story that followed a series of pictures, and then the student used the pictures to retell the story.
- **Present and Discuss Information, grades six through twelve:** The student viewed a graph, chart, or image that provided information. The student was prompted to read the information and then asked to respond to two prompts. The first prompt asked for a summary of the information in the graph, chart, or image. The second prompt asked for the students to state whether a claim was supported or unsupported based on the information in the graph or chart.
- **Summarize an Academic Presentation, K through grade twelve:** The student listened to an academic presentation while looking at a related picture(s). The student was prompted to summarize the main points of the presentation using the illustration(s) and key terms of the presentation, if provided.

### 3.1.4.3. Reading Task Types

Reading task types assessed the ability of an EL to read, analyze, and interpret a variety of grade-appropriate literary and informational texts. The following are descriptions of the nine Reading task types:

- **Read-Along Word with Scaffolding, K:** With scaffolding from the test examiner, the student provided the individual letter names and the initial letter sound for a decodable word. The student then answered a comprehension question about the word.
- **Read-Along Story with Scaffolding, K:** The student listened and followed along as the test examiner read aloud a literary text accompanied by three pictures for context and support. The student then answered a series of comprehension questions about the story.

- **Read-Along Information, K:** The student listened and followed along as the test examiner read aloud an informational text accompanied by three pictures for context and support. The student then answered a series of comprehension questions about the information.
- **Read and Choose a Word, grades one and two:** The student read three words and chose the word that matched a picture.
- **Read and Choose a Sentence, grades one through five:** The student read three or four sentences and chose the sentence that best described a picture.
- **Read a Short Informational Passage, grades one through twelve:** The student read a short informational text and answered MC questions related to the text.
- **Read a Student Essay, grades three through twelve:** The student read an informational essay presented as if written by a peer and answered a set of MC questions related to the essay.
- **Read a Literary Passage, grades one through twelve:** The student read a literary text and answered MC questions related to the text.
- **Read an Informational Passage, grades one through twelve:** The student read an informational text and answered MC questions related to the text.

#### 3.1.4.4. Writing Task Types

Writing task types assessed the ability of an EL to write literary and informational texts to present, describe, and explain information. The following are descriptions of the seven Writing task types:

- **Label a Picture—Word, with Scaffolding, K:** With scaffolding from the test examiner, the student wrote labels for objects displayed in a picture.
- **Write a Story Together with Scaffolding, K through grade two:** With scaffolding from the test examiner, the student collaborated with the test examiner to jointly compose a short literary text by adding letters, words, and a sentence to a story.
- **Write an Informational Text Together, grades one and two:** With scaffolding from the test examiner, the student listened to a short informational passage and then collaborated with the test examiner to jointly compose a text about the passage by writing a dictated sentence and an original sentence about the topic.
- **Describe a Picture, grades one through twelve:** In grades one and two, the student looked at a picture and wrote a brief description about what was happening. In grades three through twelve, the student looked at a picture and was prompted to examine a paragraph written by a classmate about what was happening in the picture. The student was asked to expand, correct, and combine different sentences written by a classmate before completing the final task of writing a sentence explaining what the students will do next.
- **Write About an Experience, grades three through twelve:** The student was provided with a common topic, such as a memorable classroom activity or event, and was prompted to write about the topic.

- **Write About Academic Information, grades three through twelve:** The student interpreted academic information from a graphic organizer created for a group project and answered two questions about it.
- **Justify an Opinion, grades three through twelve:** The student was asked to write an essay providing a position and appropriate supporting reasons about a school-related topic.

### 3.1.5. Updates to Item Writing Guidelines

The first pilot of the ELPAC items provided a wealth of experience with new ELPAC task types that informed subsequent item writer training and item-development efforts. ETS assessment specialists used data from the pilot to refine task types and develop descriptions of the ELPAC task types in the *Item Writing Guidelines for the ELPAC* (CDE, 2016b). These guidelines were used to train California educators to develop additional items for the ELPAC item pool at the Item Writer Training for California Educators from February 22, 2016, through February 25, 2016 and from November 6 through November 9, 2017.

## 3.2. Item Review Process

### 3.2.1. Overview

In partnership with SCOE, ETS convened ELPAC item writer trainings and item review panels to develop test items for both the Initial ELPAC and the Summative ELPAC. Select California educators were trained to write new items for the ELPAC. In addition, ETS trained a small group of experienced contractors to draft ELPAC items. After the items went through ETS internal and CDE reviews, California educators reviewed the items during Content Review Panel and Bias and Sensitivity Review Panel meetings. This subsection describes how California educators were selected and the process used to develop items for the ELPAC.

### 3.2.2. Composition of the ELPAC Item Writer and Item Review Meetings and Participant Qualifications

California educators participated in item writer workshops and item review meetings. Participant groups consisted of current and former teachers, resource specialists, administrators, curricular experts, and other education professionals. Minimum qualifications to be invited to participate were

- three or more years of teaching experience in kindergarten through grade twelve,
- expertise in language acquisition or experience teaching ELs in kindergarten through grade twelve,
- knowledge of and experience working with the 2012 ELD Standards,
- bachelor's or higher degree, and
- knowledge of and experience with the California content standards in language arts.

Preferred qualifications included

- a teaching credential authorization for ELD, specially designed academic instruction in English, or content instruction delivered in the primary language (e.g., Cross-cultural, Language, and Academic Development Certificate; Bilingual, Cross-cultural, Language, and Academic Development Certificate),

- specialized teaching certification in reading (e.g., Reading Certificate; Reading and Language Arts Specialist Certificate), and
- experience writing or reviewing test items for standardized tests, especially tests for K–12 ELs.

Every effort was made to ensure that groups of item reviewers included a representation of genders and of the geographic regions and ethnic groups in California.

[Table 3.1](#) shows the educational qualifications, present occupation, and credentials of the individuals who participated in an ELPAC item writer workshop or item review meeting.

**Table 3.1 ELPAC Item Writer Workshop (IWW) and Item Review Meeting (IRM) Qualifications, by Meeting Type and Total**

Qualification Type	Qualification	IWW	IRM	Total
Occupation	Classroom teacher	5	14	19
Occupation	English learner or literacy specialist or coach	9	18	27
Occupation	EL instructional specialist	0	1	1
Occupation	School administrator	4	1	5
Occupation	LEA or county office employee	0	5	5
Highest degree earned	Bachelor's degree	1	6	7
Highest degree earned	Master's degree	16	24	40
Highest degree earned	Doctorate	1	7	8
K–12 teaching credential	Elementary Teaching (multiple subjects)	7	16	23
K–12 teaching credential	Secondary Teaching (single subject)	1	5	6
K–12 teaching credential	Language Development Specialist	1	3	4
K–12 teaching credential	English Learner (CLAD, BCLAD)	17	16	33
K–12 teaching credential	Other	0	4	4

**Note:** Numbers may not match the totals because participants may have multiple occupations or teaching credentials, or are currently working toward earning their highest degree. The information is self-reported and may not reflect all of the experience and earned credentials.

### 3.2.3. Selection of Item Writers

California educators were recruited through email communications and by letter. To ensure broad representation, an email message and letter announcing the opportunities to write items and to review items were sent by the CDE to the following groups:

- The CDE's ELPAC listserv (includes California English Language Development Test District Coordinators and Title III county leads)
- The Bilingual Coordinators Network
- The CDE's California Assessment of Student Performance and Progress Coordinator listserv
- The ELPAC Technical Advisory Group

The email and letter directed applicants to fill in an online application in SurveyMonkey, a third-party, online survey provider. The application allowed California educators to apply for any or all of the events. The information from the application was loaded into a database that was used for the review and selection process.

During the selection process, applications were selected from current and retired California educators who had the following minimum qualifications:

- Bachelor's degree
- Expertise in language acquisition or experience teaching ELs in K through grade twelve
- Knowledge of, and experience working with, the 2012 ELD Standards

Additional desirable qualifications included the following:

- A teaching credential authorization for English language development, specially designed academic instruction in English, or content instruction delivered in the primary language (e.g., Cross-cultural, Language, and Academic Development Certificate; or Bilingual, Cross-cultural, Language, and Academic Development Certificate)
- Specialized teaching certification in reading (e.g., Reading Certificate or Reading and Language Arts Specialist Certificate)
- Experience writing or reviewing test items for standardized tests, especially tests for ELs in K through grade twelve
- Recent experience administering the CELDT

Selections were made to ensure representation from different cultural and linguistic groups, various-sized local educational agencies (LEAs) and county offices of education, and different geographical regions of the state, and with regard to the travel budget allowable in the contract. ETS and SCOE made preliminary selections, which were reviewed by the CDE, adjusted as needed, and then approved. Forty-two educators were selected for item writer training, along with 14 alternates. Forty-two educators were selected for Content Review Panels, along with 14 alternates. Ten educators were selected for Bias and Sensitivity Review Panels, along with three alternates.

SCOE contacted and invited the participants and contacted the alternates as necessary. Once all participants confirmed, SCOE notified those who were not selected.

### **3.2.4. Item Writing by Educators**

Item writer training was divided into two sets of meetings, each of which lasted four days.

A total of 42 educators were trained to develop items during the first set of item writer training meetings in 2016. Twenty-four educators from K through grade five were trained on Monday and Tuesday, February 22 and 23, 2016. Eighteen educators from grades six through twelve were trained on Wednesday and Thursday, February 24 and 25, 2016.

A total of 20 educators were trained to develop items during the second set of item writer training meetings in 2017. Twelve educators from K through grade five were trained on Monday and Tuesday, November 6 and 7, 2017. Eight educators from grades six through twelve were trained on Wednesday and Thursday, November 8 and 9, 2017.

The educators at each set of meetings represented a mix of rural, suburban, and urban LEAs across California.

#### **3.2.4.1. Introduction to Item Writing**

During each of the two-day meetings, educators received training and then drafted ELPAC items. At the start of day one, a PowerPoint presentation was used to provide information to the educators about topics regarding the ELPAC and item development. Topics covered during the presentation included an overview of the ELPAC, general principles of item development, a review of the 2012 ELD Standards, the overall item development process, and the process for drafting and submitting items. After the PowerPoint presentation, ETS trainers provided educators with examples of task types that are shared across grade levels and grade spans.

ETS trainers facilitated brainstorming sessions, during which educators listed topics that served as a basis for item development. Educators were asked to propose topics for item content that are covered during prior grades to ensure that topics were appropriate. After brainstorming, educators worked as a whole group to assign topics to appropriate grade levels or grade spans. Educators then split up into grade-level groups to draft items corresponding to the topics from their brainstorming session. This pattern was followed for all domains (Listening, Speaking, Reading, and Writing).

#### **3.2.4.2. Process**

After educators divided into their grade-level groups, ETS trainers provided them with *Item Writing Guidelines for the ELPAC* (CDE, 2016b), sample items, and item templates. The *Item Writing Guidelines for the ELPAC* provided details about the type of information that is required when drafting items, such as the length of any Listening stimuli or Reading passages, the number of items within the set, and the types of English language knowledge, skills, and abilities to be assessed by the items.

The sample items were developed by ETS assessment specialists to serve as examples of the task types to be developed. The item templates were Word files that contained areas for entering information. The item templates assured that items were drafted in a standardized manner and that all needed item information was entered. ETS trainers used the *Item Writing Guidelines for the ELPAC*, sample items, and item templates as training materials to provide clear expectations regarding the information needed when drafting each task type, as well as the level of quality that was expected.

#### **3.2.4.3. Assignment**

After the first set of item writer training meetings in 2016, ETS trainers gave educators item writing assignments to be completed during the two-day training. ETS trainers remained within the training rooms when educators were drafting items to answer questions and to provide feedback regarding initial drafts of items.

Educators were also given the opportunity to take an item writing assignment to be completed in the weeks after the two trainings. They were provided with the printed training materials needed to complete the assignment and given two weeks to complete their assignments. Educators were required to return all secure printed training materials at the time their assignments were submitted. Secure printed training materials were returned via secure express delivery.



To submit assignments, educators saved their assignments in password-protected files and copied them to a secure ETS server. After ETS confirmed receipt of the files, educators were prompted to delete the files from their personal devices.

In 2017, all items developed by educators were drafted according to assignments that were given during the item writer training meetings. Educators were not given additional assignments to be completed after the meetings.

### 3.2.5. Item Writing by Contractors

In both 2016 and 2017, ETS assessment specialists worked with five contractors (i.e., outside item writers) who are fully trained, experienced item writers with a record of developing quality items for other ETS English language assessments. Because there was a limited amount of time to train California educators to develop Listening and Reading sets, ETS contractors developed the Listening task types with relatively long stimuli and the Reading task types with relatively long passages. The focus of the contractors was to develop the following task types:

- *Listening—Listen to a Story*
- *Listening—Listen to an Oral Presentation*
- *Reading—Read a Literary Passage*
- *Reading—Read an Informational Passage*

The contractors delivered all items to a secure ETS server. After ETS confirmed receipt of the files, contractors were prompted to delete the files from their personal devices.

### 3.2.6. Item Review Panels

Before ELPAC items were designated as field test ready, the draft versions underwent a thorough ETS internal review process, including two content reviews, a fairness review, and an editorial review; external reviews by item review panels; and a CDE review and final approval. This section describes the reviews conducted by two sets of item review panels: the Content Review Panel and the Bias and Sensitivity Review Panel.

To help establish content validity for the ELPAC and to develop test materials that are fair to all students, the set of approximately 2,000 ELPAC test items was reviewed by a Content Review Panel and a Bias and Sensitivity Review Panel from August 1 through August 5, 2016, and the set of about 200 items was reviewed from February 19 through February 22, 2018. Content Review Panel reviewed items to ensure that items were aligned with the 2012 ELD Standards (CDE, 2014), items were appropriate for the grade level or grade span, items addressed the construct being tested, and selected-response items had one and only one correct answer. Bias and Sensitivity Review Panel reviewed items to ensure that they did not contain content that would result in bias to identified groups or be considered potentially offensive.

#### 3.2.6.1. Meeting Plan and Training

The CDE and ETS agreed to hold the first set of panel meetings on an overlapping schedule within a single week. This approach allowed ELPAC items to be developed on time for stand-alone sample field testing in 2016–2017 while ensuring that appropriate procedures were followed to produce a high-quality pool of items.

Two trainings for the panel participants were conducted during the meetings and prior to the item reviews: educators serving on the Content Review Panel were trained on Monday, August 1, 2016, and educators serving on the Bias and Sensitivity Review Panel were

trained on Wednesday, August 3, 2016. The Content Review Panel meeting began on August 1, 2016, and finished on August 5, 2016. The Bias and Sensitivity Review Panel meeting began on August 3, 2016, and finished on August 5, 2016.

Since the first set of panel meetings in 2016 allowed for thorough and efficient reviews of items, the second set of panel meetings in 2018 was also held in quick succession. Fewer items to review in 2018 than in 2016 meant that the panel meetings in 2018 were held on a back-to-back schedule within a single week, instead of in an overlapping schedule. The Content Review Panel meeting began on February 19, 2018, and finished on February 20, 2018. The Bias and Sensitivity Review Panel meeting began on February 21, 2018, and finished on February 22, 2018.

### 3.2.6.2. Process

The Bias and Sensitivity Review Panel members reviewed items as revised by the Content Review Panel. Members of the Bias and Sensitivity Review Panel needed to read and understand the comments of the Content Review Panel before providing comments on bias and sensitivity issues. Facilitators were responsible for transferring comments from the Content Review Panel to the Bias and Sensitivity Review Panel during designated times. Notetakers projected the Content Review Panel comments on a screen to allow members of the Bias and Sensitivity Review Panel to read them.

Facilitators monitored the progress of the panel reviews to ensure all items were reviewed by the last day of the panel meetings. As planned, the Content Review Panels finished their reviews and delivered their comments to the Bias and Sensitivity Review Panels. This allowed the Bias and Sensitivity Review Panel members to conduct their reviews while considering the final Content Review Panel comments. Bias and Sensitivity Review Panel members then completed their reviews of the items by the end of their sessions.

### 3.2.6.3. Outcome

Educators at both the Content Review Panel meeting and the Bias and Sensitivity Review Panel meeting had the option of making one of three decisions regarding each stimulus and item:

1. Approve as is
2. Approve with revisions
3. Reject

[Table 3.2](#) provides the status of the stimuli and items after the 2016 item review panel meetings.

**Table 3.2 Status of Stimuli and Items After the 2016 Item Review Panel Meetings**

Grade Level or Grade Span	Approved As Is	Approved with Revisions	Rejected
Kindergarten	202	91	0
Grade 1	213	87	0
Grade 2	240	90	5
Grade span 3–5	284	90	1
Grade span 6–8	216	132	0
Grade span 9–10	253	118	2
Grade span 11–12	270	88	12
<b>Totals:</b>	<b>1,678</b>	<b>696</b>	<b>20</b>

[Table 3.3](#) provides the status of the stimuli and items after the 2018 item review panel meetings.

**Table 3.3 Status of Stimuli and Items After the 2018 Item Review Panel Meetings**

<b>Grade Level or Grade Span</b>	<b>Approved As Is</b>	<b>Approved with Revisions</b>	<b>Rejected</b>
Kindergarten	27	15	0
Grade 1	14	13	0
Grade 2	17	16	0
Grade span 3–5	22	25	0
Grade span 6–8	16	21	0
Grade span 9–10	21	24	0
Grade span 11–12	23	14	0
<b>Totals:</b>	<b>140</b>	<b>128</b>	<b>0</b>

After each set of item review panel meetings, the CDE reviewed the proposed revisions to items, made any adjustments needed, and then approved the revisions. Both the 2016 and 2018 item writing efforts yielded high percentages of approved items. In 2016, 99 percent of the 2,394 items were approved. In 2018, 100 percent of the 268 items were approved. During both item development efforts, educators enhanced the quality of the item pool by providing suggestions for revising items during Content Review Panel meetings and Bias and Sensitivity Review Panel meetings. All revisions that were approved by the CDE were implemented before the items were field tested.

### 3.3. Item Banking

The ETS Item Banking Information System (IBIS) was used as the database of record throughout the item-development process. IBIS was used to store item text, graphics, scripts for audio recordings, scoring information, and metadata. After ETS assessment development staff drafted and reviewed items in IBIS, the CDE used IBIS to review items in preparation for item review panels. After the CDE approved proposed revisions from the item review panel meetings, CDE staff confirmed the items in IBIS to ensure that revisions were implemented correctly before the items were approved for field testing.

## References

- California Department of Education. (2014). *California English Language Development Standards: Kindergarten through grade 12*. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndpublication14.pdf>
- California Department of Education. (2015). Proposed test blueprints for the English Language Proficiency Assessments for California. In *California State Board of Education November 2015 agenda; Subject English Language Proficiency Assessments for California: Approve the test blueprints for the English Language Proficiency Assessments for California* (pp. 23–43). Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/be/ag/ag/yr15/documents/nov15item12.doc>
- California Department of Education. (2016b). *Item writing guidelines for the English Language Proficiency Assessments for California*. [Unpublished report]. Sacramento, CA: California Department of Education.
- California Department of Education. (2016a). *Report on the First Pilot of Items for the ELPAC*. [Unpublished report]. Sacramento, CA: California Department of Education.

# Chapter 4: Test Development

---

## 4.1. Test Design

This chapter describes the development of the 2018–2019 Summative English Language Proficiency Assessments for California (ELPAC) forms, including the revisions to the Summative ELPAC test blueprints based on the field test results, the rules for item selection, the structure of the test forms, and the development of the test materials. Each form of the Summative ELPAC assesses the four domains of Listening, Speaking, Reading, and Writing. All items included on the 2018–2019 Summative ELPAC were administered first in a stand-alone field test. Refer to the *Summative ELPAC Technical Report, 2017–18 Administration* (California Department of Education [CDE], 2019) for more details about the fall 2017 stand-alone field test.

### 4.1.1. Revision of the Test Blueprints

All items included on the 2018–2019 Summative ELPAC were administered in a stand-alone field test. After the administration of the stand-alone field test, items went through statistical analysis and the *Proposed Test Blueprints for the ELPAC* (CDE, 2015) were revised. Based on the statistical performance of the items in the stand-alone field test, Educational Testing Service (ETS) adjusted the number of items in the test blueprint.

The State Board of Education (SBE) had adopted the *Proposed Test Blueprints for the ELPAC* on November 4, 2015, which was prior to the first pilot test of items. Revisions to the test blueprints from the first pilot of items and the stand-alone field test were compiled and then presented to the SBE for review. The SBE approved and adopted the updated *Summative Assessment Test Blueprints for the ELPAC* on September 14, 2017 (CDE, 2017b).

The next two subsections provide an overview of the analyses performed in making the decisions for the final blueprint.

#### 4.1.1.1. Statistical Analysis

The Summative ELPAC stand-alone field test was held from March 6 through April 14, 2017. After the administration, all items from the stand-alone field test underwent statistical item analysis. The ETS Psychometric Analysis & Research group used student responses to compile item statistics and flagged any items that fell outside of acceptable parameters. Assessment specialists reviewed each flagged item and made one of three recommendations:

1. Keep the flagged item as is and classify it as operationally ready
2. Revise the flagged item and classify it as field test ready for a future form
3. Reject the flagged item and discontinue using it

After the field test items went through statistical item analysis, ETS delivered the item analysis results to the CDE.

#### 4.1.1.2. Analysis of Results

As part of the test design process, an evidence-centered design (ECD) approach was used. ECD is a principled framework that “ensures that the way in which evidence is gathered and interpreted bears on the underlying knowledge and purposes the assessment is intended to address” (Mislevy, Steinberg, & Almond, 1999, p. 1). Through this approach, the performance of the Summative ELPAC task types was reviewed. Those task types that

were most appropriate for use in the upcoming operational assessment were retained at each grade.

The overall number of Summative ELPAC task types remained at 27, but adjustments were made to the number of task types and items at each grade level or grade span. This was particularly the case in the Writing domain at kindergarten (K), grade one, and grade two. The Summative ELPAC test blueprints for Writing were adjusted to include fewer Writing items than were included in the stand-alone field test. The reasons for this were as follows:

- Avoid a Writing domain that is overly burdensome on students and test examiners
- Ensure that the Writing domain elicits appropriate evidence of students' skills in relation to the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards), reflecting information learned about each task type from the field test
- Ensure that the Writing domain contributes appropriately to valid and reliable score reporting

## 4.2. Item Selection

The development of the Summative ELPAC necessitated fulfilling the requirements of the test blueprints as well as meeting the statistical and psychometric criteria specified, as described in this section.

### 4.2.1. Test Development Specifications

The development of the 2018–2019 Summative ELPAC began with the creation of test development specifications. ETS created the test development specifications that the CDE reviewed and approved after revision. The test development specifications for the 2018–2019 Summative ELPAC described the goals of the assessment, the content criteria for selecting items, the psychometric criteria for selecting items, the test development process, and a timeline for major activities.

Two types of operational items were used to develop the 2018–2019 Summative ELPAC:

1. Items that were field-tested in the 2016–2017 stand-alone summative field test and were then used as operational items in the 2017–2018 Summative ELPAC
2. Items that were field-tested in the 2016–2017 stand-alone summative field test but were not used in the 2017–2018 Summative ELPAC

One operational form was created for each of the seven grade levels and grade spans. Approximately 70 percent of the operational items were reused from the 2017–2018 Summative ELPAC and approximately 30 percent of the items were operationally ready items from the 2016–2017 stand-alone summative field test. Each form assessed all four domains of Listening, Speaking, Reading, and Writing.

The test development specifications for the 2018–2019 Summative ELPAC included plans to develop five field test forms at each grade and grade span to administer embedded field test items. Development of the field test forms began as scheduled in March 2018 and proceeded according to schedule until August 2018, when the CDE directed ETS to discontinue work on the embedded field test forms. The work was discontinued because new plans to administer the 2019–2020 Summative ELPAC in a computer-based format removed the need to field test items in a paper-based format.

### 4.2.2. Content Criteria

Test validity requires that content coverage adheres to test blueprints. The blueprints specify the number of items from each task type to include in each domain and which 2012 ELD Standards are assessed in each domain. ETS assessment specialists used the *Summative Assessment Test Blueprints for the ELPAC* (CDE, 2017a) as the basis to select task types and items for the 2018–2019 Summative ELPAC. Assessment specialists selected items that covered a variety of content areas and topics to ensure that balanced forms were created.

ETS assessment specialists used the Item Banking Information System (IBIS) to develop form planners for the 2018–2019 Summative ELPAC. A form planner is an Excel spreadsheet that contains information about each of the items included in a test form. The form planners include information such as the item’s accession number (i.e., the unique item identification code), grade, domain, correct answer (for multiple-choice items), score scale (for constructed-response items), and alignment to the 2012 ELD Standards. After form planners were created, ETS reviewed them internally. An ETS assessment specialist who did not participate in test assembly performed a full review of each test form to ensure that an appropriate set of items was selected. After this review was completed, the form planners were delivered to ETS psychometricians for review.

### 4.2.3. Statistical and Psychometric Criteria

The statistical specifications provided guidelines for selecting items and developing tests with appropriate psychometric properties. Statistics from the 2016–2017 stand-alone field test were used to inform the development of the 2018–2019 Summative ELPAC.

Each ELPAC test form conformed with the following psychometric criteria:

- Individual items had  $p$ -values—a measure of item difficulty—that ranged from 0.20 to 0.95.
- The collection of items within each domain represented an overall difficulty level with average  $p$ -values from 0.5 to 0.7.
- Point-biserial correlations—a measure of reliability—for each item was greater than 0.15.
- Differential item functioning (DIF) analyses were conducted to detect possible test bias and locate items for which one group of students performed significantly better than another group of students of similar ability.

ETS assessment specialists assembled the 2018–2019 Summative ELPAC test forms based on the classical statistics obtained from the 2016–2017 stand-alone field test. ETS psychometricians then reviewed the composition of the test forms and compiled distribution tables, which showed the distribution of items according to difficulty, to ensure that correct numbers and distributions of items were selected. Having a broad distribution of item difficulties ensured that there was reasonable measurement power across the range of difficulty.

After ETS psychometricians reviewed the composition of the 2018–2019 Summative ELPAC test forms, ETS assessment specialists revised the composition of the test forms based on psychometric review, as needed.

#### 4.2.4. CDE Review of Item Selection

After revisions were made to the form planners during internal ETS reviews, the form planners and distribution tables were delivered to the CDE for review. CDE staff had access to item content and metadata via IBIS, through which they reviewed the item content, form planners, and distribution tables.

The CDE made recommendations for replacing items within the test forms. ETS adjusted the form planners as needed and then submitted the revised form planners to the CDE for review and approval.

### 4.3. Forms Development

This section describes the development of the paper-based test materials, including the production of audio recordings for Listening and Speaking items and the development of the large-print and braille versions as well as the breach edition.

#### 4.3.1. Developing Paper-based Test Materials

This subsection describes the development of audio recordings and paper-based test materials for the 2018–2019 Summative ELPAC.

##### 4.3.1.1. Audio Recordings

ETS worked with professional recording studios and voice actors to develop all audio recordings used at grades three through twelve for Listening items and *Speaking—Summarize an Academic Presentation* items. All audio recordings for Listening operational items were developed prior to the field test administration according to both the quality standards established during the development of a demonstration reel and the CDE confirmation of the field test recordings.

The item-level audio recordings from the field test administrations were used to develop the 2018–2019 Summative ELPAC Listening test forms for grades three through twelve. Based on feedback from the field test, professional audio recordings of the grades three through twelve *Speaking—Summarize an Academic Presentation* items were also developed according to the same standards that were established during the development of the demonstration reel.

After the form planners for each test form of the 2018–2019 Summative ELPAC were approved, test-length audio files were developed for the Listening domain at grades three through twelve. To prepare for the development of the test (domain)-level audio files, item-level scripts were compiled to create test-length scripts, including section directions, task type directions, practice items, and operational items. The test-length scripts and item-level audio recordings were delivered to a professional studio for compilation.

After the studio compiled the item-length recordings into test-length recordings, ETS proofed the test-length audio recordings to ensure they were compiled accurately. The grades three through twelve *Speaking—Summarize an Academic Presentation* audio files were compiled separately because they were the only audio files for the Speaking domain.

##### 4.3.1.2. Paper-based Test Materials

After the form planner for each 2018–2019 Summative ELPAC test form was approved, ETS assessment specialists delivered the form planners and item content to the ETS production team. ETS production staff used the instructions provided by the assessment specialists to compile the item content and create the paper-based test materials. The collaboration between the two teams resulted in the development of all paper-based test



materials, including seven *Examiner's Manuals* (all grade levels and grade spans), seven Test Books (all grade levels and grade spans), and four Answer Books (grade spans three through five, six through eight, nine and ten, and eleven and twelve).

After the ETS production teams composed the paper-based test materials, the materials were subject to internal ETS reviews before they were delivered to the CDE for review and approval.

## 4.3.2. Developing Special Version Forms

### 4.3.2.1. Braille

The goal of the ELPAC braille versions of the forms is to provide valid and reliable measurement of English language proficiency (ELP) for students who use braille by including scoring tables with the same performance level threshold scores as the standard version of the assessment. The same braille forms that were developed for the 2017–2018 Summative ELPAC were used in the administration of the 2018–2019 Summative ELPAC. This subsection describes the development of the braille forms used during both the 2017–2018 Summative ELPAC and the 2018–2019 Summative ELPAC administrations.

ETS assessment specialists collaborated with members of the ETS Accessibility and Alternate Formats (AAF) team to develop the braille forms.

#### 4.3.2.1.1. Criteria

A foundational step in developing the braille forms was to review the ELPAC task types at a high level and determine which task types were amenable to braille, which needed to be revised to become amenable to braille, and which were not amenable to braille. Each ELPAC task type was analyzed for suitability for administration in a braille form. Solutions were proposed at the task-type level.

In developing the proposed solutions, the ETS team endeavored to minimize changes needed to task types to make them accessible to English learners (ELs) with visual impairment. Any necessary adaptations were designed to preserve the target construct and measure the same ELP standards and targeted performance level descriptors.

ETS staff analyzed the Summative ELPAC task types and documented the process to be used to develop braille and large-print versions in the *Process for Development of Special Test Versions* (CDE, 2017b). The CDE reviewed and approved the document before ETS began development of the braille versions.

To begin, ETS reviewed individual items to ensure that item content was sensitive to the experiences of ELs with visual impairment. Reviews of individual items were also needed to confirm that the cognitive load of the braille item remained comparable with the original item. Once items were selected for the 2017–2018 Summative ELPAC, ETS staff reached agreement on the adaptations needed for the braille forms. For those items needing adaptations, variants of the items were created in IBIS, adapted, and then reviewed in IBIS to confirm accuracy. The same items that were used in the braille version of the 2017–2018 Summative ELPAC were administered in the braille version of the 2018–2019 Summative ELPAC.

#### 4.3.2.1.2. Process

After items were adapted for the braille forms, ETS provided the braille vendor with the information needed to produce the braille forms. Before the braille forms were produced, ETS communicated with the braille vendor to confirm the exact specifications for the final

deliverable according to the *Rules for Unified English Braille* (2013) set forth by the Round Table on Information Access for People with Print Disabilities Inc. and International Council on English Braille, a collective that includes the Braille Authority of North America. The braille vendor was responsible for ensuring the quality of the braille forms. The quality-control measures included two proofs of all test materials.

In addition to reviewing the ELPAC task types, instructions for test examiners, such as test administration and domain-specific procedures, were reviewed and adapted as needed for the braille administration. Test developers, in consultation with the AAF team, reviewed and adapted the directions by grade level and grade span in parallel with item-level evaluation. Final directions were provided to the CDE for approval prior to certification.

#### **4.3.2.2. Large-Print**

The goal of the ELPAC large print versions of forms is to provide valid and reliable measurement of ELP for students who use large-print materials. This subsection describes the development of the large-print forms used during the 2018–2019 Summative ELPAC administration.

ETS assessment specialists collaborated with members of the ETS AAF team to develop the large-print forms.

##### **4.3.2.2.1. Criteria**

Form 1 of the 2018–2019 Summative ELPAC was selected for use as the large-print ELPAC. ETS assessment specialists worked with the AAF group to agree upon the content to be enlarged.

Marked-up content was delivered to page composers, who created the large-print versions. The assessment specialists and AAF team then reviewed the large-print forms, requesting any revisions needed before the materials were delivered to the large-print vendor. The vendor produced proofs, which ETS then reviewed. The vendor made any requested adjustments before the large-print forms were submitted to the CDE for review.

##### **4.3.2.2.2. Process**

All student-facing test content was enlarged to develop the large-print forms. Most of the student-facing content was found in the Test Books and Answer Books, although some was in a grade-level *Examiner's Manual*. Any student-facing content that was in an *Examiner's Manual* was enlarged and placed in the Answer Book at K through grade two and in the Test Book at grades three through twelve.

ETS enlarged all student-facing text to 18- to 20-point font in grades two through twelve. In K and grade one, where the font size is already 18 points or larger, the font size was increased by four points, (e.g., 18-point text was increased to 22 points in the large-print forms).

##### **4.3.2.3. Breach Edition**

After the 2017–2018 Summative ELPAC was developed, a breach form was developed for each grade level and grade span for use in a breach edition. The breach edition was developed as a print-ready edition that could be printed and distributed in the case of large-scale, public exposure of an operational ELPAC form. Items that were approved as operationally ready during the item analysis of the 2016–2017 stand-alone field test were used to populate the forms in the breach edition. To use the ELPAC item pool efficiently, approximately 30 percent of the items were shared between the 2017–2018 Summative ELPAC and the breach edition.

Because there were no security breaches during the 2017–2018 Summative ELPAC administration, the same breach form was kept available during the administration of the 2018–2019 Summative ELPAC. During the 2018–2019 Summative ELPAC administration, there were no security breaches and the breach form was not used.

#### **4.3.3. CDE Review of Assembled Forms**

After revisions were made to the form planners during internal ETS reviews, the form planners and distribution tables were delivered to the CDE for review. Using the metadata available in IBIS, the CDE reviewed the item content, form planners, and distribution tables. The CDE made recommendations for replacing items within the test forms. ETS adjusted the form planners as needed and then submitted the revised form planners to the CDE for review and approval.

## References

- California Department of Education. (2015). Proposed test blueprints for the English Language Proficiency Assessments for California. In *California State Board of Education November 2015 agenda; Subject English Language Proficiency Assessments for California: Approve the test blueprints for the English Language Proficiency Assessments for California* (pp. 23–43). Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/be/ag/ag/yr15/documents/nov15item12.doc>
- California Department of Education. (2017b). *Process for development of special test versions*. [Unpublished report]. Sacramento, CA: California Department of Education.
- California Department of Education. (2017a). Summative assessment test blueprints for the English Language Proficiency Assessments for California. In *California State Board of Education September 2017 agenda; Subject English Language Proficiency Assessments for California: Approve the Revised Test Blueprints, the Revised General Performance Level Descriptors, and the Reporting Hierarchy* (pp. 9–21). Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/be/ag/ag/yr17/documents/sep17item18.doc>
- California Department of Education. (2019). *Summative English Language Proficiency Assessments for California technical report, 2017–18 administration*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/ta/tg/ep/documents/elpac18techrpt.pdf>
- Mislevy, R. J., Steinberg, L. S., & Almond, R. (1999). *Evidence-centered assessment design*. Princeton, NJ: Educational Testing Service.
- Round Table on Information Access for People with Print Disabilities Inc. and International Council on English Braille. (2013). *The rules of Unified English Braille second edition 2013*. Retrieved from [http://www.iceb.org/Rules%20of%20Unified%20English%20Braille%202013%20\(linked\).pdf](http://www.iceb.org/Rules%20of%20Unified%20English%20Braille%202013%20(linked).pdf)

## Chapter 5: Test Administration

---

### 5.1. Procedures to Maintain Standardization

To maintain standardization during the 2018–2019 English Language Proficiency Assessments for California (ELPAC) administration, ELPAC staff at local educational agencies (LEAs) were provided with several forms of communication and training. Educational Testing Service (ETS) produced and provided the *Summative ELPAC Test Administration Manual*, which detailed the process and policies for a secure and standardized administration, as well as other quick-reference guides describing various aspects of ELPAC administration. Additionally, the Sacramento County Office of Education (SCOE) provided several trainings across the state to site ELPAC coordinators and ELPAC test administrators. These trainings provided a hands-on opportunity for participants to learn about and ask questions regarding ELPAC administration. SCOE also provided training for test examiners who administered the Speaking and Listening sections of the ELPAC.

#### 5.1.1. LEA ELPAC Coordinator

An LEA ELPAC coordinator was designated by the district superintendent at the beginning of the 2018–2019 school year. LEAs include public school districts, statewide benefit charter schools, State Board of Education–authorized charter schools, county office of education (COE) programs, and direct funded charter schools.

LEA ELPAC coordinators were responsible for ensuring the proper and consistent administration of the ELPAC. In addition to the responsibilities set forth in *California Code of Regulations (CCR)*, Title 5, Section 11518.40, their responsibilities included

- adding site ELPAC coordinators and ELPAC test examiners to the Test Operations Management System (TOMS);
- attending, or assigning staff to attend, an annual 2018–2019 California Department of Education (CDE)–sponsored Summative ELPAC Administration and Scoring training;
- ensuring that the site ELPAC coordinators and test examiners in their LEA were appropriately trained regarding the administration of the ELPAC, including security policies and procedures;
- ensuring that all site ELPAC coordinators submitted signed *ELPAC Test Security Affidavits* and *ELPAC Test Security Agreements*;
- entering and verifying the correct shipping address for materials and reporting address for score reports in TOMS;
- reporting to the California Technical Assistance Center all test security irregularities and breaches that occurred before, during, or after test administration within 24 hours of discovery;
- ensuring that correct testing procedures were followed;
- ensuring that test materials were distributed to the schools and kept in a locked, secure area at all times;
- ordering test materials, pre-identification (Pre-ID) labels, and supplemental test materials in TOMS;

- ensuring adequate test materials were on hand and redistributed throughout the LEA during the testing window as needed;
- shipping all materials back for scoring;
- securely destroying secure, nonscannable materials locally or shipping them back to ETS for destruction; and
- distributing Student Score Reports (SSRs) to test sites electronically.

The 2018–2019 LEA ELPAC coordinator was required to sign the *ELPAC Test Security Agreement* (5 CCR 11518.50[b]).

### **5.1.2. Site ELPAC Coordinator**

A site ELPAC coordinator was trained by the LEA ELPAC coordinator for each test site (5 CCR Section 11518.40[b][7]). The 2018–2019 site ELPAC coordinator was required to sign both the *ELPAC Test Security Agreement* and the *ELPAC Test Security Affidavit* (5 CCR Section 11518.45[b][3]).

In addition to the responsibilities set forth in 5 CCR Section 11518.45, their responsibilities included

- identifying test examiners, proctors, and any other persons with access, as appropriate, and ensuring that they had submitted signed *ELPAC Test Security Affidavits*, as appropriate;
- retaining for up to 12 months the signed *ELPAC Test Security Affidavits* from test examiners and proctors;
- adding test examiners into TOMS;
- ensuring that all test examiners and proctors had been trained and certified to administer the 2018–2019 Summative ELPAC;
- assuming general oversight responsibilities for all administration activities in their school and for all test examiners and other school staff;
- viewing student information in TOMS prior to testing to ensure that the students' English language acquisition status (ELAS) is EL in the California Longitudinal Pupil Achievement Data System because only these students are eligible for the Summative ELPAC;
- coordinating with the test examiners so that all domains of the 2018–2019 Summative ELPAC are administered to each student;
- ensuring the proper administration of all testing procedures;
- mitigating and reporting all test security incidents to the LEA ELPAC coordinator in a manner consistent with ELPAC policies;
- maintaining the security of all test materials at the site; and
- assuring the proper packing and return of test materials to the LEA ELPAC coordinator.

### 5.1.3. ELPAC Test Examiner

Test examiners were identified by ELPAC site coordinators as individuals who administered the Summative ELPAC and were an employee or contractor of an LEA. A test examiner was proficient in English with complete command of pronunciation, intonation, and fluency, and had certified that training in the administration and scoring of the ELPAC had been completed. Proctors assisted test examiners during group administration.

Prior to handling testing materials, a test examiner and any other individual handling 2018–2019 Summative ELPAC testing materials was required to sign a *Test Security Affidavit* (5 CCR Section 11518.50[d]), which was provided at the Administration and Scoring Training workshop and also available on the ELPAC Forms web page at <https://www.elpac.org/test-administration/forms/>.

A test examiner's duties may have included

- reading and signing the *ELPAC Test Security Affidavit* and then returning it to the site ELPAC coordinator;
- completing annual 2018–2019 Summative ELPAC training and reviewing all ELPAC policy and administration documents prior to administering any tests;
- ensuring the physical conditions of the testing room met the criteria for a secure test environment;
- viewing student information in their local student information system prior to testing to ensure that the students' ELAS was EL;
- administering one or more domains of the 2018–2019 Summative ELPAC;
- reporting all test security incidents to the site ELPAC coordinator and LEA ELPAC coordinator in a manner consistent with ELPAC, state, and LEA policies;
- fully complying with all directions provided in the *Examiner's Manual*; and
- returning all test materials to the site ELPAC coordinator after testing.

### 5.1.4. Instructions for Test Administration

#### 5.1.4.1. Examiner's Manuals

These were grade-level or grade-span manuals that described the standardized testing procedures used by test examiners to administer the 2018–2019 Summative ELPAC to students. Test examiners were required to follow the procedures in the manuals so that all students were given an equal opportunity to demonstrate their English language proficiency.

The *Examiner's Manuals* provided directions and guidelines for filling in student demographic information on each student Answer Book prior to the test if the LEA did not use the Pre-ID service.

During the test, test examiners read, word-for-word, the directions and scripts for administration. Test examiners also used the Speaking rubrics and anchor samples in the *Examiner's Manual* to evaluate students' responses and assign scores. At grades three through twelve, where recorded audio is played for the Listening domain and for *Speaking—Summarize an Academic Presentation*, the manuals described procedures for playing the recorded audio.

#### **5.1.4.2. Summative ELPAC Test Administration Manual**

The *Summative ELPAC Test Administration Manual* (CDE, 2019) contained information and instructions on overall procedures and guidelines for all LEA and test site staff involved in the administration of the ELPAC. Sections included the following topics:

- Dates for ordering materials and testing
- Roles and responsibilities of those involved with ELPAC testing
- Test administration resources
- Test security
- Administration preparation and planning
- General test administration
- Instructions for steps to take before, during, and after testing
- Guidelines for handling materials

#### **5.1.4.3. TOMS Guide for the ELPAC**

TOMS is a web-based application that allowed LEA ELPAC coordinators to add and manage users and order materials for the Summative ELPAC. In 2018–2019, test examiners used TOMS to play the audio recordings used during the Listening and Speaking portions of the ELPAC in grades three through twelve.

TOMS modules used for Summative ELPAC administration that are described in the *TOMS Guide for the ELPAC* included the following (CDE, 2018a):

- **Adding and Managing Users**—This module allowed LEA ELPAC coordinators to add ELPAC test site coordinators and test administrators to TOMS so that the designated user could administer, monitor, and manage the ELPAC
- **Ordering Test Materials**—This module allowed LEA ELPAC coordinators to approve orders, view summary orders, view and track orders, and place supplemental orders within specified windows
- **Ordering Pre-ID Labels**—This module allowed LEA ELPAC coordinators to request Pre-ID labels that were affixed to Answer Books and used to track student testing and assign results within specified windows
- **Playing Audio Modules**—This module allowed test examiners access to the audio files that were part of the Listening and Speaking portions of the ELPAC in grades three through twelve

## **5.2. Training**

SCOE provided several trainings across the state to site ELPAC coordinators and ELPAC test administrators. These trainings provided a hands-on opportunity for participants to learn about and ask questions regarding ELPAC administration. SCOE also provided training for test examiners who administered the Speaking and Listening sections of the ELPAC.

### **5.2.1. General Test Administration**

The online Moodle training site was developed as a restricted site that could be accessed only by LEA trainers and others requiring general training in the administration of the ELPAC. (Moodle is a free, learning-management, open-source software.) The site contained all resources needed to conduct a training, such as training presentations along with the presenters' scripts.



## 5.2.2. Scoring Training of Trainers Workshops

All LEAs in California were required to send a trainer to the all-day, CDE-sponsored, statewide 2018–2019 Summative ELPAC Administration and Scoring Training (AST), which employed the “training-of-trainers” model.

### 5.2.2.1. Goals

The goals of the 2018–2019 Summative ELPAC AST were to do the following:

1. Standardize the administration of the ELPAC at all domains (i.e., Listening, Speaking, Reading, and Writing)
2. Train test examiners to score the Speaking items accurately and reliably
3. Train LEA trainers to train other qualified persons locally to administer and score the ELPAC

The training covered the test administration of all grade levels and grade spans as well as all domains. However, most of the training day was spent on the administration and scoring of the Speaking domain. Extensive training was provided because Speaking scores were given “in the moment” by test examiners, so the standardization of the scoring is critical. Refer to subsection [7.4 Constructed-Response Scoring for Speaking](#) for details about this aspect of the training.

### 5.2.2.2. Locations

The Summative ELPAC AST trainings were held at 24 locations throughout California from October 2018 through December 2018. All participants completing the Summative ELPAC AST were sent, via email, certificates of completion. A total of 2,493 educators attended, representing a total of 1,451 LEAs (refer to [table 5.1](#)).

**Table 5.1 2018 AST Training**

2018 Date	Location	Attended
October 9	Sacramento	145
October 11	Redding	79
October 16	Santa Barbara	90
October 17	Burbank	153
October 18	Montebello	145
October 19	San Diego	116
October 23	Monterey	52
October 24	San Jose	120
October 25	Redwood City	64
October 30	Santa Rosa	112
October 31	Santa Ana	133
November 1	Palm Springs	68
November 2	Riverside	106
November 6	Stockton	93
November 7	Merced	56
November 8	Madera	81
November 9	Visalia	141
November 13	Burbank	163

Table 5.1 (continuation)

2018 Date	Location	Attended
November 14	Santa Ana	63
November 15	Torrance	79
November 16	Pomona	93
November 27	Concord	112
November 29	Sacramento	117
<b>Total:</b>	<b>N/A</b>	<b>2,493</b>

An additional 205 LEAs were trained at COE-sponsored regional trainings. There were 28 regional trainings held by 16 COEs throughout the state. SCOE sold training materials on a cost-recovery basis to these county offices for their regional trainings to standardize all trainings.

Three hundred and fifty-three LEAs had no participation data available, indicating they did not attend one of the scheduled training sessions. The list of LEAs that did not attend training was provided to the CDE.

### 5.2.2.3. Availability of Materials

The online Moodle training site was developed as a restricted site that could be accessed only by LEA trainers and test examiners. The site contained all resources needed to conduct an LEA test examiner training session, such as downloadable training manuals, training presentations, training videos, scoring rubrics, as well as training and calibration quizzes for Speaking scoring. LEA trainers downloaded materials to prepare for their training sessions and shared access to the site with the test examiners within the LEA. Test examiners used the site to review training materials and to calibrate in preparation for Speaking scoring.

## 5.2.3. Scoring Rubrics

Scoring rubrics provide guidance to the raters who evaluate student responses. The *Speaking Rubrics for the ELPAC* (CDE, 2017) and the *Writing Rubrics for the ELPAC* (CDE, 2018b) are essential components in the design of the ELPAC Speaking and Writing items.

### 5.2.3.1. Creation

Draft rubrics for scoring responses to the ELPAC Speaking and Writing items were designed in tandem with the design of task types for the ELPAC. The draft rubrics were designed to be used to score responses to several task types. As part of the 2014–2015 pilot of ELPAC task types, the draft rubrics were used to evaluate student responses. After being modified as a result of further study after their first use, the revised rubrics were used to support the 2015–2016 ELPAC item writing effort in which the item pool for the stand-alone field tests was developed.

During the item writing effort, the rubrics were further refined. The most significant change was that the rubrics were revised to be specific to each task type. This change was made based on the judgment that the use of task-specific rubrics, rather than generic rubrics, would increase the ease of internalization and usability by raters and help support efficient and reliable scoring.

### 5.2.3.2. Range Finding and Approval

After the item pool for the Summative ELPAC field test was developed, the *Speaking Rubrics for the ELPAC* (CDE, 2017) and the *Writing Rubrics for the ELPAC* (CDE, 2018b) were reviewed during meetings held in Sacramento in 2016. CDE, ETS, and SCOE staff practiced scoring student responses from the pilot to evaluate the usability of the rubrics. After revisions were applied, the rubrics were approved for use during Speaking range finding meetings held in October 2016 and Writing range finding meetings held in May 2017. The approved rubrics were used to score student responses to the Summative ELPAC field test and the 2018–2019 Summative ELPAC operational assessment.

The purposes of the Speaking range finding meetings and Writing range finding meetings were to select sample responses that were used to train raters and to calibrate them prior to scoring. During the range finding meetings, educators reviewed the rubrics and refined them. CDE staff reviewed and approved the revisions to the rubrics and selected samples that aligned with the rubrics while the range finding meetings were in session.

#### 5.2.3.2.1. Composition of ELPAC Speaking Range Finding and Writing Range Finding Meetings and Participant Qualifications

California educators participated in range finding meetings. Participant groups consisted of current and former teachers, resource specialists, administrators, curricular experts, and other education professionals. Minimum qualifications to be invited to participate were

- three or more years of teaching experience in kindergarten through grade twelve,
- expertise in language acquisition or experience teaching English learners (ELs) in kindergarten through grade twelve
- knowledge of and experience working with the 2012 ELD Standards,
- bachelor's or higher degree, and
- knowledge of and experience with the California content standards in language arts.

Preferred qualifications included

- a teaching credential authorization for ELD, specially designed academic instruction in English, or content instruction delivered in the primary language (e.g., Cross-cultural, Language, and Academic Development Certificate; Bilingual, Cross-cultural, Language, and Academic Development Certificate),
- specialized teaching certification in reading (e.g., Reading Certificate, Reading and Language Arts Specialist Certificate), and
- experience writing or reviewing test items for standardized tests, especially tests for K–12 ELs.

Every effort was made to ensure that groups of range finding participants included a representation of genders and of the geographic regions and ethnic groups in California.

[Table 5.2](#) shows the educational qualifications, present occupation, and credentials of the individuals who participated in ELPAC Speaking range finding and Writing range finding meetings.

**Table 5.2 ELPAC Speaking Range Finding (SpRF) and Writing Range Finding (WrRF) Meetings Qualifications, by Meeting Type and Total**

Qualification Type	Qualification	SpRF	WrRF	Total
Occupation	Classroom teacher	4	0	4
Occupation	English learner or literacy specialist or coach	9	0	9
Occupation	EL instructional specialist	0	0	0
Occupation	School administrator	0	0	0
Occupation	LEA or county office employee	0	0	0
Highest degree earned	Bachelor's Degree	4	0	4
Highest degree earned	Master's Degree	9	0	9
Highest degree earned	Doctorate	0	0	0
K–12 teaching credential	Elementary Teaching (multiple subjects)	7	0	7
K–12 teaching credential	Secondary Teaching (single subject)	1	0	1
K–12 teaching credential	Language Development Specialist	1	0	1
K–12 teaching credential	English Learner (CLAD, BCLAD)	12	0	12
K–12 teaching credential	Other	0	0	0

**Note:** Numbers may not match the totals because participants may have multiple occupations or teaching credentials, or are currently working toward earning their highest degree. The information is self-reported and may not reflect all of the experience and earned credentials.

### 5.3. Testing Students with Disabilities

The ELPAC provided a number of accessibility resources to enable all students to participate in the 2018–2019 Summative ELPAC administration. ETS produced large-print Test Books as well as braille Test Books in contracted and uncontracted braille. The CDE's Matrix Four provided LEAs with guidance on available accessibility resources and accommodations (CDE, 2018c).<sup>3</sup> Because the 2018–2019 Summative ELPAC was a paper-pencil assessment, embedded resources typically available on a computer-based assessment, such as text-to-speech, closed captioning, American Sign Language videos, and embossed braille, were not available.

<sup>3</sup> This technical report is based on the version of Matrix Four that was available during the 2018–2019 Summative ELPAC administration.

### 5.3.1. Locally Determined Alternate Assessments

Individualized education program (IEP) teams may have determined that a student with the most significant cognitive disabilities was unable to participate in one or more domains of the ELPAC, even with accommodations, due to short- or long-term disabilities. In this instance, the student may have been tested with a locally determined alternative assessment per the student's IEP.

A statewide Summative Alternate ELPAC, for students with the most significant cognitive disabilities, was not developed at the time of this test administration.

## 5.4. Test Security and Confidentiality

For the 2018–2019 Summative ELPAC administration, every person who worked with the assessments, communicated test results, or received testing information was responsible for maintaining the security and confidentiality of the tests, including CDE staff, ETS staff, ETS subcontractors, LEA ELPAC coordinators, site ELPAC coordinators, and test examiners.

ETS' Code of Ethics requires that all test information, including tangible materials (e.g., test items and test books), confidential files (e.g., those containing personally identifiable student information), processes related to test administration (e.g., the packing and delivery of test materials), and activities are kept secure. ETS has systems in place that maintain tight security for test items, test books, and test results, as well as for student data. To ensure security for all the tests that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI), which is described in the next subsection.

All tests within the ELPAC system, as well as the confidentiality of student information, should be protected to ensure the validity, reliability, and fairness of the results. As stated in *Standard 7.9* of the *Standards for Educational and Psychological Testing*, "The documentation should explain the steps necessary to protect test materials and to prevent inappropriate exchange of information during the test administration session" (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).

### 5.4.1. ETS' Office of Testing Integrity

The OTI is a division of ETS that provides quality-assurance services for all testing programs managed by ETS; this division resides in the ETS legal department. The Office of Professional Standards Compliance at ETS publishes and maintains *ETS Standards for Quality and Fairness* (ETS, 2014), which supports the OTI's goals and activities. The *ETS Standards for Quality and Fairness* provides guidelines to help ETS staff design, develop, and deliver technically sound, fair, and beneficial products and services and to help the public and auditors evaluate those products and services.

The OTI's mission is to

- minimize any testing security violations that can impact the fairness of testing,
- minimize and investigate any security breach that threatens the validity of the interpretation of test scores, and
- report on security activities.

The OTI helps prevent misconduct on the part of students and administrators, detects potential misconduct through empirically established indicators, and resolves situations

involving misconduct in a fair and balanced way that reflects the laws and professional standards governing the integrity of testing. In its pursuit of enforcing secure practices, the OTI strives to safeguard the various processes involved in a test development and administration cycle.

#### **5.4.2. Test Delivery**

Because the 2018–2019 Summative ELPAC was a paper-pencil assessment, there were logistics involved to ensure the timely delivery of test materials to LEAs across the state. To manage the materials ordering process, ETS used TOMS, a secure website that permitted ELPAC users to perform a number of tasks for the ELPAC program. Through TOMS, users could perform the following activities:

- Confirm or update an LEA shipping address, add a score report shipment address, and indicate whether an LEA can receive pallet shipments
- Order test materials in the primary test materials order window, including braille and large-print forms, in either Round 1 or Round 2; and order additional test materials as needed, in the Supplemental window
- Add site ELPAC coordinators and test examiners
- Order Pre-ID labels
- Administer the Listening domain and the *Speaking—Summarize Academic Presentations* item for grades three through twelve

The ETS warehouse team prepared shipments based on orders submitted by each LEA. Materials were tracked using closed-loop tracking and United Parcel Service tracking methods to ensure timely delivery of ELPAC test materials. Shipping notices were included in each delivery. These notices provided LEAs with an inventory of the number of Test Books, Answer Books, and other materials included in the shipment. Additionally, LEAs were provided with return materials that included Group Identification Sheets, which were precoded, scannable forms facilitating identification of materials when they were received at ETS; and shipping labels that allowed tracking of materials that were returned to ETS for scoring.

#### **5.4.3. Security of Electronic Files Using a Firewall**

A firewall is software that prevents unauthorized entry to files, email, and other organization-specific information. All ETS data exchanges and internal email remain within the ETS firewall at all ETS locations, ranging from Princeton, New Jersey, to San Antonio, Texas, to Concord and Sacramento, California.

All electronic applications that are included in TOMS remain protected by the ETS firewall software at all times. Due to the sensitive nature of the student information processed by TOMS, the firewall plays a significant role in maintaining assurance of confidentiality among the users of this information.

#### **5.4.4. Transfer of Scores via Secure Data Exchange**

Due to the confidential nature of test results, ETS currently uses secure file transfer protocol (SFTP) and encryption for all data file transfers; test data is never sent via email. SFTP is a method for the reliable and exclusive routing of files. Files reside on a password-protected server that only authorized users can access. ETS shares an SFTP server with the CDE. On that site, ETS posts Microsoft Word and Excel files, Adobe Acrobat PDFs, or other

document files for the CDE to review; the CDE returns reviewed materials in the same manner. Files are deleted upon retrieval.

The SFTP server is used as a conduit for the transfer of files; secure test data is stored only temporarily on the shared SFTP server. Industry-standard secure protocols are used to transfer test content and student data from the ETS internal data center to any external systems. For the 2018–2019 Summative ELPAC, ETS entered information about the deliverable in a web form on a SharePoint website when a file was posted. A CDE staff member checked this log throughout the day for updates on the status of deliverables and downloads and deleted the file from the SFTP server when its status showed it has been posted.

#### **5.4.5. Data Management**

ETS currently maintains a secure database to house all student demographic data and assessment results. Information associated with each student has a database relationship to the LEA, school, and grade codes as the data is collected during operational testing. Only individuals with the appropriate credentials can access the data. ETS builds all interfaces with the most stringent security considerations, including interfaces with data encryption for databases that store test items and student data. ETS applies best and up-to-date security practices, including system-to-system authentication and authorization, in all solution designs.

All stored test content and student data is encrypted. ETS complies with the Family Educational Rights and Privacy Act (20 *United States Code [USC]* § 1232g; 34 *Code of Federal Regulations* Part 99) and the Children’s Online Privacy Protection Act (15 *USC* §§ 6501–6506, P.L. No. 105–277, 112 Stat. 2681–1728).

In TOMS, staff at LEAs and test sites were given different levels of access appropriate to the role assigned to them.

#### **5.4.6. Statistical Analysis on Secure Servers**

The 2018–2019 Summative ELPAC results were scanned or entered by human raters. After scoring constructed-response items, the Information Technology team at ETS loaded data files from the SFTP site and then loaded them into a database. The ETS Data Quality Services staff extracted the data from the database and performed quality-control procedures before passing files to the ETS Psychometric Analysis & Research (PAR) group. The PAR group kept the files on secure servers. All staff members involved with the data adhered to the ETS Code of Ethics and the ETS Information Protection Policies to prevent any unauthorized access to data.

#### **5.4.7. Student Confidentiality**

To meet the requirements of the Every Student Succeeds Act as well as state requirements, LEAs must collect demographic data about students’ ethnicity, disabilities, parent/guardian education, and so forth. ETS took every precaution to prevent any of this information from becoming public or being used for anything other than testing purposes. These procedures were applied to all documents in which student demographic data appeared, including reports and the Pre-ID files and response booklets used in paper-pencil testing.

## 5.4.8. Student Test Results

### 5.4.8.1. Types of Results

Electronic SSRs were produced for each student administered the Summative ELPAC and were made available to the LEAs every two weeks. Paper SSRs were produced for the students of the LEAs that requested paper SSRs. When requested, LEAs received two paper copies of the SSR: one in English and the other in the SSR-supported language requested by the LEA.

Additionally, ETS produced aggregate data files containing ELPAC test result data for schools.

### 5.4.8.2. Security of Results Files

ETS took measures to protect files and reports that showed students' scores and ELPAC levels. ETS is committed to safeguarding all secure information in its possession from unauthorized access, disclosure, modification, or destruction. ETS has strict information security policies in place to protect the confidentiality of both student and client data. ETS staff access to production databases was limited to personnel with a business need to access the data. User IDs for production systems were person-specific or for systems use only.

ETS has implemented network controls for routers, gateways, switches, firewalls, network tier management, and network connectivity. Routers, gateways, and switches represent points of access between networks. However, these do not contain mass storage or represent points of vulnerability, particularly for unauthorized access or denial of service.

ETS has many facilities, policies, and procedures to protect computer files. Software and procedures such as firewalls, intrusion detection, and virus control are in place to provide for physical security, data security, and disaster recovery. ETS is certified in the BS 25999-2 standard for business continuity and conducts disaster recovery exercises annually. ETS routinely backs up all data to either disks through deduplication or to tapes, all of which are stored off site.

Access to the ETS Computer Processing Center is controlled by employee and visitor identification badges. The Center is secured by doors that can only be unlocked by the badges of personnel who have functional responsibilities within its secure perimeter. Authorized personnel accompany visitors to the ETS Computer Processing Center at all times. Extensive smoke detection and alarm systems, as well as a preaction fire-control system, are installed in the Center.

### 5.4.8.3. Security of Individual Results

ETS protects individual students' results on both electronic files and paper reports during the following events:

- Scoring
- Transfer of scores by means of secure data exchange
- Reporting
- Analysis and reporting of erasure marks
- Posting of aggregate data
- Storage

In addition to protecting the confidentiality of testing materials, ETS' Code of Ethics further prohibits ETS employees from financial misuse, conflicts of interest, and unauthorized



appropriation of ETS property and resources. Specific rules are also given to ETS employees and their immediate families who may be administered a test developed by ETS (e.g., an ELPAC). The ETS OTI verifies that these standards are followed throughout ETS. This verification is conducted, in part, by periodic on-site security audits of departments, with follow-up reports containing recommendations for improvement.

#### 5.4.9. Security and Test Administration Incident Reporting Process

The LEA ELPAC coordinator was responsible for reporting all testing incidents and security breaches immediately.

If an irregularity or security breach occurred at the school, the test examiner was required to report the incident to the site ELPAC coordinator, who would then report the incident to the LEA ELPAC coordinator. Testing irregularities relate to incidents that occur during the administration of the ELPAC that were likely to impact the reliability and validity of test score interpretations.

**Testing irregularities** included *but were not limited to*:

- Cheating by students
- Failing to follow test administration directions
- Rushing students through the test or parts of the test
- Coaching students, including, *but not limited to*, the following:
  - Discussing questions with students before, during, or after testing
  - Giving or providing any clues to the answers
- Administering the wrong grade level or grade span test to a student or using mismatched test materials
- Writing on the scannable Answer Book by a test examiner that would cause the Answer Book to be unscorable and therefore need transcription to a new Answer Book
- Leaving instructional materials on walls in the testing room that may assist students in answering test questions
- Allowing students to have additional materials or tools (e.g., books, tables) that are **not** specified in an IEP, Section 504 plan, or approved by the CDE as an allowed testing accommodation

**Security breaches** included, *but were not limited to* the following:

- Site ELPAC coordinators, test examiners, proctors, or students using electronic devices such as cell phones during testing
- Posting pictures of test materials on social media sites
- Missing test materials
- Copying or taking a photo of any part of the test materials
- Permitting eligible students access to test materials outside of the testing periods
- Failing to maintain security of all test materials

- Sharing test items or other secure materials with anyone who has not signed the *Test Security Affidavit*
- Discussing test content or using test materials outside of training and administration
- Allowing students to take the test out of the designated testing area
- Allowing test examiners to take test materials home
- Allowing untrained personnel to administer the test

If an incident occurred, the LEA ELPAC coordinator was instructed to notify ETS within 24 hours of the incident. Additionally, the coordinator was required to complete the *ELPAC Testing Irregularities and Security Breach Report Form*. The CDE and ETS collaborated on defining next steps and providing the LEA ELPAC coordinator with instructions on how to mitigate the incident.

## References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- California Department of Education. (2017). *Speaking rubrics for the English Language Proficiency Assessments for California*. [Unpublished manuscript]. Sacramento, CA: California Department of Education.
- California Department of Education. (2018c). *Matrix Four: Universal tools, designated supports, and accommodations for the English Language Proficiency Assessments for California*. Retrieved from <https://www.cde.ca.gov/ta/tg/ca/accessibilityresources.asp>
- California Department of Education. (2018a). *Test Operations Management System Guide for the ELPAC*. Sacramento, CA: California Department of Education. Retrieved from <https://elpac.org/s/pdf/ELPAC.TOMS-users-guide.2018-19.pdf>
- California Department of Education. (2018b). *Writing rubrics for the English Language Proficiency Assessments for California*. Sacramento, CA: California Department of Education.
- California Department of Education. (2019). *Summative ELPAC Test Administration Manual*. Sacramento, CA: California Department of Education. Retrieved from <https://elpac.org/s/pdf/ELPAC.Test-Admin-Manual-Summative.2018-19.pdf>
- Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>

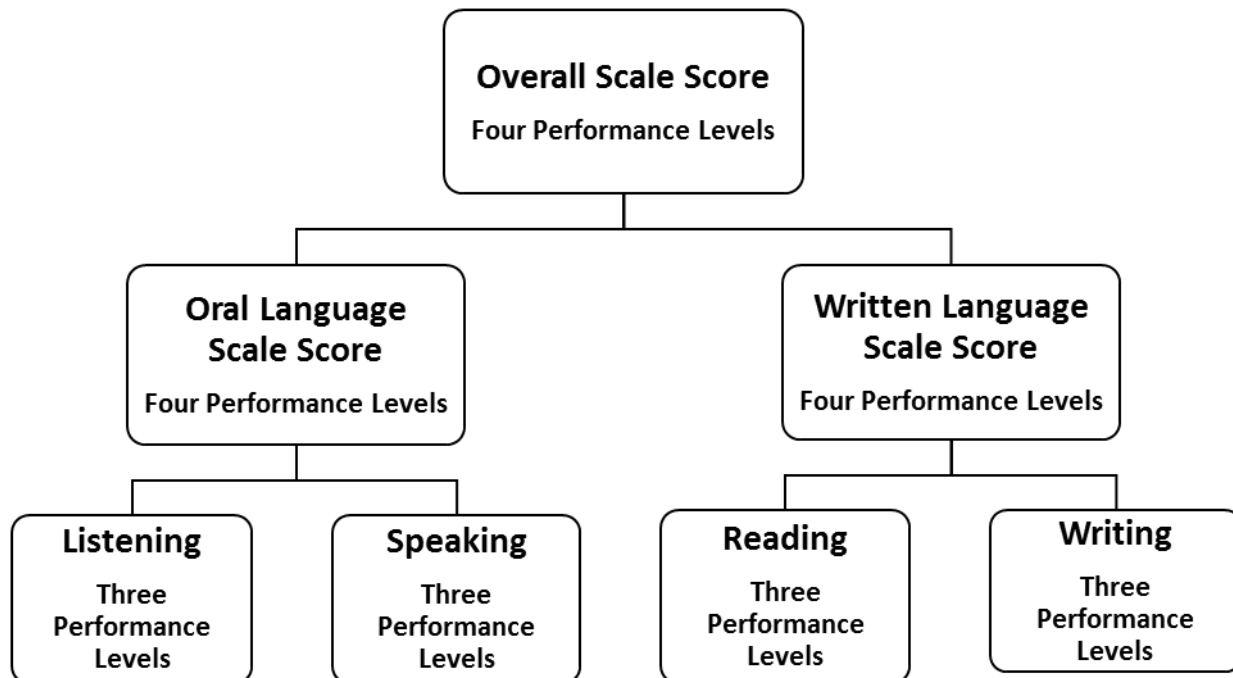
## Chapter 6: Standard Setting

This chapter summarizes the standard setting process through which Summative English Language Proficiency Assessments for California (ELPAC) performance levels were established. Included are a background of the development of ELPAC, an overview of the standard setting methodology, a summary of the standard setting procedures, the description of the performance level descriptors (PLDs), and the results. The detailed standard setting information for the Summative ELPAC is described in the *Standard-Setting Technical Report for the Summative ELPAC* (California Department of Education [CDE], 2018).

### 6.1. Background

Implementation of the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014) and the administration of the new Summative ELPAC required a standard setting process to evaluate students' English language proficiency against the new expectations.

[Figure 6.1](#) presents the score reporting hierarchy for the Summative ELPAC, approved in September 2017 by the California State Board of Education (SBE). As depicted in this figure, four performance levels must be reported for three composite scores: scale scores and performance levels for the overall, oral language, and written language scores. The oral language scale score branches off into the Listening and Speaking domains, which each have three performance levels. The written language scale score branches off into the Reading and Writing domains, which each have three performance levels.



**Figure 6.1 Summative ELPAC Score Reporting Hierarchy**

To develop threshold score recommendations aligned with the score reporting hierarchy, Educational Testing Service (ETS) conducted standard setting workshops in Sacramento,

California, for the seven Summative ELPAC grade levels and grade spans on October 17 through October 20, 2017 (kindergarten [K], grade one, and grade two), and October 23 through October 26, 2017 (grade spans three through five, six through eight, nine and ten, and eleven and twelve). Standard setting for K through grade two was conducted in week one and for grades three through twelve in week two. All four domains and the overall score were considered in the standard setting process.

## 6.2. Performance Level Descriptors

The Summative ELPAC general (policy) PLDs describe short policy descriptors that convey the expectation at each performance level, across all grades tested (CDE, 2016). They were provided to the panelists for prereading prior to the standard setting workshop.

After the general PLDs were available, a team of educators familiar with both students taking the Summative ELPAC and the 2012 ELD Standards reviewed the general PLDs for the ELPAC target population. They developed more detailed grade- and content-specific PLDs for the range of expectations at each performance level (range PLDs). Panelists referenced the SBE-approved general PLDs and the range PLDs as part of the standard setting process.

## 6.3. Standard Setting Methodologies

Standard setting refers to a class of methodologies by which one or more performance threshold scores are used to determine performance levels. The purpose of the standard setting process for the Summative ELPAC was to collect recommendations from California educators for the placement of the threshold scores for review by the CDE, with final determination and approval by the SBE.

ETS conducted standard setting workshops in fall 2017, following the field test administration of the Summative ELPAC. The overall approach used for setting standards for the Summative ELPAC aligned with the 2012 ELD Standards, which reflect the interdependence of the language domains.

By design, the Summative ELPAC and the standard setting methodologies explicitly support a treatment of skills such as Speaking and Listening in combination, rather than as isolated skills. Educators working on standard setting panels considered the assessment by domain, articulated skills that are expected in Listening, Speaking, Reading, and Writing, and made final threshold score recommendations by considering the interdependence of these skills.

The Bookmark method (Lewis, et al., 1998; Mitzel, et al., 2001) was applied to the Reading and Listening domains; a Performance Profile approach was applied to the Writing and Speaking domains (Baron & Papageorgiou, 2014; Tannenbaum & Cho, 2014; Tannenbaum & Baron, 2010; Wan, Bay, & Morgan, 2017). In the final round, panelists were instructed to think holistically across the four domains and consider consequence data when they made the overall threshold score recommendations.

### 6.3.1. Bookmark Method (Reading and Listening Domains)

The Summative ELPAC standard setting process employed the Bookmark method for the seven grade levels and grade spans (K, grade one, grade two, and grade spans three through five, six through eight, nine and ten, and eleven and twelve) for the Reading and

Listening domains, which consisted of dichotomously scored multiple-choice items.<sup>4</sup> This portion of the workshop resulted in recommendations for threshold scores for these two domains.

The Bookmark method has its basis in item response theory (IRT) analysis. IRT is used to estimate item difficulties. These estimates are used to order the items from easiest to hardest in a booklet known as an ordered item booklet (OIB) and to place item difficulty estimates on the score scale. One benefit of this approach is that, once panelists make judgments in the OIB, the difficulty values associated with each item have a built-in relationship to scale scores through theta (the ability parameter in IRT), which allows results to be provided to score users and policy makers on the familiar metric of scale score. Panelists completed two rounds of Bookmark judgments for Reading and Listening for their assigned grade level or grade span. Then, the panelists began work on the Speaking and Writing domains for the same assigned grade level or grade span.

### **6.3.2. Performance Profile Method (Speaking and Writing Domains)**

The Summative ELPAC standard setting process employed the Performance Profile method for the Speaking and Writing domains, which consisted of constructed-response items. This portion of the workshop resulted in recommendations for threshold scores for these two domains.

The Performance Profile method is a holistic method that requires panelists to make decisions or judgments based on an examinee's score profiles, or overall performance, rather than on each separate test item or task. This method has been used in standard setting studies for English learner (EL) assessments and other types of K–12 statewide assessments throughout the United States (e.g., Baron & Papageorgiou, 2014; ETS, 2014).

In this approach, panelists reviewed actual samples of student responses across multiple tasks, such as Speaking video samples of student performance on the Speaking tasks, and multiple Writing responses. Item scores for a student's set of responses to the items form a profile; panelists considered the performance at each total score represented by the profiles of responses across tasks. Writing profiles were sampled from field test responses, and speaking profiles were sampled from scorer-training videos developed by the Sacramento County Office of Education in June 2017. Profiles were selected to represent the full range of scores and the most frequently occurring score patterns.

In each of two rounds of judgments, all panelists independently selected total scores associated with score profiles and marked the score representing the expected knowledge and skills at the threshold of each performance level, using the definitions of borderline students. The instruction to the panelists was to base decisions about which total score aligns best with the definition of the borderline student on the full set of evidence provided across all test items in Speaking (the same process was followed for Writing).

Panelists recorded their Round 1 recommended Speaking or Writing total score for each threshold score. After Round 1, each panelist's individual threshold score recommendations were shared with the panel and discussed; panel judgments were summarized and discussed prior to the next round of judgments.

---

<sup>4</sup> Grade two included two items that were multipoint items. These items appeared twice in the ordered item booklet, according to the RP67 theta value associated with each score point.

## 6.4. Standard Setting Procedures

This subsection describes what occurred prior to, and during, the standard setting workshop.

### 6.4.1. Panelists

Prior to the standard setting, panelists were recruited to include a diverse, representative group of California educators with both experience in the education of students who will take the ELPAC and familiarity with the 2012 ELD Standards. An additional goal was to recruit subject-area teachers working with these students in grades six and above, because these teachers provide a perspective on content-specific learning goals for the students taking the ELPAC. Educators were selected using the following criteria:

- Educators who were working with ELs in the grade level(s) assigned to the panel
- English-language specialists
- Educators teaching any or all of the subject areas of mathematics, science, and social studies

The final decision on the panelists selected for the workshops was made by the CDE.

For the Summative ELPAC, there were six panels of educators. Three panels—kindergarten and grades one and two—met during the first week of the workshop. Three panels—grade spans three through five, six through eight, and nine through twelve—met in the second week. There were 71 panelists; the number in each panel ranged from 8 to 11.

Because standard setting is based on expert judgment—informed by performance data—it was important that panelists collectively reflected 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 ELPAC educator population in particular. Panels included a sample across genders, ethnic and racial backgrounds, and geographical regions in California. A majority of the educators indicated they had more than five years of experience working with ELs. Most panels included educators with experience teaching mathematics, science, social studies, and English.

### 6.4.2. Materials

All panelists, regardless of the standard setting methodology, were provided the following materials prior to the workshop:

- A letter describing the purpose and procedures of the standard setting workshop
- A preworkshop assignment specific to their panel assignment
- A notetaking form for the assignment
- A link to the SBE-approved general PLDs
- The domain and grade- and grade-span-specific PLDs for the tests the panelists would be reviewing

At the standard setting workshop, panelists received training materials, a set of operational materials, and evaluation forms. The definitions of borderline students were developed by the panelists themselves during the workshop.

The operational materials panelists received at the workshop included the following:

- The OIB (for the Bookmark method)
- An item map
- Judgment recording forms
- Performance samples for Speaking and Writing
- Rubrics used operationally to score student responses

At the end of the training for each method, and at the end of the workshop, panelists completed evaluation forms. Evaluations included questions about training, understanding the tasks, the influence of different aspects of the standard setting process, and panelists' beliefs about the final recommended threshold scores.

#### **6.4.2.1. Descriptions of Materials**

##### ***6.4.2.1.1. Materials for the Bookmark Method***

The OIB is a booklet of all items included in the standard setting judgments, ordered by difficulty on the basis of student performance. For each item, the page of the OIB shows the item (along with any short passage or graphic), the possible responses, and the correct answer. For the items that are associated with a passage, a separate passage booklet was included with the OIB for panelists to reference for items associated with a passage.

The item map is a summary document displaying relevant information regarding each item. It shows the ordered item number, the original item number in the test, the correct answer, a difficulty value, and the passage title and score-level scale. The item map provided was ordered by difficulty in the same manner as the OIB.

##### ***6.4.2.1.2. Materials for the Performance Profile Method***

Performance Profile samples are complete student responses to the Speaking and Writing tests. For Speaking, video files of students responding to all tasks were displayed; each student score was known to the panelists, allowing them to visualize a sample of students across the range of performance. All student videos played for each panel showed students taking the same items.

For Writing, copies of students' Writing responses to the full set of Writing tasks were provided in a booklet of Writing samples. The Writing sample book included the prompt and written response for a range of Writing domain scores. More than one score sample was displayed when available for both Speaking and Writing.

Panelists also used the scoring rubrics for Speaking and Writing in their discussions and in their individual judgments.

#### **6.4.3. Process**

The workshop process included a general session, where all panelists were provided an overview of the purpose of the meeting, their role and the roles of facilitators, and an explanation of the two approaches used in the standard setting for the ELPAC. Educators were then guided to grade-span-specific panel rooms, where they completed the training and judgment process (Baron & Papageorgiou, 2016; Morgan, 2004).



### 6.4.3.1. Training

Training was provided on the following topics:

- Test familiarization
- Development of Borderline Student Definitions
- Standard setting judgment process for both bookmark and performance profile
  - Training and practice prior to the first round of judgments
- Review of ordered items and practice in method of bookmark placement
- Review of Speaking videos for performance profile judgments
- Feedback and discussion and Round 2 judgments for each domain
- Round 3 integrated holistic judgments on the overall score

### 6.4.3.2. Judgments and Feedback

The Reading and Writing section scores were combined into the written composite score, and the Speaking and Listening section scores were combined into the oral composite score. The feedback to the panelists after Round 2 judgments were complete included each of the four domain score recommendations as well as the recommended threshold scores for the two composites. Panelists received training on how the domain scores were combined and how to consider the data provided for the domains and composites in the Round 3 integrated judgments.

Panelists made recommendations for three threshold scores on the Summative ELPAC overall score and were instructed to consider all of the information provided and then make a recommendation for the overall score performance level expectations.

## 6.5. Standard Setting Results

Results from the ELPAC standard setting after Round 3 included a recommended threshold score for each composite (oral and written) and the overall composite for each test (kindergarten, grades one and two, and grade spans three through five, six through eight, nine and ten, and eleven and twelve). The *Standard-Setting Technical Report for the Summative ELPAC* (CDE, 2018) presents details about the following results from the standard setting workshops:

- The median threshold score recommendations for each domain at the end of each round
- Standard errors of judgment, scale scores, and conditional standard errors of measurement in the Bookmark metric for Reading and Listening
- Standard errors of judgment, scale scores, and conditional standard errors of measurement in the Performance Profile metric for Speaking and Writing

[Table 6.1](#) through [table 6.7](#) show the projected percentage of students statewide who would be placed at this performance level on the basis of the results of the 2016–2017 field test administration of the Summative ELPAC. The threshold scale score is the minimum standard setting scale score needed to achieve a performance level.

Scales provided in these tables were presented and used in the standard setting process and are more user-friendly than scores in the theta metric. However, it should be noted that the scores presented are not the ELPAC-reported scale scores. The scale was created, based on the 2016–2017 field test data, for standard setting prior to the approval of the official scale for the Summative ELPAC and was used as a tool for the standard setting process.

**Table 6.1 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Kindergarten**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	10.6
Level 2	338	20.4
Level 3	380	35.8
Level 4	428	33.2

**Table 6.2 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade One**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	9.1
Level 2	381	14.0
Level 3	411	21.3
Level 4	441	55.5

**Table 6.3 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Two**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	5.1
Level 2	389	10.2
Level 3	424	33.6
Level 4	475	51.1

**Table 6.4 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Three Through Five**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	8.1
Level 2	441	21.4
Level 3	490	52.8
Level 4	569	17.7

**Table 6.5 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Six Through Eight**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	6.5
Level 2	451	24.6
Level 3	516	40.7
Level 4	577	28.3

**Table 6.6 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Nine and Ten**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	16.7
Level 2	484	25.6
Level 3	544	31.9
Level 4	607	25.9

**Table 6.7 Projected Distribution of 2016–2017 Students Based on Round 3 Recommendations: Grade Span Eleven and Twelve**

<b>Performance Level</b>	<b>Threshold Scale Score</b>	<b>Percentage</b>
Level 1	N/A	13.4
Level 2	486	24.3
Level 3	547	36.4
Level 4	618	25.8

Results presented in the *Standard-Setting Technical Report for the Summative ELPAC* are based on the standard setting workshop and panel-recommended threshold scores at the end of the workshop. Following the standard setting workshop, the SBE reviewed both the panel recommendations and the State Superintendent of Public Instruction’s recommendations for threshold scores.

## References

- Baron, P. A., & Papageorgiou, S. (2014). *Mapping the TOEFL® Primary™ test onto the Common European Framework of Reference* (Research Memorandum RM-14-05). Princeton, NJ: Educational Testing Service.
- Baron, P. and Papageorgiou, S. (2016). *Setting language proficiency score requirements for English-as-a-second-language placement decisions in secondary education*. (Research Report 16-17). Princeton, NJ: Educational Testing Service. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1124774.pdf>
- California Department of Education. (2014). *2012 California English language development standards: Kindergarten through grade 12*. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndpublication14.pdf>
- California Department of Education. (2016). *Summative ELPAC general PLDs*. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/ta/tg/ep/elpacgpld.asp>
- California Department of Education. (2018). *Standard-setting technical report for the Summative English Language Proficiency Assessments for California*. [Unpublished report]. Sacramento, CA: California Department of Education.
- Educational Testing Service. (2014). *Standard-setting technical report; Proficiency Assessments for Wyoming Students [PAWS] and Student Assessment of Writing Skills [SAWS]*. [Unpublished report]. Princeton, NJ: Educational Testing Service.
- Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (1998). *The Bookmark standard setting procedure: Methodology and recent implementations*. Paper presented at the 1998 annual meeting of the National Council on Measurement, San Diego, CA.
- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.). *Setting Performance Standards: Concepts, Methods, and Perspectives* (pp. 249–81). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Morgan, D.L. (2004, June). *The performance profile method (PPM): A unique standard setting method as applied to a unique population*. Presented at the annual meeting of the Council of Chief State School Officers in Boston, MA.
- Tannenbaum, R. J., & Baron, P. A. (2010). *Mapping TOEIC® test scores to the STANAG 6001 language proficiency levels* (Research Memorandum RM-10-11). Princeton, NJ: Educational Testing Service.
- Tannenbaum, R. J., & Cho, Y. (2014). Critical factors to consider in evaluating standard-setting studies to map language test scores to frameworks of language proficiency. *Language Assessment Quarterly*, 11, 233–249.
- Wan, L., Bay, L., & Morgan, D. (2017) *Validity evidence for the Performance Profile standard setting method*. Presentation at the National Council for Measurement in Education Annual Conference, San Antonio, TX.

## Chapter 7: Scoring and Reporting

---

### 7.1. Procedures for Maintaining and Retrieving Individual Scores

The local educational agency (LEA) English Language Proficiency Assessments for California (ELPAC) coordinator was responsible for returning all materials to Educational Testing Service (ETS) for scoring. When materials were received at the ETS warehouse, several quality checks were implemented. These included verifying there was no damage to the Answer Books prior to scanning as well as capturing issues such as double marks and inconsistencies between the pre-identification label and the marked information.

Once received, all Answer Books were scanned and writing items were routed to trained raters at ETS for scoring. Once student responses were scored, a Student Score Report (SSR) was produced.

### 7.2. Multiple-Choice Scoring

After the certification of student records for scoring, ETS transferred the records to the scoring management system. These records contained all relevant response data and identifying information for matching against the correct scoring keys. The ETS scoring engine then processed the records and produced the multiple-choice (MC) raw scores before permanently storing the results in the students' records.

### 7.3. Constructed-Response Scoring for Writing

Prior to operational use, for all ELPAC Writing items, a range of professionals that included California educators carefully developed and vetted the rubrics, benchmark sample responses, and rater training materials over the course of field testing and additional reviews.

#### 7.3.1. Scorer Training

It is critical for the success of the ELPAC constructed-response (CR) scoring to have well-defined scorer recruitment, training, and certification processes with staff in place to control scoring quality.

##### 7.3.1.1. Procedures

The procedures ETS used in training ELPAC scorers included the following:

- **Rigorous Training for the Scoring Leaders**—ETS developed training materials and helped select benchmark and training samples during range finding for the purpose of training scoring leaders and scorers. ETS hired scoring leaders with experience and familiarity in scoring similar programs at ETS. Scoring leaders were given materials to study independently.
- **Extensive Training of Scorers**—Scorers were trained to properly apply the appropriate rubric for scoring each task type, following generic sample responses that exemplified the quality required for each score point. This ensured that every prompt was scored using the same general criteria. The ETS Online Network for Evaluation (ONE) scoring system supported scorer training with a full-service menu of options, including orientation materials, program-specific information, and training on how to use the platform, as well as interactive training that included practice scoring for both potential and qualified scorers.

There were two types of training sets offered within ONE:

- **Feedback Sets** provided users with feedback after each response. Users could also access the overall results at the end of the set.
- **Practice Sets** mimicked the actual scoring, and users do not have access to any score results until they have completed the set.

ETS provided role-based training modules for using ONE. For example, ETS expected scoring leaders to review and study more modules (on topics such as monitoring) than scorers. Scoring leaders and scorers were required to review PDFs as well as training videos that covered the critical functions required for their individual scoring roles.

#### **7.3.1.2. Certification**

Certification occurred after training and was intended to determine how well scorers could adopt and apply the scoring standards. Scorers' ratings on certification responses were compared to predetermined correct scores to ascertain whether scorers successfully applied the scoring standards reflected in the rubrics (scoring guides).

As part of the initial qualification for scoring ELPAC prompts, every scorer had to successfully complete training and pass a certification test consisting of a set of prescored responses. If a scorer was unsuccessful on the first certification attempt, that scorer was retrained prior to making a second attempt. If a scorer was unsuccessful at the second attempt, that scorer was not added to the potential ELPAC scorer pool. This process increased the likelihood of securing a highly proficient scorer pool.

#### **7.3.1.3. Regular Calibration**

Calibration is a short test of reader accuracy that occurs regularly at the beginning of a scoring session to determine whether scorers are ready to begin scoring the assessment. Calibration is a proven method to mitigate scoring drift and promote the quality of scoring over time.

Before calibration, scorers were directed to review relevant training materials (rubrics and benchmarks). During calibration, scorers assigned scores to a prescored set of responses to determine their ability to accurately apply scores for a particular task type.

As with the certification process described previously, scorers had two opportunities for correct calibration. If a scorer was unsuccessful on his or her first calibration attempt, that scorer conferred with his or her scoring leader for advice and guidance. The scoring leader had access to the scorer's performance results and could mentor the scorer on specific areas of scoring inaccuracy. The scoring leader advised the scorer to refer to training content and read over practice responses prior to a second attempt at calibration. Scorers who were unsuccessful after two attempts at calibration were not allowed to score that particular prompt on that day.

#### **7.3.1.4. Training and Certification Samples**

All Writing scoring staff were trained to follow the principles of holistic scoring using sample responses that were drawn from previous ELPAC administrations. After training, scoring staff were required to complete a certification test that also used ELPAC prompts and samples. Only those people who passed the certification test were allowed to score operational responses.

### 7.3.2. Scoring Practices

ETS adhered to the following scoring practices and procedures:

- New scorers had to demonstrate their accuracy by passing a certification test before they were scheduled to score the ELPAC. Scorers had to then pass a shorter, more focused calibration test before each scheduled scoring session. Certification and calibration were described in the previous subsection.
- Scorers underwent training in appropriately applying the rubric for each specific task type, following the generic sample responses that exemplified the quality required for each score point. This ensured that every prompt was scored using the same general criteria.
- If scorers disagreed on a score, the decision moved up to a scoring leader and, if needed, the content scoring leader. Scoring leaders provided adjudications of discrepant scores (i.e., scores more than one score point apart). A scorer discovered to be scoring inaccurately was additionally monitored and might have been required to have additional training. In some cases, if scoring inaccuracies were a problem, the scorer was dismissed.
- ETS trained all levels of scoring leadership, not only on the prompts, rubrics, and related scoring material, but also on how best to monitor the quality of the scoring.
- A second independent score was assigned to 10 percent of the responses. This allowed scoring managers to monitor and verify interrater reliability.
- Scoring leaders read behind and monitored scorers. Some responses were scored a second time (i.e., “read behind”) to check agreement among raters (refer to [10.7.2 Read-Behinds](#) for additional information). Scoring leaders had the option of evaluating responses a scorer had previously scored, with or without the knowledge of the score given (“informed” versus “blind” back-rating).
- Prescored validity samples were inserted into the scoring queues as an ongoing check for scoring accuracy.
- ONE provided operational data on scorers and teams who were reading at unusually slow or rapid rates, allowing scoring leadership to investigate and provide counseling and guidance, if warranted.
- During each scoring session, highly skilled content scoring leaders monitored scoring leaders and their virtual teams by reviewing interrater agreement rates as well as back-scoring agreement rates between scorers and scoring leaders. Content scoring leaders adjudicated any discrepant scores that arose and provided feedback to scorers and scoring leaders as needed.
- ETS assessment specialists and top scoring leadership analyzed interrater reliability statistics to verify that scorers were scoring consistently.

### 7.3.3. Managing Scoring

ETS invited applicants to score the Summative ELPAC Writing domain largely from its existing rater pool of more than 30,000 experienced raters. Raters who accepted the invitation met the rater qualifications and demonstrated their scoring accuracy by passing a certification test before being selected to score for the ELPAC.

### 7.3.3.1. Rater Qualification

The ELPAC pool consisted of two types of raters:

1. Many raters had previous experience with holistic scoring of writing responses from scoring the SAT®, the TOEFL®, or both. These raters were required to complete training using ELPAC-specific materials prior to operational scoring.
2. Other raters who were recruited specifically for ELPAC scoring were also required to complete training using ELPAC-specific materials prior to operational scoring.

[Table 7.1](#) summarizes the rater pool for ELPAC.

**Table 7.1 Summary of Characteristics of ELPAC Raters**

Criteria	Number
Experience teaching in a kindergarten (K)–12 school	341
Currently works in a K–12 school in California	98
Others—Not meeting any of the previous criteria	917

California educators serving as raters should meet the following qualifications:

- Must have a current California teaching credential (although California charter school teachers may or may not have a teaching credential)
- May be retired educators or other administrative staff with a teaching credential who are not current classroom teachers
- Must have achieved, at minimum, a bachelor’s degree

All team leaders and raters are required to qualify before scoring and are informed of what they are expected to achieve in order to qualify.

ETS makes a distinction between training sets and calibration (qualification) sets. Training sets are nonconsequential as the sets provide the raters the opportunity to score sample papers and receive feedback, including the correct score point and rationale associated with that score point and the sample paper. Training sets are a learning tool that the raters are required to complete. Nonadjacent scores may occur in the training sets as minimum agreement standards are not part of training sets.

Upon completion of the required training sets, raters move on to a consequential calibration set that will determine rater eligibility for operational scoring of a particular item type. Calibration (qualification) sets have minimum agreement levels that are enforced, and nonadjacent scores are not allowed. All 0–4 and 0–3 point items adhere to the ELPAC recommendation of a 70 percent exact and 0 percent discrepant (nonadjacent) agreement rate to score.



The standards, provided in [table 7.2](#), are qualification expectations for the various score point ranges and the qualification standard in terms of the percent of exact agreement. This qualification set, like the validity papers discussed in the next subsection ([Monitoring Scoring](#)), has been scored previously by scoring experts. Raters must score the papers in the same manner according to the percentage of agreements listed in [table 7.2](#).

**Table 7.2 Rater Qualification Standard for Agreement with Correct Scores**

Score Point Range	Qualification Standard (Exact Agreement)
0–1	90%
0–2	80%
0–3	70%
0–4	60%

The qualification process is conducted through an online system that captures the results electronically for each individual trainee.

### 7.3.4. Monitoring Scoring

There are proven processes in place for monitoring ELPAC scoring. During the 2018–2019 operational scoring, raters passed a regular calibration test that measured the rater’s ability to accurately apply scores to responses for a particular prompt or task. Scoring leadership mentored the raters with feedback during shifts and completed back-rating of rater scores.

ETS had communication channels in place among raters, leadership, and ETS staff to share information related to operational scoring and personnel concerns. ETS staff monitored these communications and investigated all scoring accuracy and personnel concerns.

## 7.4. Constructed-Response Scoring for Speaking

### 7.4.1. Scorer Training for Speaking

Participants in the Summative ELPAC Administration and Scoring Training (AST), described in subsection [5.2.2 Scoring Training of Trainers Workshops](#), received training on the administration and scoring of the Speaking domain. The training agenda primarily focused on Speaking task types.

#### 7.4.1.1. Methods of Approaching Item Types

Workshop trainers presented each of the six Speaking task types using the following strategies:

- Video of student being administered the task type
- 2012 *California English Language Development Standards, Kindergarten Through Grade 12* associated with the task type
- Test administration procedures
- Rubric overview
- Scoring and prompting guidelines
- Anchors
- Practice scoring

### **7.4.1.2. Agenda**

What follows is the agenda used during the scoring training:

- Section 1—Overview
- Section 2—Test Administration
  - Test Administration
  - Grades 3–12 Listening video
  - Grades 3–12 Reading video
  - Grades 3–12 Writing video
  - Moodle Training Site
- Section 3—Speaking Overview
  - Speaking Overview
  - Full Speaking video
- Section 3—Talk About a Scene (K–12)
- Section 4—Speech Functions (2–12)
- Section 5—Support an Opinion (K–2)
- Section 5—Support an Opinion (3–12)
- Section 6—Retell a Narrative (K–5)
- Section 7—Present and Discuss Information (6–12)
- Section 8—Summarize an Academic Presentation (K–5)
- Section 8—Summarize an Academic Presentation (6–12)
- Section 9—Full Speaking video (K–5)
- Section 9—Full Speaking video (6–12)
- Section 10—K–1 Administration of Reading, Writing, and Listening

### **7.4.1.3. Training Materials**

To establish consistency in statewide local training, training materials were developed and provided to all LEAs. Each person attending training received a printed training binder with access to a PDF provided on the Moodle training website. Participants were also provided with administration training videos and training presentations, with scripts posted on the Moodle training website for LEA trainers to use for their local training of test examiners. The training materials were primarily focused on scoring the Speaking task types. Training materials are described in the next subsections.

#### **7.4.1.3.1. Training Binder**

A Summative ELPAC AST binder was provided to participants in the training. Each binder contained the following sections:

- **Section 1—Introduction**
  - Overview of the program
  - Contact information
  - Program resources

- **Section 2—Test Administration**
  - Testing materials and irregularities training
  - Overview of use of the Test Operations Management System (TOMS) for streaming Listening domain and Speaking
  - Group test administration
  - Matrix Four: Universal Tools, Designated Supports, and Accommodations
  - Other logistics including Moodle training site
- **Section 3—Talk About a Scene**
  - Prompting and scoring guidelines
  - Rubrics
  - Each scene
  - Anchor charts
- **Section 4—Speech Functions**
  - Prompting and scoring guidelines
  - Rubrics
  - Anchors, with more than 96 audio tracks as samples for training and calibration
- **Section 5—Support an Opinion**
  - Prompting and scoring guidelines
  - Rubrics
  - Anchors, with more than 127 audio tracks as samples for training and calibration
- **Section 6—Retell a Narrative**
  - Prompting and scoring guidelines
  - Rubric
  - Anchors, with more than 96 audio tracks as samples for training and calibration
- **Section 7—Present and Discuss Information**
  - Prompting and scoring guidelines
  - Rubrics
  - Anchors, with more than 78 audio tracks as samples for training and calibration
- **Section 8—Summarize an Academic Presentation**
  - Prompting and scoring guidelines
  - Rubrics
  - Anchors, with more than 168 audio tracks as samples for training and calibration
- **Section 9—Video Scoring Practice**
  - Seven full Speaking videos for scoring practice of an entire administration by grade level or grade span
- **Section 10—K–1 Administration**
  - Narrated training video
  - PowerPoint talking points slides

**7.4.1.3.2. Training Videos**

Five test administration videos were created and presented during statewide training; these were made available with the other training materials. Videos used are listed in [table 7.3](#).

**Table 7.3 Available Scoring Training Videos**

<b>Topic</b>	<b>Description</b>
Kindergarten and grade one	The video, which includes narration, presents a kindergarten student being administered all four domains and a grade one student being administered the Writing domain.
Reading, grades three through twelve	The narrated video was recorded with high school students being administered the Reading domain in a group setting.
Writing, grades three through twelve	The narrated video was recorded with high school students being administered the Writing domain in a group setting.
Grade two administration	This narrated, grade two video incorporates small-group directions and testing of grade two Reading, Writing, and Listening.
Listening	This narrated video includes test administration practices and was recorded with middle school students taking the Listening domain with audio streamed through TOMS.
Video scoring practice	Seven full Speaking administration videos were created; each presents a student in a different grade level or grade span being administered the Speaking test.

**7.4.1.3.3. Training Presentations**

Ten training presentations were created for LEA ELPAC trainers to use for local training. These training presentations included all of the Speaking video and audio files to be embedded into the presentations. Most of these presentations focused on training and scoring the Speaking task types.

[Table 7.4](#) includes a list of the training presentations available to LEAs.

**Table 7.4 Available Training Presentations for Speaking**

<b>Binder Section</b>	<b>Training Presentations</b>
Section 1	Introduction Training Presentation
Section 2	Test Administration Training Presentation
Section 3	Talk About a Scene Training Presentation
Section 4	Speech Functions Training Presentation
Section 5	Support an Opinion Training Presentation
Section 6	Retell a Narrative Training Presentation
Section 7	Present and Discuss Information Training Presentation
Section 8	Summarize an Academic Presentation Training Presentation
Section 9	Scoring Video Practice (one presentation per grade level or grade span)
Section 10	K–1 Administration of Reading-Writing-Listening

#### 7.4.1.3.4. Online Training Resources

Moodle provides a password-protected, online platform where course materials can be developed and made available. The ELPAC Moodle training site provides California LEAs with necessary training resources to train test examiners to score the ELPAC. There were 20,694 users as of the close of the Summative ELPAC test administration window on May 31, 2019.

To give test examiners an opportunity to refresh and test their knowledge prior to administering the Summative ELPAC, the online training site included more than 53 training and calibration quizzes with more than 400 audio samples.

To access the ELPAC Moodle training site, LEA users required individual user accounts. Each LEA had its own district group; the LEA ELPAC coordinator was issued a unique enrollment key for the training course and could view the results of the quizzes taken by test examiners, to monitor scoring calibration.

The training quizzes allowed a test examiner to listen to the audio, select a score, and receive feedback. The Moodle quiz provided the correct score, justification, and feedback after the test examiner completed 10 samples.

For items that included artwork, such as *Retell a Narrative* and *Present and Discuss Information*, the picture stimulus was included in the quiz for the test examiner's reference while listening to the audio. A replay feature allowed the test examiner to replay the audio as necessary.

Upon completion of the calibration quiz, the "Pass/Fail" and "Percent correct" notifications were posted for the test examiner.

[Table 7.5](#) shows a list of the training and calibration quizzes by task type created and posted to the Moodle training site.

**Table 7.5 Training and Calibration Quizzes by Task Type**

Task Type	Training Quizzes	Calibration Quizzes
<i>Talk About a Scene</i>	<ul style="list-style-type: none"> <li>• Kindergarten video quiz</li> <li>• Grade 1 video quiz</li> <li>• Grade 2 video quiz</li> <li>• Grades 3–5 video quiz</li> <li>• Grade 6–8 video quiz</li> <li>• Grade 9–10 video quiz</li> <li>• Grade 11–12 video quiz</li> </ul>	[None]
<i>Speech Functions</i>	<ul style="list-style-type: none"> <li>• Grades 2–12</li> <li>• Grades 2–5</li> <li>• Grades 6–8</li> <li>• Grades 9–12</li> </ul>	<ul style="list-style-type: none"> <li>• Grades 2–12</li> <li>• Grades 2–5</li> <li>• Grades 6–8</li> <li>• Grades 9–12</li> </ul>
<i>Support an Opinion</i>	<ul style="list-style-type: none"> <li>• Grades K–2</li> <li>• Grades 3–5</li> <li>• Grades 6–8</li> <li>• Grades 9–12</li> <li>• Grades 3–12</li> </ul>	<ul style="list-style-type: none"> <li>• Grades K–2</li> <li>• Grades 3–5</li> <li>• Grades 6–8</li> <li>• Grades 9–12</li> <li>• Grades 3–12</li> </ul>

Table 7.5 (continuation)

Task Type	Training Quizzes	Calibration Quizzes
<i>Retell a Narrative</i>	<ul style="list-style-type: none"> <li>• Kindergarten</li> <li>• Grade 1</li> <li>• Grade 2</li> <li>• Grades 3–5</li> </ul>	<ul style="list-style-type: none"> <li>• Kindergarten</li> <li>• Grade 1</li> <li>• Grade 2</li> <li>• Grades 3–5</li> </ul>
<i>Present and Discuss Information</i>	<ul style="list-style-type: none"> <li>• Grades 6–8</li> <li>• Grades 9–10</li> <li>• Grades 11–12</li> </ul>	<ul style="list-style-type: none"> <li>• Grades 6–8</li> <li>• Grades 9–10</li> <li>• Grades 11–12</li> </ul>
<i>Summarize an Academic Presentation</i>	<ul style="list-style-type: none"> <li>• Kindergarten</li> <li>• Grade 1</li> <li>• Grade 2</li> <li>• Grades 3–5</li> <li>• Grades 6–8</li> <li>• Grades 9–10</li> <li>• Grades 11–12</li> </ul>	<ul style="list-style-type: none"> <li>• Kindergarten</li> <li>• Grade 1</li> <li>• Grade 2</li> <li>• Grades 3–5</li> <li>• Grades 6–8</li> <li>• Grades 9–10</li> <li>• Grades 11–12</li> </ul>

### 7.4.2. Scorer Qualifications for Speaking

The Speaking domain was scored by test examiners “in the moment.” All test examiners were required to receive the Speaking scoring training from an LEA trainer.

## 7.5. Types of Scores

### 7.5.1. Raw Scores

Raw scores for each domain were obtained by summing the number of MC and machine-scorable CR items answered correctly and adding the total number of points obtained on the hand-scored CR items within the Speaking and Writing domains and the K and grade one Reading domain.

The domain raw scores from Listening and Speaking were summed to compute the oral language raw score. The domain raw scores from Reading and Writing were summed to compute the written language raw score. The number and percentage of students at each raw score, and the associated level, are reported for each domain in table 7.A.1 through table 7.A.52 in [appendix 7.A](#).

### 7.5.2. Scale Scores

Raw scores are not directly comparable from administration to administration because each raw score is based on a set of items that may differ in difficulty. Student performance on the ELPAC is reported in terms of scale scores that express student proficiency in terms of a constant metric. Thus, a scale score of 1350 in one language skill area in one administration represents the same level of proficiency as 1350 on the same language skill area in another administration, even though each scale score may represent a different raw score.

ELPAC scale scores are expressed as four-digit numbers that range from 1150 to 1950 across grade levels and grade spans. Lower scores indicate lesser proficiency and higher scores indicate greater proficiency.

### 7.5.2.1. Scale Score Conversions

For each language skill area, the following steps are used to establish the raw-score-to-scale-score relationship. The process begins by inverting the test characteristic curve (Stocking, 1996) where each possible raw score is mapped to a corresponding theta score. These theta scores represent a student's ability level on a particular language skill and are transformed onto their respective language skill area through a linear transformation as described in equation 7.1:

$$\text{Scale score} = \text{Intercept} + \text{Slope} \times (\text{theta score}) \quad (7.1)$$

Refer to subsection *11.5.6 Developing Summative ELPAC Reporting Scales* in the *Summative English Language Proficiency Assessments for California Technical Report, 2017–18 Administration* (CDE, 2019) for applicable scaling constraints (e.g., slope and intercept terms) for converting theta scores to the oral language and written language scales.

Through this process, raw-to-scale-score conversion tables are established. The complete raw-to-scale-score conversion tables for oral and written language skills are presented in the tables in [appendix 8.D](#).

### 7.5.2.2. Overall Scale Score

The overall scale score is calculated as the weighted average of the scale scores of the oral and written language skills scale scores. For K, the overall scores are calculated as the weighted average scores of the two composite scores as shown in equation 7.2:

$$0.70 \times \text{Oral language score} + 0.30 \times \text{Written language score} \quad (7.2)$$

For grades one through twelve, the overall scores are calculated as the average scores of the two composite scores as shown in equation 7.3:

$$0.50 \times \text{Oral language score} + 0.50 \times \text{Written language score} \quad (7.3)$$

The frequency distributions of raw score, scale score, and level for composite language skills are presented in [appendix 7.B](#), in table 7.B.1 through table 7.B.26. Additionally, [appendix 7.C](#) provides the overall scale score distribution for each grade.

Refer to subsection *11.5.6 Developing Summative ELPAC Reporting Scales* in the *Summative English Language Proficiency Assessments for California Technical Report, 2017–18 Administration* (CDE, 2019) for more details regarding how the Summative ELPAC reporting scales were established.

## 7.5.3. ELPAC Levels

Reporting scales for the Summative ELPAC's two composite language skills and overall scores classify each student's performance into one of the four levels, which are as follows:

1. Level 1—Beginning stage of developing English skills
2. Level 2—Somewhat developed English skills
3. Level 3—Moderately developed English skills
4. Level 4—Well developed English skills (indicating the highest level of performance)

[Appendix 7.D](#) provides a summary of student ELPAC levels for each of the composite language skills. Each table presents the number and percentage of students at each ELPAC level for K through grade twelve.



To guide the interpretation of the scale scores for each domain, the range of possible scale scores for each domain is divided into three levels:

1. Level 1—Beginning
2. Level 2—Somewhat/Moderately Developed
3. Level 3—Well Developed

[Appendix 7.E](#) provides a summary of student levels for each language domain. Each table presents the number and percentage of students at each level for K through grade twelve.

The scale score ranges defining the various reporting levels and grade levels or grade spans are presented in [table 7.6](#).

**Table 7.6 Composite Language Skills and Overall Reporting Scale Score Ranges for Each Reporting Level by Grade Level or Grade Span**

Grade Level or Grade Span	Test	Level 1	Level 2	Level 3	Level 4
Kindergarten	Overall	1150–1373	1374–1421	1422–1473	1474–1700
Kindergarten	Oral Language	1150–1385	1386–1426	1427–1477	1478–1700
Kindergarten	Written Language	1150–1345	1346–1409	1410–1462	1463–1700
Grade 1	Overall	1150–1410	1411–1454	1455–1506	1507–1700
Grade 1	Oral Language	1150–1407	1408–1450	1451–1492	1493–1700
Grade 1	Written Language	1150–1413	1414–1458	1459–1519	1520–1700
Grade 2	Overall	1150–1423	1424–1470	1471–1531	1532–1700
Grade 2	Oral Language	1150–1413	1414–1459	1460–1509	1510–1700
Grade 2	Written Language	1150–1432	1433–1480	1481–1553	1554–1700
Grade 3	Overall	1150–1447	1448–1487	1488–1534	1535–1800
Grade 3	Oral Language	1150–1434	1435–1465	1466–1511	1512–1800
Grade 3	Written Language	1150–1460	1461–1508	1509–1556	1557–1800
Grade 4	Overall	1150–1458	1459–1498	1499–1548	1549–1800
Grade 4	Oral Language	1150–1438	1439–1471	1472–1521	1522–1800
Grade 4	Written Language	1150–1477	1478–1524	1525–1574	1575–1800
Grade 5	Overall	1150–1466	1467–1513	1514–1559	1560–1800
Grade 5	Oral Language	1150–1446	1447–1476	1477–1532	1533–1800
Grade 5	Written Language	1150–1486	1487–1549	1550–1586	1587–1800
Grade 6	Overall	1150–1474	1475–1516	1517–1566	1567–1900
Grade 6	Oral Language	1150–1449	1450–1483	1484–1541	1542–1900
Grade 6	Written Language	1150–1498	1499–1549	1550–1591	1592–1900
Grade 7	Overall	1150–1480	1481–1526	1527–1575	1576–1900
Grade 7	Oral Language	1150–1455	1456–1497	1498–1553	1554–1900
Grade 7	Written Language	1150–1504	1505–1555	1556–1597	1598–1900
Grade 8	Overall	1150–1485	1486–1533	1534–1589	1590–1900
Grade 8	Oral Language	1150–1460	1461–1504	1505–1568	1569–1900
Grade 8	Written Language	1150–1509	1510–1561	1562–1609	1610–1900



Table 7.6 (continuation)

Grade Level or Grade Span	Test	Level 1	Level 2	Level 3	Level 4
Grade span 9–10	Overall	1150–1492	1493–1544	1545–1605	1606–1950
Grade span 9–10	Oral Language	1150–1464	1465–1511	1512–1578	1579–1950
Grade span 9–10	Written Language	1150–1519	1520–1577	1578–1631	1632–1950
Grade span 11–12	Overall	1150–1499	1500–1554	1555–1614	1615–1950
Grade span 11–12	Oral Language	1150–1469	1470–1513	1514–1582	1583–1950
Grade span 11–12	Written Language	1150–1528	1529–1594	1595–1645	1646–1950

The threshold scores in [table 7.6](#) are updates to the 2017–2018 threshold scores adopted by the State Board of Education (SBE) in November 2017 for the 2017–2018 administration of the Summative ELPAC. The original threshold scores established through an ELPAC standard setting workshop were revised based on the results of the Summative Threshold Score Validation Study (CDE, 2018) and other analyses. These changes were adopted by the SBE in November 2018 for the 2018–2019 administration and beyond.

## 7.6. Overview of Score Aggregation

To provide meaningful results to the stakeholders, test scores for a given grade are aggregated at the school, LEA or direct funded charter school, county, and state levels. The aggregated scores are generated for the selected groups of interest (gender, ethnicity, primary disability, etc.) and for the total population. This subsection of the report contains a description of the types of aggregation that are performed on the Summative ELPAC test summary scores.

### 7.6.1. Individual Student Score Distributions and Summary Statistics

Summary statistics that describe student performance on a test are presented in [table 7.7](#) mean and standard deviation of overall, written, and oral language scale scores. Included in

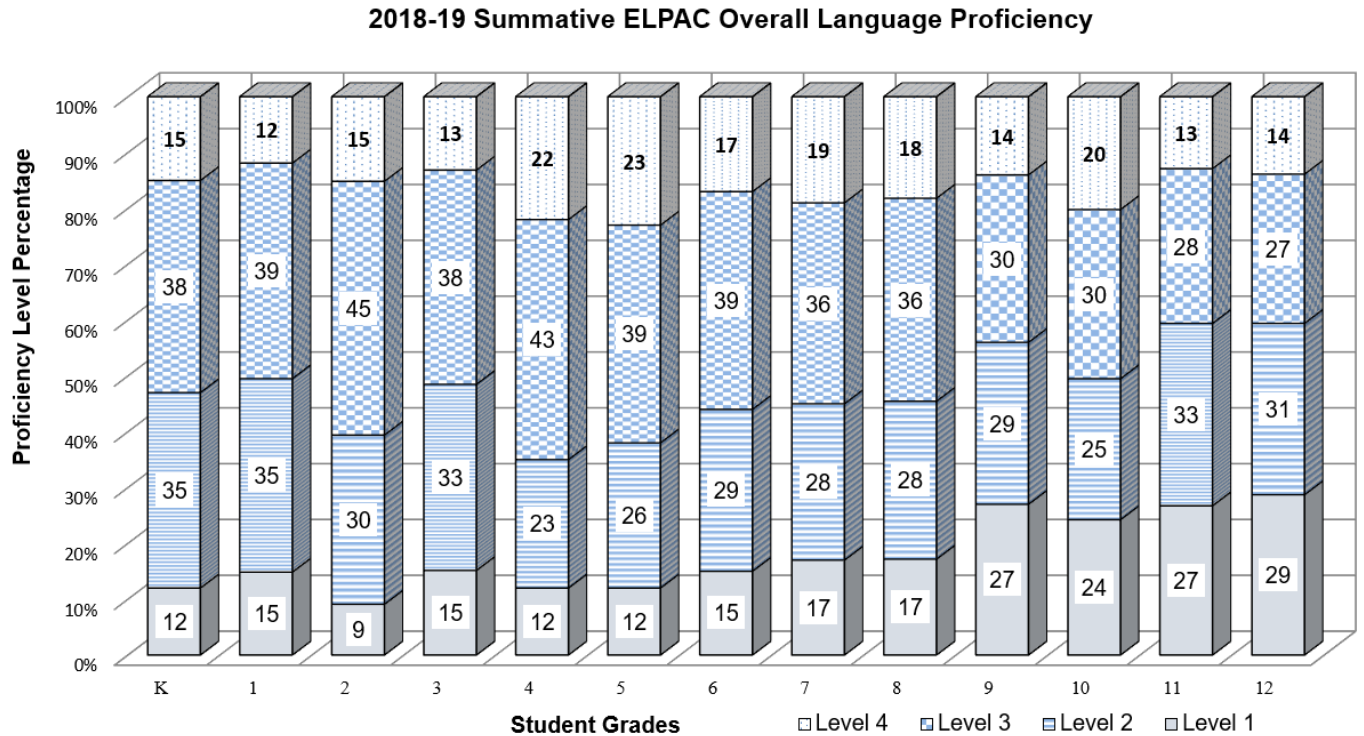
the table are the number of students taking each test and the means and standard deviations of student scores expressed in terms of scale scores.

**Table 7.7 Mean and Standard Deviation of Overall, Written, and Oral Language Scale Scores**

<b>Grade</b>	<b>Number of Students Tested</b>	<b>Overall Scale Score Mean</b>	<b>Overall Scale Score SD</b>	<b>Oral Scale Score Mean</b>	<b>Oral Scale Score SD</b>	<b>Written Scale Score Mean</b>	<b>Written Scale Score SD</b>
Kindergarten	154,118	1426	67	1434	67	1405	89
Grade 1	127,703	1453	64	1460	64	1446	78
Grade 2	115,895	1482	64	1486	68	1479	72
Grade 3	100,010	1485	61	1482	69	1488	62
Grade 4	101,702	1509	65	1506	75	1512	64
Grade 5	90,236	1521	69	1517	80	1524	68
Grade 6	76,325	1517	73	1514	89	1519	68
Grade 7	70,245	1526	80	1523	98	1528	73
Grade 8	58,729	1533	85	1529	104	1536	78
Grade 9	54,965	1524	99	1519	123	1529	87
Grade 10	47,296	1537	102	1534	127	1539	90
Grade 11	41,958	1531	98	1519	114	1542	94
Grade 12	36,600	1516	128	1507	140	1524	127

The percentage of students at each proficiency level for overall, written, and oral language are presented in [table 7.8](#) through [table 7.10](#) and presented in [figure 7.1](#) through [figure 7.3](#). Note that the percentage for the proficiency levels may not sum to exactly 100 because of rounding.

Figure 7.1 presents the percentages of students at each proficiency level by grade for 2018–2019 Summative ELPAC Overall proficiency levels. This figure was generated using data from table 7.8.



**Figure 7.1 Percentage of students at each overall proficiency level**

**Table 7.8 Percentage of Students in Each Proficiency Level—Overall**

Grade Level	Proficiency Level 1	Proficiency Level 2	Proficiency Level 3	Proficiency Level 4
Kindergarten	12	35	38	15
Grade 1	15	35	39	12
Grade 2	9	30	45	15
Grade 3	15	33	38	13
Grade 4	12	23	43	22
Grade 5	12	26	39	23
Grade 6	15	29	39	17
Grade 7	17	28	36	19
Grade 8	17	28	36	18
Grade 9	27	29	30	14
Grade 10	24	25	30	20
Grade 11	27	33	28	13
Grade 12	29	31	27	14

Figure 7.2 presents the percentages of students at each oral language proficiency level by grade. This figure was generated using data from table 7.9.

2018-19 Summative ELPAC Oral Language Skill

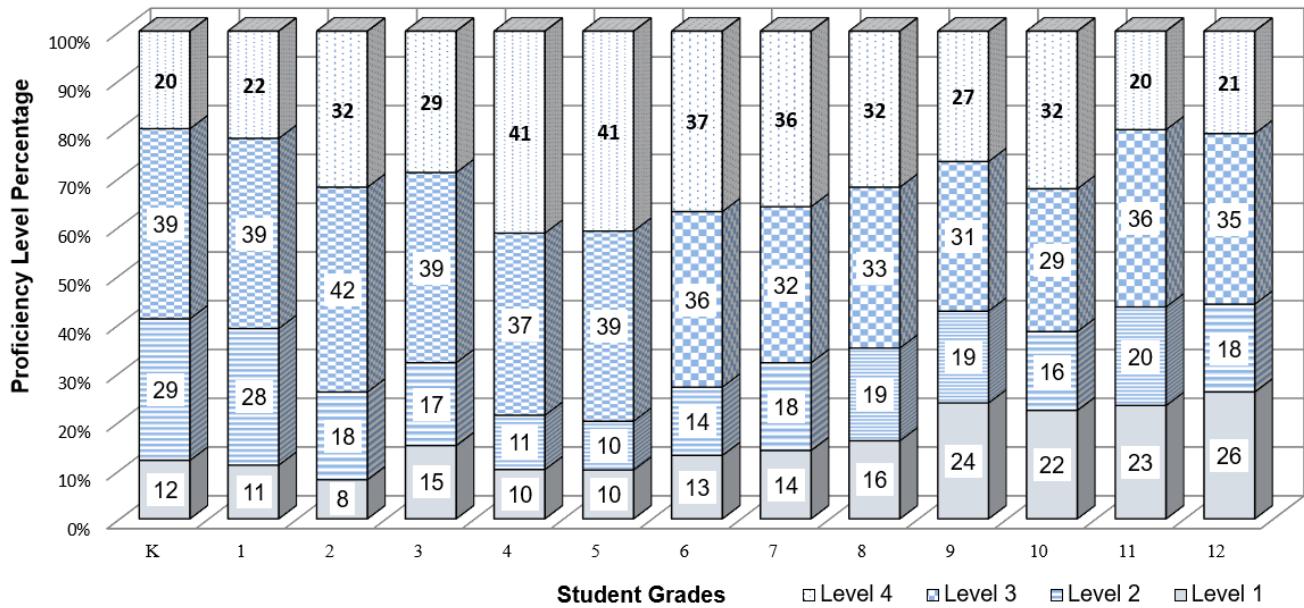
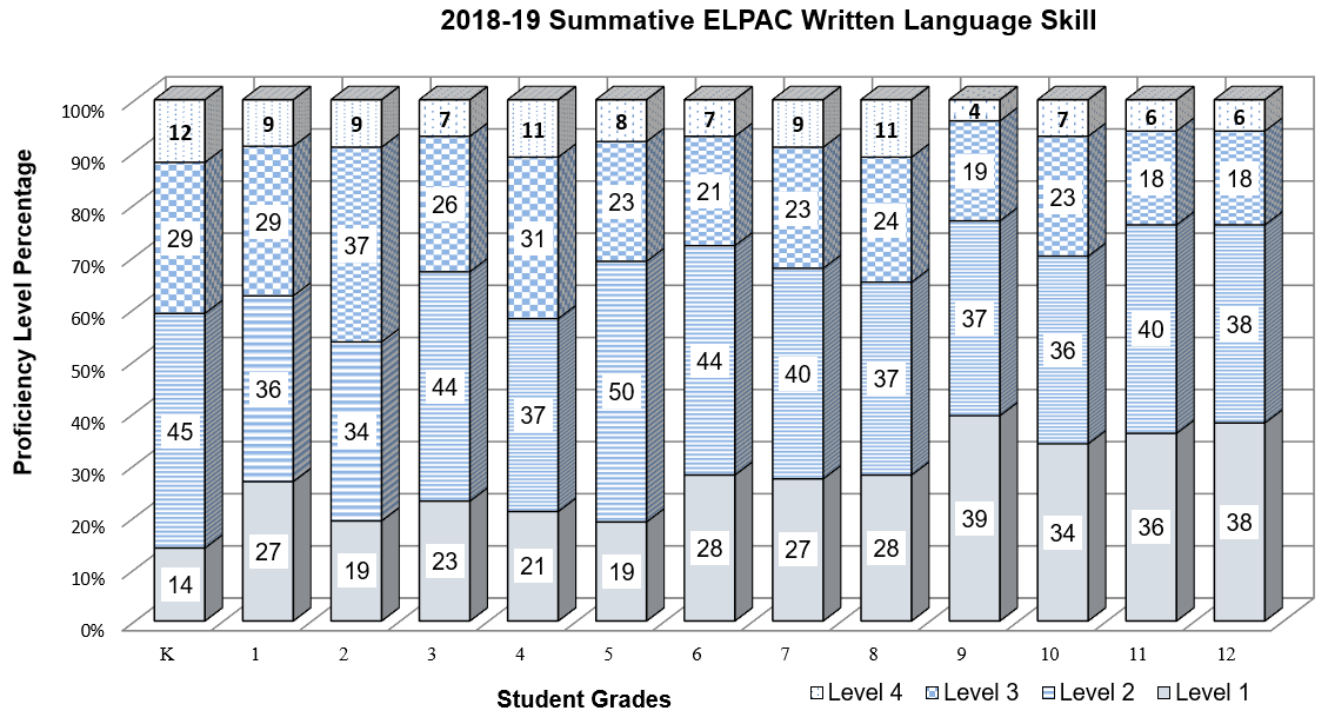


Figure 7.2 Percentage of students at each oral language proficiency level

Table 7.9 Percentage of Students in Each Proficiency Level—Oral Language Composite

Grade Level	Proficiency Level 1	Proficiency Level 2	Proficiency Level 3	Proficiency Level 4
Kindergarten	12	29	39	20
Grade 1	11	28	39	22
Grade 2	8	18	42	32
Grade 3	15	17	39	29
Grade 4	10	11	37	41
Grade 5	10	10	39	41
Grade 6	13	14	36	37
Grade 7	14	18	32	36
Grade 8	16	19	33	32
Grade 9	24	19	31	27
Grade 10	22	16	29	32
Grade 11	23	20	36	20
Grade 12	26	18	35	21

Figure 7.3 presents the percentages of students at each written language skill proficiency level by grade. This figure was generated using data from table 7.10.



**Figure 7.3 Percentage of students at each written language proficiency level**

**Table 7.10 Percentage of Students in Each Proficiency Level—Written Language Composite**

Grade Level	Proficiency Level 1	Proficiency Level 2	Proficiency Level 3	Proficiency Level 4
Kindergarten	14	45	29	12
Grade 1	27	36	29	9
Grade 2	19	34	37	9
Grade 3	23	44	26	7
Grade 4	21	37	31	11
Grade 5	19	50	23	8
Grade 6	28	44	21	7
Grade 7	27	40	23	9
Grade 8	28	37	24	11
Grade 9	39	37	19	4
Grade 10	34	36	23	7
Grade 11	36	40	18	6
Grade 12	38	38	18	6

The summary performances for the two composite language skills and overall scores for selected groups of students are provided in appendix 7.F, in table 7.F.1 through table 7.F.39. In these tables, students are grouped by demographic characteristics,

including gender, ethnicity, economic status (disadvantaged or not), migrant status, and special education services status.

For each student group, the number tested, scale score means, standard deviations, and the percentage of students in each level are reported.

[Table 7.11](#) provides definitions of demographic student groups.

**Table 7.11 Demographic Student Groups Reported**

Category	Student Groups
Gender	<ul style="list-style-type: none"> <li>• Male</li> <li>• Female</li> </ul>
Ethnicity	<ul style="list-style-type: none"> <li>• American Indian or Alaska Native</li> <li>• Asian</li> <li>• Native Hawaiian or Other Pacific Islander</li> <li>• Filipino</li> <li>• Hispanic or Latino</li> <li>• Black or African American</li> <li>• White</li> <li>• Two or more races</li> </ul>
Special Education Service Status	<ul style="list-style-type: none"> <li>• Students not receiving special education services</li> <li>• Students receiving special education services</li> </ul>
Economic Status	<ul style="list-style-type: none"> <li>• Not economically disadvantaged</li> <li>• Economically disadvantaged</li> </ul>
Migrant Status	<ul style="list-style-type: none"> <li>• Eligible for the Title I Part C Migrant Program (Migrant)</li> <li>• Not eligible for the Title I Part C Migrant Program (Nonmigrant)</li> </ul>

## 7.7. Reports Produced and Scores for Each Report

### 7.7.1. Online Reporting

TOMS is a secure website hosted by ETS that permits LEA users to manage aspects of the ELPAC administration and report delivery. This system used a role-specific design to restrict access to certain tools and applications based on the user’s designated role. Specific functions of TOMS included the following:

- Manage user access privileges
- Manage test material orders
- Run and download various reports

### 7.7.2. Special Cases

All students identified as English learners (ELs) were required to take the Summative ELPAC. There were no special cases that excuse a student from receiving a score. In instances where a student’s individualized education program or Section 504 plan specified that the student had a disability for which there were no appropriate accommodations for assessment in one or more of the Speaking, Listening, Reading, and Writing domains, the student was assessed in the remaining domains in which it was possible to assess the student, per the *Code of Federal Regulations*, Title 34, Section 200.6. For the domains that



were not possible to assess, the student's Answer Book was marked with a domain exemption. Scores for the composite containing the exempt domain were calculated using the remaining domain, if it was not exempt.

If the student was administered a locally available alternate form of the assessment for a domain, the student's Answer Book was marked with an alternate assessment for that domain. The domain(s) marked with alternate assessment received the lowest obtainable scale score.

Note the following about special reporting cases:

- A student may have been assigned an overall score only if assessed in both oral and written language. To be considered as having been assessed in oral language, the student must have been assessed in either Speaking or Listening. To be considered as having been assessed in written language, the student must have been assessed in either Reading or Writing.
- A valid score could only be provided in those instances where the student tested in at least one of the domains for oral language and written language. In all instances where the overall score resulted in No Score, the student counted as tested without a valid score. Students who had domain exemptions for both domains in a single composite could not be given an overall score; however, these students were counted as tested.

### 7.7.3. Types of Score Reports

The following is a list of score reports produced for the 2018–2019 Summative ELPAC:

- SSR—The SSR was the official student score report for the parents or guardians and described the student's results.
- Tested LEA student data files and corresponding aggregate files—Aggregate files were used for public web reporting on the Test Results for California's Assessments website at <https://caaspp-elpac.cde.ca.gov/elpac/> and for CDE apportionment. LEA student data files were available for download on demand by the LEA in TOMS to coincide with the SSRs.
- State student data files—The state student data files were the full operational file and included 100 percent of the student scores and eligibility data. This file was provided to the CDE and was used for apportionment.

#### 7.7.3.1. Student Score Reports

The SSR was the official score report for the parents or guardians and describes the student's results. For the 2018–2019 administration, SSRs were made available to the LEAs in English, Spanish, Filipino, Chinese (Traditional), and Vietnamese. An SSR in a supported language was created if the student's primary language as reported in the California Longitudinal Achievement Data System was one of these supported languages.

The LEAs that received SSRs in supported languages received one SSR in English and another in an SSR-supported language. The LEAs that requested purchased paper SSRs received two paper versions, one in English and another in an SSR-supported language. These reports were also available as PDFs for the LEA to download from TOMS.

The SSR included the following information:

- Overall score and reporting level
- Oral language score and reporting level
- Written language score and reporting level
- Domain performance levels

As mentioned previously, overall score—the oral language score and written language score—placed a student within one of the four ELPAC reporting levels as Beginning Stage, Somewhat Developed, Moderately Developed, or Well Developed. For each domain, a student is placed within one of three proficiency levels as Beginning, Somewhat/Moderately Developed, or Well Developed.

#### **7.7.3.2. School Reports**

Site ELPAC coordinators could download individual SSRs or bulk download a compressed (.zip) file of student SSRs for the school from TOMS.

#### **7.7.3.3. LEA Reports**

LEAs had the option of downloading the following ELPAC reports from TOMS:

- LEA student data files
- LEA-level aggregate files

### **7.7.4. Score Report Applications**

Summative ELPAC results provided parents and guardians with information about their child’s progress toward English proficiency. The results were a tool for increasing communication and collaboration between parents or guardians and teachers.

Summative ELPAC results were one of the components schools could use to help make decisions about how best to support student progress. Summative ELPAC results, however, should never be used as the only source of information to make important decisions about a child’s education.

### **7.7.5. Criteria for Interpreting Test Scores**

An LEA may use ELPAC results to help make decisions about student placement in programs that support the student’s ongoing development toward English proficiency. However, it is important to remember that a single test can provide only limited information. Other relevant information should be considered as well. It is advisable for parents or guardians to evaluate their child’s progress by looking at classroom work and progress reports in addition to the child’s ELPAC results.

LEAs may use ELPAC results to help make decisions about student placement in EL programs, student exit from EL programs, and student growth in proficiency while in EL programs. The ELPAC, however, is a single measure of student performance and is intended to be used in combination with other relevant information in the decision-making process. Test scores must be interpreted cautiously when making decisions about student or program performance.

2018–2019 Summative ELPAC reporting levels represented broad ranges of proficiency with wide gradations between the lowest and highest possible scores in each range that were reflected in student performance. While statistical procedures were carefully applied to ensure a continuous scale throughout the full range of the common scale, ETS recommends using caution in comparing individual student performance across



nonadjacent grade spans. Although the common scales have the same general properties across domains, numeric comparisons across domains cannot be made—a student scoring 400 in oral language and 420 in written language is not necessarily doing better in terms of written skills.

### **7.7.6. Criteria for Interpreting Score Reports**

Summative ELPAC scores represented only one view of a child’s progress toward language proficiency. It is advisable for parents/guardians to evaluate their child’s progress by looking at classroom work and progress reports in addition to the child’s ELPAC results before making reclassification decisions.

Because the Summative ELPAC results were vertically scaled, scale scores for a test may be compared to scale scores for the same student or groups of students in different years, as well as for between specific grade levels. This allows users to say that proficiency for a given grade was higher or lower one year as compared with another. For example, the grade two Summative ELPAC scale scores in 2017–2018 and 2018–2019 may be compared, as can the grade five Summative ELPAC scale score in 2017–2018 and the grade six Summative ELPAC scale score in 2018–2019, because of the vertical scale.

## References

- California Department of Education. (2018). *English Language Proficiency Assessments for California threshold score validation study final report*. [Unpublished report]. Sacramento, CA: California Department of Education.
- California Department of Education. (2019). *Summative English Language Proficiency Assessments for California technical report, 2017–18 administration*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/ta/tq/ep/documents/elpac18techrpt.pdf>
- Stocking, M. L. (1996). An alternative method for scoring adaptive tests. *Journal of Educational and Behavioral Statistics*, 21, 365–89.

# Chapter 8: Test Analyses and Results

---

This chapter summarizes the item- and test-level statistics from the analyses conducted for the 2018–2019 operational administration of the English Language Proficiency Assessments for California (ELPAC).

## 8.1. Background

This chapter provides information on the psychometric analyses of the 2018–2019 Summative ELPAC operational data. It describes the data samples used for statistical analyses and presents the results of the item and test analyses, such as classical item analyses and differential item functioning (DIF). It includes explanations for all statistical procedures implemented during the psychometric analyses, including reliability estimates, standard errors of measurement, and decision consistency and accuracy of the performance-level classifications. Information on the procedures designed to ensure the validity of scores and interpretations is also provided.

### 8.1.1. Summary of the Analyses

Each of these sets of analyses for the Summative ELPAC is presented in the body of the text and in the listed appendices.

1. **Classical Item Analyses**—Classical item analysis for the Summative ELPAC is discussed in subsection [8.2 Classical Item Analysis Statistics](#). [Appendix 8.A](#) presents results of the classical item analyses, including item difficulty indices, item-total correlation coefficients, and the omission rates for multiple choice (MC) and constructed-response (CR) items. In addition, the distribution of score points for the CR with multiple score points and the machine-scored MC items is provided.
2. **Item Response Theory (IRT) Analyses**—IRT analyses, including calibrations, are presented in subsection [8.3 Item Response Theory Analyses](#). [Appendix 8.B](#) includes distribution of IRT *a*-values, *b*-values, and item statistics by domain.
3. **DIF Analyses**—DIF analysis is described in subsection [8.4 Differential Item Functioning \(DIF\)](#). [Table 8.14](#) presents the results of the DIF analyses for all items of the Summative ELPAC.
4. **Reliability Analyses**—Reliability estimation is illustrated in subsection [8.5 Reliability Analyses](#). The following results of the analyses are presented:
  - [Appendix 8.C](#) provides results of the reliability analyses of total test scores for the selected student groups of interest (e.g., gender, ethnicity).
  - [Appendix 8.D](#) presents the raw-score-to-scale-score conversion tables with the conditional standard errors of measurement (CSEM) for the oral and written language composites.
  - [Appendix 8.E](#) provides interrater reliability statistics showing the agreement between two raters for Writing items.
  - [Appendix 8.F](#) presents statistics describing the decision accuracy and decision consistency of the performance classifications.

### 8.1.2. Samples for the Analyses

In general, analyses included in the technical report are based on all valid students' scores in the tested population. The actual data sample used depended on the availability date and content of the data file. Additionally, a student data file was selected to meet an analysis timeline. Students taking the braille version were excluded from these item analyses.

[Table 8.1](#) shows the number of students tested by grade level. The data includes the Summative ELPAC population comprised of students who have been identified as English learners in kindergarten through grade twelve. The N counts here may not match those in other reports, nor will they always match those shown in other tables and appendices of this report, due to different reporting specifications requiring demographic information that may be missing from some records. Students with an include indicator of “Y” in [table 8.1](#) were used for the [chapter 8](#) analyses. [Table 8.1](#) also presents the number of excluded students using braille at each grade level.

**Table 8.1 Number of Students Tested by Include Indicator and Grade Level**

Grade Level	Y	Total Number Tested	Braille Count
Kindergarten	154,118	154,122	3
Grade 1	127,703	127,705	5
Grade 2	115,895	115,898	5
Grade 3	100,010	100,015	7
Grade 4	101,702	101,708	5
Grade 5	90,236	90,239	8
Grade 6	76,325	76,328	6
Grade 7	70,245	70,247	6
Grade 8	58,729	58,735	7
Grade 9	54,965	54,965	0
Grade 10	47,296	47,298	1
Grade 11	41,958	41,963	0
Grade 12	36,600	36,603	2

**Note:** “Y” indicates students who were enrolled during the active testing window and completed at least one of the two domains in each composite.

## 8.2. Classical Item Analysis Statistics

Many of the statistics that are commonly used for evaluating tests, such as  $p$ -values, point-biserial correlations, DIF classifications, and reliability coefficients arise from classical test theory. These item analyses were conducted for each item across all domains. The students who took the braille version were excluded from these item analyses.

Detailed results of these item analyses are presented in [appendix 8.A](#) and are summarized in the tables in this chapter.

## 8.2.1. Description of Classical Item Analysis Statistics

The classical item analyses include the item difficulty indices and the item-total correlation indices. Flagging rules associated with these statistics identify items that are not performing as expected. The omit rate of each item, the proportion of test takers choosing each distractor, the correlation of each distractor with the total score, and the distribution of students at each score point for the CR items are also included in the classical item analyses.

### 8.2.1.1. Classical Item Difficulty Indices ( $p$ -value)

For MC items, item difficulty is indicated by the  $p$ -value, which is the proportion of students who answer an item correctly. The range of  $p$ -values is from 0.00 to 1.00, inclusive. Items with higher  $p$ -values are easier items; those with lower  $p$ -values are more difficult items.

The formula for  $p$ -value for an MC item is:

$$p\text{-value}_{MC} = \frac{\sum X_{ic}}{N_i}, \quad (8.1)$$

Refer to the [Alternative Text for Equation 8.1](#) for a description of this equation.

where,

$X_{ij}$  is the score received for a given MC item  $I$  for student  $j$ , and

$N_i$  is the total number of students who were presented with item  $i$ .

For CR items, difficulty is indicated by the average item score (AIS). The AIS can range from 0.00 to the maximum total possible points for an item. To facilitate interpretation, the AIS values for CR items or machine-scorable CR items are often expressed as the proportion of the maximum possible score, which is analogous to the  $p$ -values of dichotomous items.

For CR items, the  $p$ -value is defined as:

$$p\text{-value}_{CR} = \frac{\sum X_{ij}}{N_i \times \text{Max}(X_i)}, \quad (8.2)$$

Refer to the [Alternative Text for Equation 8.2](#) for a description of this equation.

where,

$X_{ij}$  is the score received for a given CR item  $i$  for student  $j$ ,

$\text{Max}(X_i)$  is the maximum score for item  $i$ , and

$N_i$  is the total number of students who were presented with item  $i$ .

The Summative ELPAC  $p$ -values were generally within the expected range of above 0.20 and below 0.95; most were also in the desired difficulty range of 0.30 to 0.90. These ranges were defined to produce items that represented item difficulties throughout the range of student proficiency.

Mean item  $p$ -values are presented in [table 8.2](#). In general, the mean  $p$ -values were within acceptable ranges. The mean  $p$ -values indicate that many of the tests were relatively easy. The lowest mean  $p$ -values were observed for Reading with grade spans six through eight, nine and ten, and eleven and twelve demonstrating mean  $p$ -values of 0.436, 0.514, and 0.522, respectively.

**Table 8.2 Mean  $p$ -values**

Grade Level or Grade Span	Listening Mean $p$ -value	Speaking Mean $p$ -value	Reading Mean $p$ -value	Writing Mean $p$ -value
Kindergarten	0.719	0.736	0.679	0.759
Grade 1	0.785	0.784	0.665	0.732
Grade 2	0.800	0.828	0.733	0.750
Grade span 3–5	0.720	0.857	0.600	0.685
Grade span 6–8	0.729	0.834	0.436	0.706
Grade span 9–10	0.699	0.789	0.514	0.695
Grade span 11–12	0.648	0.809	0.522	0.690

**8.2.1.2. Item-Total Correlation**

An important indicator of item discrimination is the point-biserial correlation (i.e., item-total correlation), defined as the correlation between student scores on an individual item and student “total” scores on the test (after excluding the scores of the item in question). These statistics are included in the item analysis tables in [appendix 8.A](#).

To calculate point-biserial correlations by domain, domain scores are used instead of total scores. In general, the item-total correlation ranges from -1.0 (a perfect negative relationship) to 1.0 (a perfect positive relationship). A relatively high positive item-total correlation is desired, as it indicates that students with higher scores on the test tended to perform better on the item than students with lower test scores. A negative item-total correlation typically signifies a problem with the item, because it indicates that students with low scores on the test are getting higher scores on the item than students with high scores on the test.

To avoid artificially inflating the correlation coefficients, the contribution of the item being analyzed was first removed from the total score when calculating each of the correlations. Thus, performance on each Listening item was correlated with the total Listening score minus the score on the item in question. Likewise, performance on each Reading item was correlated with the total Reading score minus the score on the item in question, and so on for the Speaking and Writing items.

[Table 8.3](#) reports the mean point-biserial correlations by grade level or grade span and domain. Some of the mean point-biserial correlations were very low. For example, the mean correlations for the Listening and Reading domains for grade span six through eight were 0.381 and 0.376, respectively. The mean correlations for grade span eleven and twelve were 0.399 for Listening and 0.393 for Reading.

Desired values for this correlation are positive and larger than 0.20. Negative item-total correlations indicate that low-ability students obtain higher scores on the item than high-ability students, an indication that the scoring key may be incorrect. Items with item-total correlations below 0.20 were flagged for review.

As shown in [Table 8.3](#), mean point-biserial correlations were within acceptable ranges across the four domains.

**Table 8.3 Mean Point-Biserial Correlation**

Grade Level or Grade Span	Listening Mean Point-Biserial Correlations	Speaking Mean Point-Biserial Correlations	Reading Mean Point-Biserial Correlations	Writing Mean Point-Biserial Correlations
Kindergarten	0.479	0.724	0.523	0.780
Grade 1	0.479	0.677	0.628	0.697
Grade 2	0.438	0.670	0.501	0.704
Grade span 3–5	0.416	0.714	0.449	0.724
Grade span 6–8	0.381	0.736	0.376	0.702
Grade span 9–10	0.444	0.805	0.433	0.694
Grade span 11–12	0.399	0.759	0.393	0.687

### 8.2.2. Summary of Classical Item Analysis Flagging Criteria

Items are flagged for review if the item analysis yields any of the following results, including both MC and CR items:

1. The  $p$ -value is above 0.95.
2. The  $p$ -value is below 0.20.
3. Item-total correlation (point-biserial) is below 0.20.
4. Among the highest-performing students (the top 20 percent), the number of students choosing any distractor is greater than the number choosing the key.

Educational Testing Service (ETS) psychometric staff and content assessment development staff carefully reviewed each of the items flagged after the 2018–2019 Summative ELPAC administration and summarized the results for the California Department of Education, with recommendations for subsequent analyses. These results were also entered into the item bank and used by the assessment development team for test assembly for future operational administrations.

### 8.2.3. Classical Item Analysis Results Summary

This subsection presents tables of the classical item analysis results for the 2018–2019 test items.

[Table 8.4](#) presents the summary of the number of items with classical item analysis flags in the 2018–2019 Summative ELPAC. It indicates there were very few items with flags appearing on the 2018–2019 ELPAC summative test.

**Table 8.4 Number of Items with Classical Item Analysis Flag by Domain**

Grade Level or Grade Span	Listening		Speaking		Reading		Writing	
	Total Items	Listening	Total Items	Speaking	Total Items	Reading	Total Items	Writing
Kindergarten	1	20	0	10	0	14	0	8
Grade 1	0	22	1	10	0	20	0	7
Grade 2	1	22	1	13	0	26	0	7
Grade span 3–5	0	22	2	13	0	26	0	6
Grade span 6–8	0	22	1	13	2	26	0	6
Grade span 9–10	0	22	0	13	0	26	0	6
Grade span 11–12	0	22	0	13	3	26	0	6

Detailed results of the item analyses for each item by grade level and grade span are presented in [appendix 8.A](#). The item statistics, including *p*-value, point-biserial correlation, and item type, are included in those tables. The distribution of item scores on each CR item is presented in table 8.A.23 and table 8.A.24.

[Table 8.5](#) presents *p*-value and item-total correlation information by grade level and grade span as well as the number of unique items in each test. Overall, the classical item analysis results were within acceptable ranges. Grade two Speaking, grade two Listening, grade span three through five Speaking, grade one Speaking, and grade span six through eight Speaking results showed maximum *p*-values greater than 0.95 meaning that some items were easier than desired. These items were flagged for additional review.

**Table 8.5 Classical Item Statistics for Each Domain**

Grade Level or Grade Span and Domain	No. of Unique Items	Mean <i>p</i> -value	Minimum <i>p</i> -value	Maximum <i>p</i> -value	Mean Point-Biserial Correlation
Kindergarten Listening	20	0.719	0.247	0.927	0.479
Kindergarten Speaking	10	0.736	0.518	0.885	0.724
Kindergarten Reading	14	0.679	0.432	0.798	0.523
Kindergarten Writing	8	0.759	0.679	0.892	0.780



Table 8.5 (continuation)

Grade Level or Grade Span and Domain	No. of Unique Items	Mean $p$ -value	Minimum $p$ -value	Maximum $p$ -value	Mean Point-Biserial Correlation
Grade 1 Listening	22	0.785	0.563	0.947	0.479
Grade 1 Speaking	10	0.784	0.474	0.962	0.677
Grade 1 Reading	20	0.665	0.425	0.863	0.628
Grade 1 Writing	7	0.732	0.596	0.938	0.697
Grade 2 Listening	22	0.800	0.382	0.960	0.438
Grade 2 Speaking	13	0.828	0.597	0.952	0.670
Grade 2 Reading	26	0.733	0.508	0.945	0.501
Grade 2 Writing	7	0.750	0.651	0.915	0.704
Grade span 3–5 Listening	22	0.720	0.461	0.908	0.416
Grade span 3–5 Speaking	13	0.857	0.643	0.961	0.714
Grade span 3–5 Reading	26	0.600	0.328	0.857	0.449
Grade span 3–5 Writing	6	0.685	0.582	0.750	0.724
Grade span 6–8 Listening	22	0.729	0.393	0.941	0.381
Grade span 6–8 Speaking	13	0.834	0.621	0.966	0.736
Grade span 6–8 Reading	26	0.436	0.213	0.718	0.376
Grade span 6–8 Writing	6	0.706	0.590	0.821	0.702
Grade span 9–10 Listening	22	0.699	0.499	0.859	0.444
Grade span 9–10 Speaking	13	0.789	0.529	0.923	0.805
Grade span 9–10 Reading	26	0.514	0.287	0.768	0.433
Grade span 9–10 Writing	6	0.695	0.572	0.776	0.694
Grade span 11–12 Listening	22	0.648	0.436	0.775	0.399
Grade span 11–12 Speaking	13	0.809	0.651	0.933	0.759
Grade span 11–12 Reading	26	0.522	0.237	0.777	0.393
Grade span 11–12 Writing	6	0.690	0.596	0.757	0.687

#### 8.2.4. Omit Rates

For both MC and CR items, examining item omission is useful for identifying potential problems with test features such as testing time and item or test layout. For the Summative ELPAC, items with omit rates greater than five percent were flagged for further investigation. Omit rates are often useful in determining whether testing times are sufficient, particularly if there is a high rate of items omitted at the end of a test section. In the case of the Summative ELPAC, where speed is not an issue because the test is untimed, high item omit rates may indicate extreme item difficulty.

The Summative ELPAC omit rates tended to be low. For items in the Listening and Speaking domains, the mean omit rates were less than 1 percent and less than 2 percent, respectively. Mean omit rates were highest for the grade one Reading domain (6.52%) and the kindergarten Writing domain (4.71%).

[Table 8.6](#) reports the mean omit rates by grade span and domain.

**Table 8.6 Mean Omit Rates**

Grade Level or Grade Span	Listening Mean Percent of Items Omitted	Speaking Mean Percent of Items Omitted	Reading Mean Percent of Items Omitted	Writing Mean Percent of Items Omitted
Kindergarten	0.93	0.94	0.78	4.71
Grade 1	0.68	0.55	6.52	3.46
Grade 2	0.35	0.96	1.04	1.29
Grade span 3–5	0.26	0.90	0.47	1.09
Grade span 6–8	0.20	1.03	0.51	1.04
Grade span 9–10	0.31	1.97	0.66	1.96
Grade span 11–12	0.31	1.31	0.65	1.53

### 8.3. Item Response Theory Analyses

IRT is built upon the item response function, which describes the probability of a given response as a function of a test-taker’s true ability. IRT can be used to implement item calibrations, link item parameters, scale test scores across different forms or test administrations, evaluate item performance, build an item bank, and assemble test forms.

The two-parameter logistic IRT model was used for the Summative ELPAC item calibrations. In particular, the generalized partial credit (GPC) model (Muraki, 1992) was applied to both dichotomous and polytomous items. The mathematical formula of the GPCM is the following:

$$P_{ih}(\theta_j) = \begin{cases} \frac{\exp(\sum_{v=1}^h Da_i(\theta_j - b_i + d_{iv}))}{1 + \sum_{c=1}^{n_i} \exp(\sum_{v=1}^c Da_i(\theta_j - b_i + d_{iv}))}, & \text{if score } h = 1, 2, \dots, n_i \\ \frac{1}{1 + \sum_{c=1}^{n_i} \exp(\sum_{v=1}^c Da_i(\theta_j - b_i + d_{iv}))}, & \text{if score } h = 0 \end{cases} \tag{8.3}$$

Refer to the [Alternative Text for Equation 8.3](#) for a description of this equation.

where,

- $P_{ih}(\theta_j)$  is the probability of student with proficiency  $\theta_j$  obtaining score  $h$  on item  $i$ ,
- $n_i$  is the maximum number of score points for item  $i$ ,

$a_i$  is the discrimination parameter for item  $i$ ,

$b_i$  is the location parameter for item  $i$ ,

$c$  is the number of nonzero score points for item  $i$ ,

$d_{iv}$  is the category parameter for item  $i$  on score  $v$ , and

$D$  is a scaling constant of 1.7 that makes the logistic model approximate the normal ogive model.

Preequated grade-level or grade-span test forms were administered for the 2018–2019 Summative ELPAC. These preequated test forms were based on calibrations and linking analyses conducted during the spring 2016–2017 field test (refer to appendix 11.A of the *Summative English Language Proficiency Assessments for California Technical Report, 2017–18 Administration* [CDE, 2019] for the field test IRT data). All IRT analyses results for the 2018–2019 preequated operational Summative ELPAC test forms are shown in [appendix 8.B](#).

The overall summary of the IRT  $a$ -value (discrimination) parameter estimates—refer to [equation 8.3](#)—used on the 2018–2019 Summative ELPAC oral language and written language skills tests are shown in [table 8.7](#) and [table 8.8](#). The mean, standard deviation (SD), minimum, and maximum values are presented, in addition to the number of items for each domain.

**Table 8.7 IRT  $a$ -values (Discrimination Parameter) for 2018–2019 Oral Language Tests by Grade Level or Grade Span**

Grade Level or Grade Span	Domain	Number of Items	Mean	Standard Deviation	Minimum	Maximum
Kindergarten	Listening	20	0.53	0.15	0.26	0.78
Kindergarten	Speaking	10	1.04	0.13	0.84	1.22
Grade 1	Listening	22	0.59	0.19	0.21	0.86
Grade 1	Speaking	10	1.01	0.21	0.83	1.49
Grade 2	Listening	22	0.57	0.26	0.17	1.30
Grade 2	Speaking	13	0.76	0.14	0.52	1.01
Grade span 3–5	Listening	22	0.41	0.11	0.19	0.63
Grade span 3–5	Speaking	13	0.72	0.16	0.46	1.02
Grade span 6–8	Listening	22	0.28	0.11	0.06	0.51
Grade span 6–8	Speaking	13	0.73	0.25	0.43	1.24
Grade span 9–10	Listening	22	0.24	0.07	0.09	0.38
Grade span 9–10	Speaking	13	0.65	0.16	0.35	0.89
Grade span 11–12	Listening	22	0.21	0.07	0.09	0.38
Grade span 11–12	Speaking	13	0.59	0.10	0.40	0.71

**Table 8.8 IRT *a*-values (Discrimination Parameter) for 2018–2019 Written Language Tests by Grade Level or Grade Span**

Grade Level or Grade Span	Domain	Number of Items	Mean	Standard Deviation	Minimum	Maximum
Kindergarten	Reading	14	0.52	0.30	0.27	1.31
Kindergarten	Writing	8	1.65	0.42	0.96	2.27
Grade 1	Reading	20	0.86	0.17	0.52	1.22
Grade 1	Writing	7	0.88	0.41	0.57	1.75
Grade 2	Reading	26	0.81	0.27	0.29	1.32
Grade 2	Writing	7	0.85	0.23	0.45	1.20
Grade span 3–5	Reading	26	0.57	0.23	0.15	1.05
Grade span 3–5	Writing	6	0.61	0.11	0.42	0.71
Grade span 6–8	Reading	26	0.37	0.12	0.13	0.59
Grade span 6–8	Writing	6	0.56	0.07	0.47	0.64
Grade span 9–10	Reading	26	0.42	0.14	0.14	0.73
Grade span 9–10	Writing	6	0.46	0.12	0.32	0.61
Grade span 11–12	Reading	26	0.35	0.16	0.04	0.63
Grade span 11–12	Writing	6	0.55	0.10	0.39	0.65

The overall summary of the IRT *b*-value (item difficulty) parameter estimates are shown in [table 8.9](#) and [table 8.10](#) for the Summative ELPAC oral language and written language skills tests. The mean, standard deviation (SD), minimum, and maximum values are presented, in addition to the number of items for each domain.

**Table 8.9 IRT *b*-values (Item Difficulty Parameter) for 2018–2019 Oral Language Tests by Grade Level or Grade Span**

Grade Level or Grade Span	Domain	Number of Items	Mean	Standard Deviation	Minimum	Maximum
Kindergarten	Listening	20	-3.19	1.33	-5.01	-0.23
Kindergarten	Speaking	10	-2.74	0.57	-3.40	-1.81
Grade 1	Listening	22	-3.00	0.98	-5.29	-1.65
Grade 1	Speaking	10	-2.38	0.84	-3.86	-1.03
Grade 2	Listening	22	-2.56	1.30	-4.58	1.30
Grade 2	Speaking	13	-2.29	0.65	-3.66	-1.11
Grade span 3–5	Listening	22	-1.97	1.22	-3.98	0.18
Grade span 3–5	Speaking	13	-2.20	0.68	-3.43	-0.74
Grade span 6–8	Listening	22	-2.14	1.61	-4.21	1.56
Grade span 6–8	Speaking	13	-1.80	0.75	-3.05	-0.42
Grade span 9–10	Listening	22	-2.21	1.39	-4.23	0.29
Grade span 9–10	Speaking	13	-1.47	0.87	-3.29	0.77
Grade span 11–12	Listening	22	-1.63	0.95	-3.47	0.81
Grade span 11–12	Speaking	13	-1.61	0.90	-3.22	-0.23

**Table 8.10 IRT *b*-values (Item Difficulty Parameter) for 2018–2019 Written Language Tests by Grade Level or Grade Span**

Grade Level or Grade Span	Domain	Number of Items	Mean	Standard Deviation	Minimum	Maximum
Kindergarten	Reading	14	-4.32	1.10	-6.46	-1.86
Kindergarten	Writing	8	-3.93	0.48	-4.98	-3.56
Grade 1	Reading	20	-2.45	0.58	-3.33	-1.48
Grade 1	Writing	7	-2.86	0.73	-4.14	-2.20
Grade 2	Reading	26	-2.02	0.63	-3.14	-0.86
Grade 2	Writing	7	-2.27	0.63	-3.21	-1.70
Grade span 3–5	Reading	26	-0.55	1.04	-3.51	1.46
Grade span 3–5	Writing	6	-1.06	0.81	-2.16	-0.27
Grade span 6–8	Reading	26	1.29	1.37	-0.72	5.01
Grade span 6–8	Writing	6	-0.76	0.79	-1.75	0.17
Grade span 9–10	Reading	26	0.81	0.79	-1.08	2.23
Grade span 9–10	Writing	6	-0.87	1.29	-2.54	0.63
Grade span 11–12	Reading	26	1.89	3.71	-0.33	18.94
Grade span 11–12	Writing	6	-0.50	0.92	-1.78	0.72

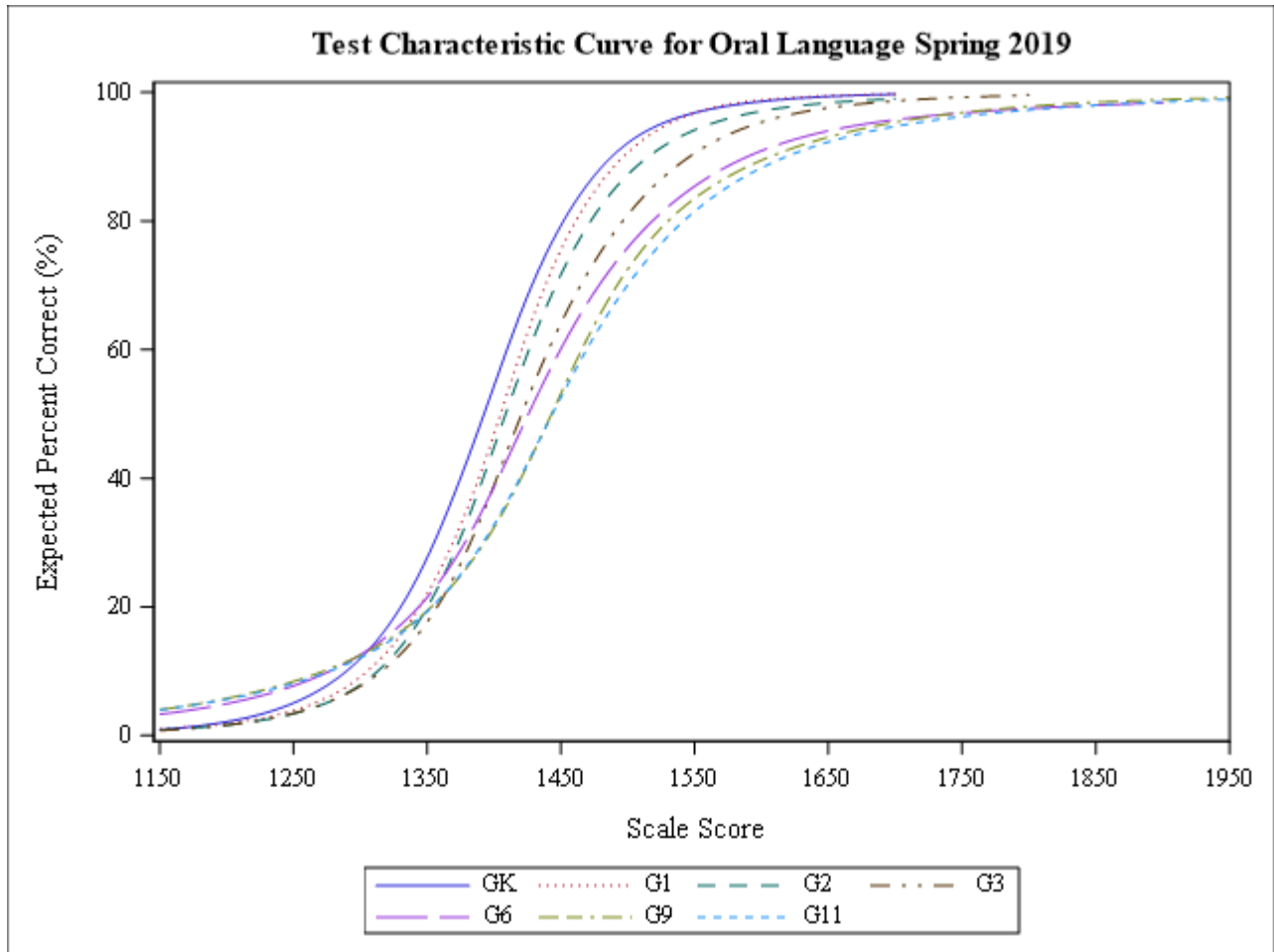
The summary of the IRT *a*-values and *b*-values indicates that the oral tests were relatively easy while test difficulty increased across the grade levels, as shown in [table 8.9](#).

[Table 8.10](#) shows that the Reading and Writing tests were more difficult in high school. All IRT *b*-parameter values were within expected ranges from -6 to +6 except one very easy Reading item in kindergarten and one very difficult Reading item in grade span eleven and twelve.

The distributions of the IRT *a*-values and *b*-values for all operational items appearing on the 2018–2019 test forms are provided in table 8.B.1 through table 8.B.4 in [appendix 8.B](#). In addition, table 8.B.5 through table 8.B.11 provide the IRT discrimination, difficulty, and step parameter estimates at the item level for each grade level or grade span for both oral language and written language skills.

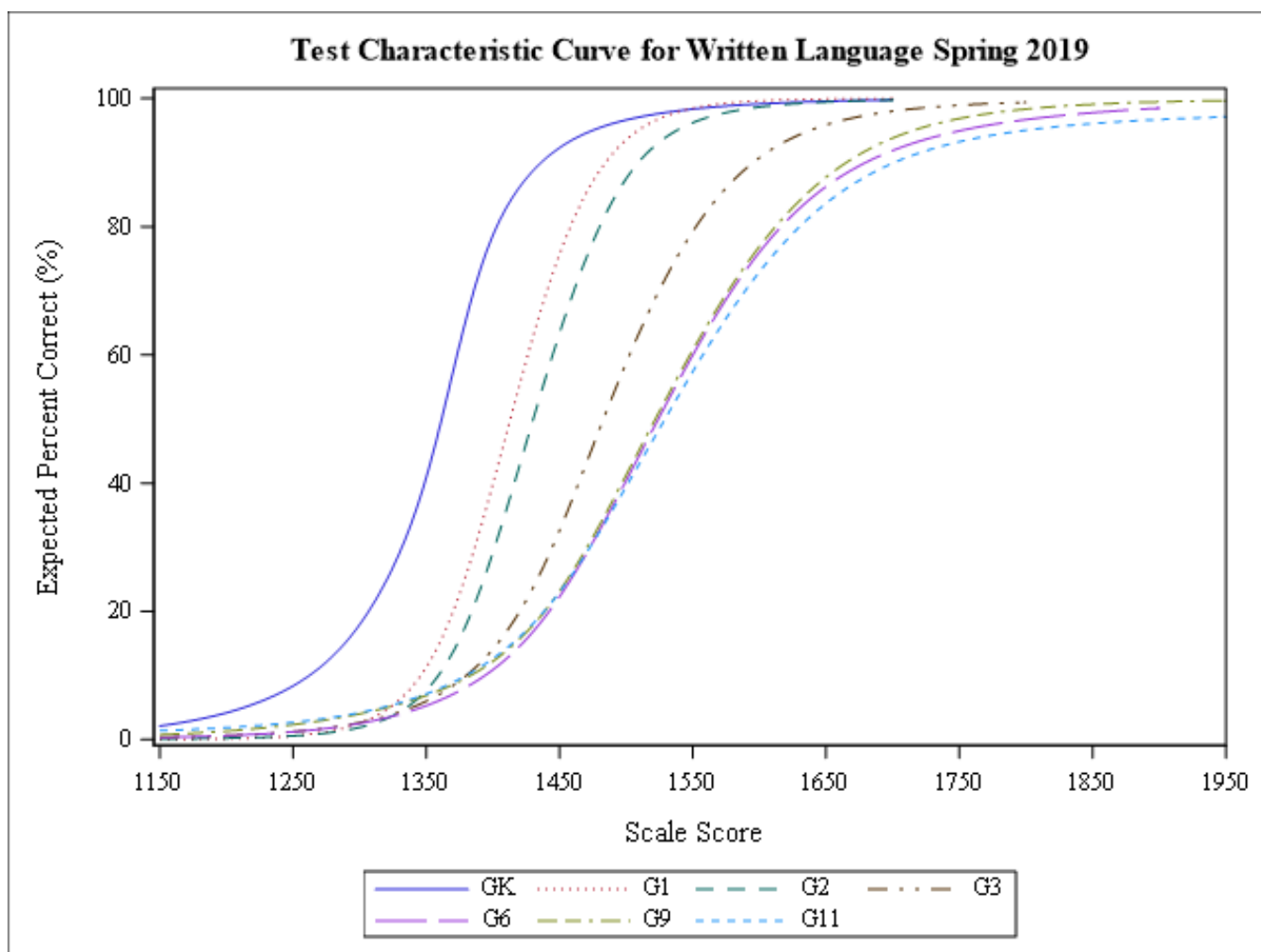
[Figure 8.1](#) and [figure 8.2](#) display the test characteristic curves (TCCs) in the reporting scale metric for the oral and written vertical scales. The expected percentage of correct responses are separated more widely at the lower grade levels, with diminishing amounts of change in the upper grade levels for both the oral and written language scales. The properties of the vertical scale are consistent with the previous ELPAC operational test in 2017–2018 that show increasing difficulty from grade to grade and the within-grade variability (SD) increase from grade to grade.

[Figure 8.1](#) shows the TCC for the oral language composite scores at each grade level or grade span. The curves in [figure 8.1](#) are derived from the data in [table 8.21](#).



**Figure 8.1** Test characteristic curve for the oral language composite

Figure 8.2 shows the TCC for the written language composite scores at each grade level or grade span. The curves in figure 8.2 are derived from the data in table 8.22.



**Figure 8.2 Test characteristic curve for the written language composite**

## 8.4. Differential Item Functioning (DIF)

In addition to the classical item analyses, DIF analyses for male and female gender and ethnicity were conducted for the Summative ELPAC data. The sample size requirements for the DIF analyses were 700 students in the combined focal and reference groups and 300 in the smaller of the two groups. The performance of male and female students was examined for gender DIF, while the performance of Hispanic or Latino students compared to all other ethnicities was examined for ethnicity DIF.

If an item performs differentially across identifiable student groups (e.g., male and female gender and ethnicity) when students are matched on ability, the item may be measuring something other than the intended construct (i.e., possible evidence of bias). It is important, however, to recognize that item performance differences flagged for DIF might be related to actual differences in relevant knowledge or skills between groups (i.e., impact) or statistical Type I error, which might falsely identify DIF in an item. As a result, DIF analysis is used mainly as a statistical tool to identify *potential* item bias. Subsequent reviews by content

experts and bias and sensitivity experts are required to determine the source and meaning of performance differences.

There are many possible reasons for DIF. The wording of an item, for example, may be such that one group interprets the question differently than the other, or the reading demands of an item are such that, although reading is not being measured (e.g., in a mathematics test), reading differences between the groups lead to differential outcomes on the item.

The Summative ELPAC DIF procedures used were the Mantel-Haenszel (MH) procedure (1959) for MC items and the standardized mean difference (SMD) procedure (Dorans, 1989) for CR items.

### 8.4.1. Multiple-choice Items

The Mantel-Haenszel differential item functioning (MH-DIF) statistic was calculated for MC items (Mantel & Haenszel, 1959; Holland & Thayer, 1985). Using the total domain raw score as the criterion score, students in each domain score category in the focal group (e.g., females and non-Hispanic or Latino) are compared with examinees in the same theta score category in the reference group (e.g., males and Hispanic or Latino).

For the MH-DIF, the examinees are split into a focal group, which is typically of prime interest, and a reference group. Each group is then further divided into  $K$  matched ability groups, often on the basis of total test raw score. That is, all examinees obtaining a raw score of 10 represented one matched ability group, for example. Then for an item,  $j$ , the data from the  $k$ th level of reference and focal group members can be arranged as a  $2 \times 2$  table as shown in [table 8.11](#).

**Table 8.11 MH Data Structure**

Group	Item $j$		Total
	Correct	Incorrect	
Reference Group	$A_k$	$B_k$	$n_{Rk}$
Focal Group	$C_k$	$D_k$	$n_{Fk}$
Total Group	$R_k$	$W_k$	$n_{Tk}$

The MH odds ratio estimate,  $\alpha_{MH}$ , for item  $j$  compares the two groups in terms of their odds of answering the item correctly and is given as follows:

$$\alpha_{MH} = \frac{\sum_k \frac{A_k D_k}{n_{Tk}}}{\sum_k \frac{B_k C_k}{n_{Tk}}} \tag{8.4}$$

Refer to the [Alternative Text for Equation 8.4](#) for a description of this equation.

To facilitate the interpretation of MH results, the common odds ratio is frequently transformed to the delta scale using the following formula (Holland & Thayer, 1988):

$$\Delta_{MH} = -2.351 \ln[\alpha_{MH}] \tag{8.5}$$

Refer to the [Alternative Text for Equation 8.5](#) for a description of this equation.



$\Delta_{MH}$  is negative when the item is more difficult for members of the focal group than it is for the comparable members of the reference group.  $\Delta_{MH}$  is positive when the item is more difficult for members of the reference group than it is for the comparable members of the focal group.

MC items are assigned one of three DIF classifications shown in [table 8.12](#).

**Table 8.12 DIF Categories for MC**

DIF category	Criteria
C (large)	$\Delta_{MH}$ is at least 1.5 and is significantly greater than 1.0.
B (moderate)	$\Delta_{MH}$ is at least 1.0 and is significantly greater than 0.0.
A (negligible)	Otherwise

Items with a “C” classification will not be used in the creation of future forms. In these cases, the items were not originally flagged with “C” DIF during field test item analyses but are now flagged with “C” DIF because the underlying student populations changed. During form construction, items with a “B” classification are used only when necessary to meet test specifications.

### 8.4.2. Constructed-Response Items

The standardization DIF (Dorans & Schmitt, 1993; Zwick, Thayer, & Mazzeo, 1997; Dorans, 2013), in conjunction with the Mantel chi-square statistic (Mantel, 1963; Mantel & Haenszel, 1959), is used to identify polytomous items with DIF. The SMD compares the item means of the two groups after adjusting for differences in the distribution of students across the values of the matching variable, using the total domain raw score as the criterion score. The SMD statistic is computed using the following formula:

$$SMD = \frac{\sum_{m=1}^M N_{fm} \times (E_f(Y | X = m) - E_r(Y | X = m))}{\sum_{m=1}^M N_{fm}} = \frac{\sum_{m=1}^M N_{fm} \times D_m}{\sum_{m=1}^M N_{fm}} \tag{8.6}$$

Refer to the [Alternative Text for Equation 8.6](#) for a description of this equation.

where,

$X$  = the criterion score,

$Y$  = the item score,

$M$  = the number of score categories on  $X$ ,

$N_{fm}$  = the number of students in the focal group in score category  $m$ ,

$E_r$  = the expected item score for the reference group,

$E_f$  = the expected item score for the focal group, and

$D_m$  = the expected item score difference between the focal group and the reference group in score category  $m$ .

These indices are indicators of the degree to which members of one group perform better or worse than expected on each CR item.

CR items are also assigned one of three DIF classifications.

A positive SMD value means that, conditional on the criterion score, the focal group has a higher mean item score than the reference group. In contrast, a negative SMD value means that, conditional upon the criterion score, the focal group has a lower mean item score than the reference group.

CR items are assigned one of three DIF classifications shown in [table 8.13](#) dif categories for cr items.

**Table 8.13 DIF Categories for CR Items**

DIF category	Criteria
C (large)	<ul style="list-style-type: none"> <li>Mantel chi-square <math>p</math>-value is <math>&lt; .05</math>; and</li> <li>The absolute value of <math> SMD/SD </math> is <math>&gt; 0.25</math>.</li> </ul>
B (moderate)	<ul style="list-style-type: none"> <li>Mantel chi-square <math>p</math>-value is <math>&lt; 0.05</math>; and</li> <li>The absolute value is <math>0.17 &lt;  SMD/SD  \leq 0.25</math>.</li> </ul>
A (negligible)	<ul style="list-style-type: none"> <li>Mantel chi-square <math>p</math>-value is <math>&lt; .05</math>; or</li> <li>The absolute value of <math> SMD/SD </math> is <math>\leq 0.17</math>.</li> </ul>

**Note:** SMD = standardized mean difference; SD = total group standard deviation of item score

These classifications were defined to be in alignment with the MC classifications in terms of stringency (Zwick, Thayer, and Mazzeo, 1997). Items with a “C” classification will not be used in the creation of future forms, and items with a “B” classification will be used only when necessary to meet test specifications.

### 8.4.3. Classification

Based on the DIF statistics and significance tests, items were classified into three categories and assigned values of A, B, or C. Category A items contained negligible DIF, Category B items exhibited slight-to-moderate DIF, and Category C items possessed moderate-to-large DIF. Items with a Category C will not be used in the creation of future forms. Items with Category “B” DIF will be used only when necessary to meet test blueprints. The classification included an indication of which group had higher performance: “-” indicated that the reference group had higher item performance and “+” indicated that the focal groups’ item performance was higher.

[Table 8.14](#) presents the summary of the DIF analysis and shows that there are no items flagged for Category C DIF by gender.

**Table 8.14 Gender DIF Classification**

Grade Level or Grade Span and Domain	Category C+	Category B+	Category A	Category B-	Category C-	Total Number of Items
Kindergarten Listening	0	0	20	0	0	20
Grade 1 Listening	0	0	22	0	0	22

<b>Grade Level or Grade Span and Domain</b>	<b>Category C+</b>	<b>Category B+</b>	<b>Category A</b>	<b>Category B-</b>	<b>Category C-</b>	<b>Total Number of Items</b>
Grade 2 Listening	0	0	21	1	0	22

Table 8.14 (continuation)

<b>Grade Level or Grade Span and Domain</b>	<b>Category C+</b>	<b>Category B+</b>	<b>Category A</b>	<b>Category B-</b>	<b>Category C-</b>	<b>Total Number of Items</b>
Grade span 3–5 Listening	0	0	22	0	0	22
Grade span 6–8 Listening	0	0	22	0	0	22
Grade span 9–10 Listening	0	0	22	0	0	22
Grade span 11–12 Listening	0	0	21	1	0	22
Kindergarten Speaking	0	0	10	0	0	10
Grade 1 Speaking	0	0	10	0	0	10
Grade 2 Speaking	0	0	13	0	0	13
Grade span 3–5 Speaking	0	0	13	0	0	13
Grade span 6–8 Speaking	0	0	13	0	0	13
Grade span 9–10 Speaking	0	0	13	0	0	13
Grade span 11–12 Speaking	0	0	13	0	0	13
Kindergarten Reading	0	0	14	0	0	14
Grade One Reading	0	0	20	0	0	20
Grade Two Reading	0	0	26	0	0	26
Grade span 3–5 Reading	0	1	25	0	0	26
Grade span 6–8 Reading	0	0	26	0	0	26
Grade span 9–10 Reading	0	0	26	0	0	26
Grade span 11–12 Reading	0	0	25	1	0	26
Kindergarten Writing	0	0	8	0	0	8
Grade 1 Writing	0	0	7	0	0	7
Grade 2 Writing	0	0	7	0	0	7
Grade span 3–5 Writing	0	0	6	0	0	6
Grade span 6–8 Writing	0	0	6	0	0	6
Grade span 9–10 Writing	0	0	6	0	0	6
Grade span 11–12 Writing	0	0	6	0	0	6

[Table 8.15](#) presents the summary of the DIF analysis and shows that there is one item flagged for Category C DIF by ethnicity. The item was reviewed. It was determined that the item asked about the meaning of an English vocabulary word that is similar to the corresponding word in Spanish. As a result, students whose first language is Spanish were probably able to understand the meaning of the word based on the common Latin root, whereas students who did not know Spanish needed to infer the meaning of the word based on the context. This item was kept in the item pool.

**Table 8.15 Ethnicity DIF Classification**

<b>Grade Level or Grade Span and Domain</b>	<b>Category C+</b>	<b>Category B+</b>	<b>Category A</b>	<b>Category B-</b>	<b>Category C-</b>	<b>Total Number of Items</b>
Kindergarten Listening	0	0	20	0	0	20
Grade 1 Listening	0	0	22	0	0	22
Grade 2 Listening	0	0	22	0	0	22
Grade span 3–5 Listening	0	0	22	0	0	22
Grade span 6–8 Listening	0	0	22	0	0	22
Grade span 9–10 Listening	0	0	21	1	0	22
Grade span 11–12 Listening	0	0	21	1	0	22
Kindergarten Speaking	0	0	21	1	0	22
Grade 1 Speaking	0	0	10	0	0	10
Grade 2 Speaking	0	0	10	0	0	10
Grade span 3–5 Speaking	0	0	13	0	0	13
Grade span 6–8 Speaking	0	0	13	0	0	13
Grade span 9–10 Speaking	0	0	13	0	0	13
Grade span 11–12 Speaking	0	0	13	0	0	13
Kindergarten Reading	0	0	14	0	0	14
Grade 1 Reading	0	0	20	0	0	20
Grade 2 Reading	0	0	26	0	0	26
Grade span 3–5 Reading	0	0	26	0	0	26
Grade span 6–8 Reading	0	0	26	0	0	26
Grade span 9–10 Reading	0	0	23	3	0	26
Grade span 11–12 Reading	0	1	24	0	1	26
Kindergarten Writing	0	0	8	0	0	8
Grade 1 Writing	0	0	7	0	0	7
Grade 2 Writing	0	0	7	0	0	7
Grade span 3–5 Writing	0	0	6	0	0	6
Grade span 6–8 Writing	0	0	6	0	0	6
Grade span 9–10 Writing	0	0	6	0	0	6
Grade span 11–12 Writing	0	0	5	1	0	6

## 8.5. Reliability Analyses

The reliability for a particular group of students' test scores estimates the extent to which the scores would remain consistent if those same students were retested with a parallel version of the same test. If the test includes CR items, reliability extends to an evaluation of the extent to which the students' scores would remain consistent if both the items and the scorers were changed.

### 8.5.1. Internal Consistency Reliability

The reliability coefficient cannot, in fact, be computed directly unless the student actually takes two parallel versions of the same test. However, with some reasonable assumptions, reliability can be estimated from the students' responses to a single version of the test.

Like other statistics, the reliability coefficient can vary substantially from one group of students to another. It tends to be larger in groups that are more diverse in the ability measured by the test and smaller in groups that are more homogeneous in the ability measured.

The Summative ELPAC test reliabilities were evaluated for each domain and the composite scores using the coefficient alpha (Cronbach, 1951) index of internal consistency, which is calculated as follows:

$$\hat{\alpha} = \frac{k}{k-1} \left[ 1 - \frac{\sum_{i=1}^k \hat{\sigma}_i^2}{\hat{\sigma}_X^2} \right] \quad (8.7)$$

Refer to the [Alternative Text for Equation 8.7](#) for a description of this equation.

where,

$k$  is the number of items on test form,

$\hat{\sigma}_i^2$  is the estimated variance of item  $i$ , and

$\hat{\sigma}_X^2$  is the estimated total test variance.

The reliability of the overall score was estimated by substituting samples estimates into the following definitional formula for composite reliability (Feldt & Brennan, 1989):

$$\hat{\alpha}_c = 1 - \frac{\sum_j w_j^2 \hat{\sigma}_j^2 (1 - \hat{\alpha}_j)}{\hat{\sigma}_c^2} \quad (8.8)$$

Refer to the [Alternative Text for Equation 8.8](#) for a description of this equation.

where,

$w_j$  is the weight of the  $j$ th component in forming the composite score,

$\hat{\sigma}_j^2$  is the variance of scores on the  $j$ th component,

$\hat{\alpha}_j$  is the reliability of scores on the  $j$ th component, and

$\hat{\sigma}_c^2$  is the variance of the composite score.

[Table 8.16](#) presents reliability coefficients for each domain and composite score of the test by grade level or grade span. Domain reliabilities ranged from .79 to .95 and composite reliabilities ranged from .86 to .94, which is good to excellent internal consistency. For grades three through twelve, the oral language composite had higher reliability coefficients than the written language composite. However, for the lower grade levels, the written language composite had slightly higher reliability coefficients than the oral language composite. The overall test reliability was high, ranging from 0.94 to 0.96.

**Table 8.16 Reliability Coefficient of Domains and Composite Scores**

Grade Level or Grade Span	Listening: Reliability-Coefficient Alpha	Speaking: Reliability-Coefficient Alpha	Reading: Reliability-Coefficient Alpha	Writing: Reliability-Coefficient Alpha	Oral: Reliability-Coefficient Alpha	Written: Reliability-Coefficient Alpha	Overall: Reliability
Kindergarten	0.832	0.884	0.814	0.914	0.904	0.911	0.935
Grade 1	0.855	0.864	0.925	0.842	0.907	0.934	0.952
Grade 2	0.836	0.899	0.897	0.849	0.915	0.924	0.951
Grade span 3–5	0.815	0.918	0.859	0.850	0.916	0.903	0.948
Grade span 6–8	0.795	0.927	0.786	0.844	0.918	0.859	0.938
Grade span 9–10	0.860	0.952	0.853	0.858	0.942	0.896	0.958
Grade span 11–12	0.855	0.951	0.848	0.877	0.940	0.900	0.958

The reliabilities of each domain and composite scores were also examined for various student groups from the population. Table 8.C.1 through table 8.C.7 present the reliabilities for the student groups based on gender, ethnicity, economic status, migrant status, and students receiving special education services status. The reliabilities for various student groups show a similar pattern as overall reliability. Grade span six through eight shows lower reliability than other grade and grade-span tests, especially Reading, which shows the lowest reliability coefficient.

### 8.5.2. Standard Error of Measurement (SEM)

The SEM is a measure of how much students’ scores would vary from the scores they would earn on a perfectly reliable test. If it were possible to compute the error of measurement for each student’s score in a large group of students, these errors of measurement would have a mean of zero. These standard errors of measurement would be an indication of how much the errors of measurement are affecting the students’ scores. The SEM is expressed in the same units as the test score, whether the units are in raw score or scale score metric. In a large group of students, approximately two-thirds of the students will earn scores within one SEM of the scores they would earn on a perfectly reliable test.

The SEM is the square root of the error variance in the scores, that is, the SD of the distribution of the differences between students’ observed scores and their true scores. The SEM is calculated by the following:

$$SEM = SD\sqrt{1-\alpha} \tag{8.9}$$

Refer to the [Alternative Text for Equation 8.9](#) for a description of this equation.

where,

$\alpha$  is the reliability estimated in [equation 8.8](#) for two composite scores of oral and written, and

$SD$  is the standard deviation of the total score of oral score or composite scores (either theta or scale score).

For grades one through twelve, the SEM for the overall score is calculated according to the following formula:

$$SEM_{overall} = \sqrt{.5^2 SEM_{Oral}^2 + .5^2 SEM_{Written}^2} \tag{8.10}$$

Refer to the [Alternative Text for Equation 8.10](#) for a description of this equation.

And for  $K$ ,

$$SEM_{overall} = \sqrt{.7^2 SEM_{Oral}^2 + .3^2 SEM_{Written}^2} \tag{8.11}$$

Refer to the [Alternative Text for Equation 8.11](#) for a description of this equation.

These SEM values are shown in [table 8.17](#). The range of raw score standard errors for the Summative ELPAC were between 1.18 and 3.17 points across all grade levels and domains. In general, this translated into an error band of about two raw score points in most domains. For example, if a student received a raw score of 25 with a standard error of 2.00 points, upon retesting, the student would be expected to obtain a score between 23 and 27 about two-thirds of the time.

**Table 8.17 SEM based on Classical Test Theory**

Grade Level or Grade Span	SEM—Listening Raw Score	SEM—Speaking Raw Score	SEM—Reading Raw Score	SEM—Writing Raw Score	SEM—Oral Raw Score	SEM—Written Raw Score	SEM—Overall Raw Score
Kindergarten	1.636	1.856	1.677	1.192	2.673	2.189	1.983
Grade 1	1.676	1.724	1.618	1.607	2.540	2.401	1.747
Grade 2	1.602	1.852	1.912	1.457	2.616	2.526	1.818
Grade span 3–5	1.859	1.773	2.159	1.518	2.760	2.816	1.972

Table 8.17 (continuation)

Grade Level or Grade Span	SEM—Listening Raw Score	SEM—Speaking Raw Score	SEM—Reading Raw Score	SEM—Writing Raw Score	SEM—Oral Raw Score	SEM—Written Raw Score	SEM—Overall Raw Score
Grade span 6–8	1.851	1.864	2.287	1.510	2.871	2.943	2.056
Grade span 9–10	1.900	1.941	2.243	1.550	3.103	2.941	2.137
Grade span 11–12	2.006	1.905	2.251	1.504	3.129	2.946	2.149

It is important to remember that assessments are not perfectly reliable and only offer an estimate of what the student is capable of in a specified domain. As [table 8.18](#) shows, the SEM scale score values for oral and written language skills averaged about 23 scale score points and 17 scale score points for overall.

**Table 8.18 SEM Based on Scale Score**

Grade Level or Grade Span	SEM—Oral	SEM—Written	SEM—Overall
Kindergarten	20.684	26.531	16.522
Grade 1	19.513	20.174	14.033
Grade 2	19.908	20.008	14.112
Grade span 3–5	22.022	20.647	15.094
Grade span 6–8	27.551	27.515	19.469
Grade span 9–10	30.086	28.471	20.711
Grade span 11–12	30.995	35.057	23.397

### 8.5.3. Conditional Standard Error of Measurement (CSEM)

Classical test theory assumes that the standard error of a test score is constant throughout the score range. While the assumption is probably reasonable in the midscore ranges, it is less reasonable at the extremes of the score distribution. IRT expands the concept by providing estimates of the standard error at each score point on the distribution.

The IRT, or conditional SEM (CSEM) for scale scores, is defined as

$$CSEM(SS) = a \frac{1}{\sqrt{I(\hat{\theta})}} \tag{8.12}$$

Refer to the [Alternative Text for Equation 8.12](#) for a description of this equation.

where,

$$SS = a \times \theta + b,$$

CSEM(SS) is the conditional standard of measurement on the scale score scale, and



$a$  and  $b$  are the scaling constants (the slope and intercept) needed to transform theta to the scale score metric.

$I(\hat{\theta})$  is the test information function at ability level  $\hat{\theta}$ . For student  $j$ , test information is calculated as

$$I(\theta_j) = \sum_{i=1}^n I_i(\theta_j) \quad (8.13)$$

Refer to the [Alternative Text for Equation 8.13](#) for a description of this equation.

where,

$I_i(\theta_j)$  is the item information of item  $i$  for student  $j$ .

Item information is calculated as

$$I_i(\theta_j) = [s_{i2}(\theta_j) - s_i^2(\theta_j)] \quad (8.14)$$

Refer to the [Alternative Text for Equation 8.14](#) for a description of this equation.

where,

$S_i(\theta_j)$  is the expected item score for item  $i$  on a theta score  $\theta_j$  calculated as

$$s_i(\theta_j) = \sum_{h=0}^{n_i} h p_{ih}(\theta_j) \quad (8.15)$$

Refer to the [Alternative Text for Equation 8.15](#) for a description of this equation.

and

$$s_{i2}(\theta_j) = \sum_{h=0}^{n_i} h^2 p_{ih}(\theta_j) \quad (8.16)$$

Refer to the [Alternative Text for Equation 8.16](#) for a description of this equation.

where,

$P_{ih}(\theta_j)$  is the probability of an examinee with  $\theta_j$  getting score  $h$  on item  $i$ , the computation of which is shown in [equation 8.3](#), and

$n_i$  is the maximum score.

The IRT's version of an SEM has an inverse normal distribution in which SEM values decrease as scores move toward the center of the range. CSEM values are reported as part of the raw-score-to-scale-score conversion tables presented in [appendix 8.D](#) for the oral and written language skills.

CSEMs vary across the scale, and are typically smaller toward the center of the scale where more items are located and typically larger at the extreme ends of the scale. For most grades and grade spans, the lowest values of CSEM are between the proficiency levels one and two; the threshold scores between proficiency levels one and two are toward the middle of the scale score ranges. The CSEMs for threshold scores between proficiency levels three and four are somewhat larger.

### 8.5.4. Writing Score Reliability

Rater consistency is critical to the scores of ELPAC writing items and their score interpretations. When two trained raters independently assign the same score (or rating) to an item response, there is evidence that the scoring standard is being applied consistently. Double scoring substantially increases the reliability of the scoring process. When used to monitor and evaluate the accuracy of rating, 10 percent of the responses are typically rated twice by two independent raters. Interrater reliability is evaluated empirically by computing the percentage of exact agreement between two raters.

Evidence that the raters' scores are consistent helps to support the inference that the scores have the intended meaning. The exact agreement data collected is used to evaluate interrater agreement. [Table 8.19](#) presents the average percent exact agreement, adjacent, and discrepant by grade level or grade span. The average percent exact agreement, adjacent, and discrepant by the possible maximum score point indicates that the average ratings meet or exceed the criteria described in [table 7.2](#).

**Table 8.19 Summary Rater Agreement by Possible Maximum Score Points and Grade Level or Grade Span**

Grade Level or Grade Span	Number of Score Points	Average of Percent Exact	Average of Percent Adjacent	Average of Percent Discrepant
Kindergarten	All Writing Items	98.73	1.25	0.03
Kindergarten	1-pt score Items	99.48	0.53	0.00
Kindergarten	2-pt score items	97.98	1.98	0.05
Grade 1	All Writing Items	92.47	7.21	0.29
Grade 1	1-pt score Items	99.30	0.70	0.00
Grade 1	2-pt score items	95.60	4.30	0.05
Grade 1	3-pt score Items	85.83	13.50	0.63
Grade 2	All Writing Items	93.17	6.69	0.13
Grade 2	1-pt score Items	99.55	0.45	0.00
Grade 2	2-pt score items	97.20	2.75	0.00
Grade 2	3-pt score Items	86.23	13.47	0.30
Grade span 3–5	All Writing Items	88.37	11.52	0.12
Grade span 3–5	2-pt score items	94.03	5.90	0.03
Grade span 3–5	3-pt score items	88.00	12.00	0.00
Grade span 3–5	4-pt score items	80.05	19.70	0.30
Grade span 6–8	All Writing Items	87.18	12.60	0.20
Grade span 6–8	2-pt score items	93.60	6.33	0.07
Grade span 6–8	3-pt score items	86.80	13.10	0.10
Grade span 6–8	4-pt score items	77.75	21.75	0.45

Table 8.19 (continuation)

Grade Level or Grade Span	Number of Score Points	Average of Percent Exact	Average of Percent Adjacent	Average of Percent Discrepant
Grade span 9–10	All Writing Items	87.68	12.10	0.22
Grade span 9–10	2-pt score items	94.80	5.20	0.00
Grade span 9–10	3-pt score items	90.30	9.50	0.20
Grade span 9–10	4-pt score items	75.70	23.75	0.55
Grade span 11–12	All Writing Items	88.72	11.12	0.17
Grade span 11–12	2-pt score items	94.77	5.20	0.03
Grade span 11–12	3-pt score items	87.40	12.30	0.30
Grade span 11–12	4-pt score items	80.30	19.40	0.30

Table 8.E.1 in [appendix 8.E](#) provides interrater agreement statistics for each Writing domain item including percent adjacent and percent discrepant on the 2018–2019 Summative ELPAC.

### 8.5.5. Decision Classification Analyses

While the reliabilities of performance-level classifications, which are criterion referenced, are related to the reliabilities of the test scores on which they are based, they are not identical. Glaser (1963) was among the first to draw attention to this distinction, and Feldt and Brennan (1989) extensively reviewed the topic. While test reliability evaluates the consistency of test scores, decision classification reliability evaluates the consistency of classification.

Consistency in classification represents how well two versions of an assessment with equal difficulty agree in the classification of students (Livingston & Lewis, 1995). This is estimated by using actual response data and total test reliability from an administered form of the assessment from which two parallel versions of the assessment are statistically modeled and classifications compared. Decision consistency, then, is the extent to which the test classification of examinees into mastery levels agrees with classifications based on a hypothetical parallel test. The examinees' scores on the second form are statistically modeled.

Note that the values of all indices depend on several factors, such as the reliability of the actual test form, distribution of scores, number of threshold scores, and location of each threshold score. The probability of a correct classification is the probability that the classification the examinee received is consistent with the classification that the examinee would have received on a parallel form. This is akin to the exact agreement rate in interrater reliability. The expectation is that this probability would be high.

Decision accuracy is the extent to which the test's classification of examinees into levels agrees with the examinees' true classification. The examinees' true scores—and, therefore, true classification—are not known but can be modeled. Consistency and accuracy are important to consider together. The probability of accuracy represents the agreement between the observed classification based on the actual test form and true classification, given the modeled form.

Commonly used indices for decision consistency and accuracy include (a) decision consistency and accuracy at each threshold score, (b) overall decision consistency and accuracy across all threshold scores, and (c) coefficient kappa.

Cohen’s kappa (Fleiss and Cohen, 1973) represents the agreement of the classifications between two parallel versions of the same test, taking into account the probability of a correct classification by chance. It measures how the test contributes to the classification of examinees over and above chance classifications. In general, the value of kappa is lower than the value of the probability of correct classification because the probability of a correct classification by chance is larger than zero.

The methodology used for estimating the reliability of classification decisions described in Livingston and Lewis (1995) is implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.14).

Overall decision accuracy and consistency—that is, classification across all threshold scores—are reported in [table 8.20](#). Decision accuracy ranged from 0.726 to 0.785 for the oral composite and from 0.704 to 0.799 for the Written composite. Decision consistency for the oral and written composites were 0.666 to 0.706 and 0.612 to 0.722, respectively. Values are consistent with those observed in previous ELPAC administrations.

**Table 8.20 Classification Accuracy and Consistency for Reported Composite Scores**

Grade Level or Grade Span	Oral Accuracy	Oral Consistency	Written Accuracy	Written Consistency
Kindergarten	0.767	0.680	0.785	0.706
Grade 1	0.764	0.678	0.799	0.722
Grade 2	0.785	0.706	0.797	0.718
Grade 3	0.745	0.666	0.757	0.668
Grade 4	0.765	0.701	0.759	0.671
Grade 5	0.767	0.706	0.770	0.691
Grade 6	0.735	0.673	0.704	0.615
Grade 7	0.735	0.679	0.706	0.612
Grade 8	0.731	0.677	0.732	0.639
Grade 9	0.760	0.710	0.777	0.695
Grade 10	0.778	0.721	0.770	0.686
Grade 11	0.749	0.690	0.762	0.677
Grade 12	0.726	0.692	0.779	0.708

Results of classification consistency and accuracy are reported in [appendix 8.F](#) by grade level or grade span and composite language skills.

## References

- California Department of Education. (2019). *Summative English Language Proficiency Assessments for California technical report, 2017–18 administration: Chapter 11 appendix*. Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/ta/tg/ep/documents/elpac18techrptappdx11.pdf>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Dorans, N. J. (2013). *ETS contributions to the quantitative assessment of item, test, and score fairness*. (ETS Scientific and Policy Contributions Series SPC-13-04). Princeton, NJ: Educational Testing Service.
- Dorans, N. J., & Schmitt, A. P. (1993). Constructed response and differential item functioning: A pragmatic approach. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 135–65). Hillsdale, NH: Lawrence Erlbaum Associates, Inc.
- Dorans, N. J. (1989). Two new approaches to assessing differential item functioning: Standardization and the Mantel-Haenszel method. *Applied Measurement in Education*, 2, 217–33.
- Feldt, L. S. & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd edition) (pp. 105–46). New York: Macmillan.
- Fleiss, J. L. & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613–19.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist*, 18, 519–32
- Holland, P. W., & Thayer, D. T. (1985). *An alternative definition of the ETS delta scale of item difficulty* (Research Report 85–43). Princeton, NJ: Educational Testing Service.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 129–45). Hillsdale, NJ: Erlbaum.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal of Educational Measurement*, 32, 179–97.
- Mantel, N. (1963). Chi-Square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58, 690–700.
- Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analyses of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–48.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16(2): 159–76.
- Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education*, 10(4), 321–44.

## Accessibility Information

### Alternative Text for Equation 8.1

The p-value for item i is equal to the sum of the ith item scores across all j students divided by the total number of students who were presented with item i.

### Alternative Text for Equation 8.2

The p-value for item i is equal to the sum of the ith item scores across all j students divided by product of the total number of students who were presented with item i and the maximum score available for item i.

### Alternative Text for Equation 8.3

If score h equals 1, 2, up to n sub i, then  $P_{ih}$  is equal to fraction where the numerator has the exponential of the summation of v from 1 to h of  $D \times a_{iv}$  times  $\theta_j - b_{iv}$ . The denominator is 1 plus the summation of c from 1 to n sub l of the exponential of sum of v from 1 to c of  $D \times a_{iv}$  times  $\theta_j - b_{iv}$ .

If score h equals 0, then  $P_{ih}$  is equal to fraction where the numerator is 1. The denominator is 1 plus the summation of c from 1 to n sub l of the exponential of sum of v from 1 to c of  $D \times a_{iv}$  times  $\theta_j - b_{iv}$ .

### Alternative Text for Equation 8.4

$\alpha_{MH}$  is equal to a fraction where the numerator is the sum over all k of a fraction where the numerator is  $A_{kD}$  and the denominator is  $n_{Tk}$ . The denominator is equal to a fraction where the numerator is the sum over all k of a fraction where the numerator is  $B_{kC}$  and the denominator is  $n_{Tk}$ .

### Alternative Text for Equation 8.5

$MH D - DIF$  equals negative 2.351 times the natural logarithm  $\alpha_{MH}$ .

### Alternative Text for Equation 8.6

SMD is equal to fraction where numerator is equal to the summation of m from 1 to M of  $N_{fm}$  times  $E_{fY}$  given  $X = m - E_{rY}$  given  $X = m$ . The denominator is the summation of m from 1 to M of  $N_{fm}$ . This is equal to fraction where the numerator is the summation of m from 1 to M of  $N_{fm}$  times  $D_M$ . The denominator is the summation of m from 1 to M of  $N_{fm}$ .

### Alternative Text for Equation 8.7

$\hat{\alpha}$  equals fraction with numerator K and denominator  $K - 1$  times  $1 - \frac{\sum_{l=1}^K \sigma_l^2}{\hat{\sigma}_X^2}$ .

### Alternative Text for Equation 8.8

$\hat{\alpha}_c$  equals 1 minus fraction with numerator sum of j of  $w_j^2$  times  $\hat{\sigma}_j$  times  $1 - \hat{\alpha}_j$  and denominator  $\hat{\sigma}_c$ .

**Alternative Text for Equation 8.9**

SEM equals total score standard deviation multiplied by the square root of 1 minus alpha where alpha is the reliability corresponding to the two composite scores.

**Alternative Text for Equation 8.10**

Overall SEM is equal to square root of the sum of the weighted composite of the squared SEMs. The weighted composite is 0.5 squared times the square of the oral SEM plus 0.5 squared times the square of the written SEM.

**Alternative Text for Equation 8.11**

Overall SEM is equal to square root of the sum of the weighted composite of the squared SEMs. The weighted composite is 0.7 squared times the square of the oral SEM plus 0.3 squared times the square of the written SEM.

**Alternative Text for Equation 8.12**

CSEM of SS equals 1 times a divided by the square root of I of theta hat.

**Alternative Text for Equation 8.13**

I of theta sub j equals the sum from I equals 1 to n of I sub I of theta sub j.

**Alternative Text for Equation 8.14**

I sub I of theta sub j equals open bracket s sub i<sup>2</sup> open parenthesis theta sub j closed parenthesis min s sub I squared open parenthesis theta sub j closed parenthesis closed bracket.

**Alternative Text for Equation 8.15**

s sub I open parenthesis theta sub j closed parenthesis is equal to summation from h equal zero to n sub i of h times p sub i h open parenthesis theta sub j closed parenthesis.

**Alternative Text for Equation 8.16**

s sub i<sup>2</sup> open parenthesis theta sub j closed parenthesis is equal to summation from h equal zero to n sub i of h squared times p sub i h open parenthesis theta sub j closed parenthesis.

## Accessibility Reference: Table Data for Test Characteristic Curves

**Table 8.21 TCC Data for the Oral Language Composite (Figure 8.1)**

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1150	0.918	0.988	0.866	0.781	3.309	4.006	3.991
1151	0.933	1.000	0.877	0.792	3.334	4.034	4.018
1152	0.949	1.013	0.888	0.804	3.360	4.063	4.046
1153	0.965	1.025	0.899	0.815	3.385	4.093	4.073
1154	0.981	1.038	0.910	0.827	3.411	4.122	4.101
1155	0.997	1.050	0.921	0.839	3.437	4.152	4.129
1156	1.014	1.063	0.933	0.851	3.464	4.182	4.157
1157	1.031	1.077	0.944	0.863	3.491	4.212	4.185
1158	1.049	1.090	0.956	0.876	3.518	4.243	4.214
1159	1.066	1.103	0.968	0.888	3.545	4.273	4.243
1160	1.084	1.117	0.980	0.901	3.572	4.304	4.272
1161	1.103	1.131	0.993	0.914	3.600	4.335	4.301
1162	1.121	1.145	1.005	0.928	3.628	4.367	4.331
1163	1.140	1.160	1.018	0.941	3.657	4.399	4.360
1164	1.160	1.174	1.031	0.955	3.686	4.430	4.390
1165	1.179	1.189	1.044	0.969	3.715	4.463	4.420
1166	1.199	1.204	1.057	0.983	3.744	4.495	4.451
1167	1.219	1.219	1.070	0.997	3.773	4.528	4.481
1168	1.240	1.234	1.084	1.012	3.803	4.561	4.512
1169	1.261	1.250	1.098	1.026	3.834	4.594	4.543
1170	1.283	1.266	1.112	1.041	3.864	4.628	4.575
1171	1.304	1.282	1.126	1.057	3.895	4.661	4.606
1172	1.327	1.298	1.140	1.072	3.926	4.695	4.638
1173	1.349	1.315	1.155	1.088	3.958	4.730	4.670
1174	1.372	1.332	1.170	1.104	3.989	4.764	4.702
1175	1.396	1.349	1.185	1.120	4.022	4.799	4.735
1176	1.419	1.366	1.200	1.136	4.054	4.834	4.768
1177	1.444	1.384	1.216	1.153	4.087	4.870	4.801
1178	1.468	1.402	1.231	1.170	4.120	4.905	4.834
1179	1.493	1.420	1.247	1.187	4.154	4.941	4.868
1180	1.519	1.439	1.264	1.205	4.188	4.978	4.901
1181	1.545	1.457	1.280	1.223	4.222	5.014	4.935
1182	1.571	1.476	1.297	1.241	4.256	5.051	4.970
1183	1.598	1.496	1.314	1.259	4.291	5.088	5.004
1184	1.626	1.515	1.331	1.278	4.327	5.126	5.039
1185	1.654	1.535	1.349	1.297	4.362	5.164	5.074



Table 8.21 (continuation one)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1186	1.682	1.555	1.367	1.316	4.398	5.202	5.110
1187	1.711	1.576	1.385	1.335	4.435	5.240	5.146
1188	1.740	1.597	1.403	1.355	4.472	5.279	5.182
1189	1.770	1.618	1.422	1.375	4.509	5.318	5.218
1190	1.801	1.640	1.440	1.396	4.546	5.357	5.254
1191	1.832	1.661	1.460	1.417	4.584	5.397	5.291
1192	1.863	1.684	1.479	1.438	4.623	5.437	5.328
1193	1.895	1.706	1.499	1.459	4.662	5.477	5.366
1194	1.928	1.729	1.519	1.481	4.701	5.517	5.403
1195	1.961	1.753	1.540	1.503	4.740	5.558	5.442
1196	1.995	1.776	1.560	1.526	4.780	5.600	5.480
1197	2.029	1.800	1.581	1.549	4.821	5.641	5.518
1198	2.064	1.825	1.603	1.572	4.862	5.683	5.557
1199	2.100	1.850	1.625	1.595	4.903	5.725	5.597
1200	2.136	1.875	1.647	1.619	4.945	5.768	5.636
1201	2.173	1.901	1.669	1.644	4.987	5.811	5.676
1202	2.211	1.927	1.692	1.668	5.030	5.854	5.716
1203	2.249	1.953	1.715	1.694	5.073	5.898	5.757
1204	2.288	1.980	1.739	1.719	5.117	5.942	5.798
1205	2.327	2.008	1.763	1.745	5.161	5.986	5.839
1206	2.368	2.036	1.787	1.771	5.205	6.031	5.881
1207	2.409	2.064	1.812	1.798	5.250	6.076	5.922
1208	2.450	2.093	1.837	1.825	5.296	6.121	5.965
1209	2.493	2.122	1.863	1.853	5.342	6.167	6.007
1210	2.536	2.152	1.889	1.881	5.388	6.213	6.050
1211	2.580	2.182	1.915	1.910	5.435	6.260	6.093
1212	2.625	2.213	1.942	1.939	5.483	6.307	6.137
1213	2.670	2.245	1.969	1.968	5.531	6.354	6.181
1214	2.717	2.277	1.997	1.998	5.579	6.402	6.225
1215	2.764	2.309	2.026	2.028	5.628	6.450	6.270
1216	2.812	2.342	2.054	2.059	5.678	6.498	6.315
1217	2.861	2.376	2.083	2.091	5.728	6.547	6.361
1218	2.910	2.410	2.113	2.123	5.778	6.597	6.407
1219	2.961	2.445	2.143	2.155	5.830	6.646	6.453
1220	3.012	2.480	2.174	2.188	5.881	6.696	6.500
1221	3.065	2.516	2.205	2.221	5.934	6.747	6.547
1222	3.118	2.553	2.237	2.255	5.986	6.798	6.594
1223	3.172	2.590	2.270	2.290	6.040	6.849	6.642
1224	3.227	2.628	2.302	2.325	6.094	6.901	6.690

Table 8.21 (continuation two)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1225	3.284	2.667	2.336	2.361	6.149	6.953	6.739
1226	3.341	2.706	2.370	2.397	6.204	7.006	6.788
1227	3.399	2.746	2.405	2.434	6.259	7.059	6.838
1228	3.458	2.787	2.440	2.471	6.316	7.112	6.888
1229	3.518	2.828	2.476	2.509	6.373	7.166	6.938
1230	3.579	2.871	2.512	2.548	6.430	7.221	6.989
1231	3.642	2.914	2.549	2.587	6.489	7.275	7.040
1232	3.705	2.957	2.587	2.627	6.548	7.331	7.092
1233	3.770	3.002	2.626	2.668	6.607	7.386	7.144
1234	3.835	3.047	2.665	2.709	6.667	7.443	7.197
1235	3.902	3.094	2.705	2.751	6.728	7.499	7.250
1236	3.970	3.141	2.745	2.794	6.790	7.556	7.303
1237	4.039	3.189	2.787	2.837	6.852	7.614	7.358
1238	4.110	3.237	2.829	2.881	6.915	7.672	7.412
1239	4.181	3.287	2.872	2.926	6.978	7.731	7.467
1240	4.254	3.338	2.915	2.971	7.042	7.790	7.523
1241	4.328	3.389	2.960	3.018	7.107	7.849	7.579
1242	4.404	3.442	3.005	3.065	7.173	7.909	7.635
1243	4.480	3.496	3.051	3.112	7.239	7.970	7.692
1244	4.558	3.550	3.098	3.161	7.307	8.031	7.750
1245	4.638	3.606	3.146	3.210	7.374	8.093	7.808
1246	4.718	3.662	3.194	3.261	7.443	8.155	7.867
1247	4.801	3.720	3.244	3.312	7.512	8.218	7.926
1248	4.884	3.779	3.295	3.363	7.582	8.281	7.986
1249	4.969	3.839	3.346	3.416	7.653	8.345	8.046
1250	5.056	3.900	3.398	3.470	7.725	8.409	8.107
1251	5.144	3.962	3.452	3.524	7.797	8.474	8.169
1252	5.234	4.025	3.506	3.579	7.871	8.539	8.231
1253	5.325	4.090	3.562	3.636	7.945	8.605	8.294
1254	5.417	4.156	3.618	3.693	8.020	8.672	8.357
1255	5.512	4.223	3.676	3.751	8.095	8.739	8.421
1256	5.608	4.291	3.734	3.810	8.172	8.806	8.485
1257	5.705	4.361	3.794	3.870	8.249	8.875	8.550
1258	5.805	4.432	3.855	3.931	8.328	8.943	8.616
1259	5.906	4.505	3.917	3.994	8.407	9.013	8.683
1260	6.009	4.579	3.981	4.057	8.487	9.083	8.750
1261	6.113	4.654	4.045	4.121	8.568	9.154	8.817

Table 8.21 (continuation three)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1262	6.219	4.731	4.111	4.186	8.649	9.225	8.886
1263	6.328	4.809	4.178	4.253	8.732	9.297	8.955
1264	6.438	4.889	4.247	4.320	8.816	9.370	9.025
1265	6.550	4.970	4.316	4.389	8.900	9.443	9.095
1266	6.664	5.053	4.388	4.459	8.986	9.517	9.166
1267	6.779	5.138	4.460	4.530	9.072	9.591	9.238
1268	6.897	5.224	4.534	4.602	9.160	9.666	9.311
1269	7.017	5.312	4.610	4.675	9.248	9.742	9.384
1270	7.139	5.402	4.687	4.750	9.338	9.819	9.458
1271	7.263	5.493	4.765	4.826	9.428	9.896	9.533
1272	7.389	5.587	4.845	4.903	9.520	9.974	9.609
1273	7.518	5.682	4.927	4.981	9.612	10.053	9.685
1274	7.648	5.779	5.010	5.061	9.706	10.132	9.763
1275	7.781	5.878	5.095	5.142	9.800	10.212	9.841
1276	7.916	5.978	5.182	5.225	9.896	10.293	9.920
1277	8.053	6.081	5.270	5.309	9.993	10.375	9.999
1278	8.193	6.186	5.360	5.394	10.091	10.457	10.080
1279	8.335	6.293	5.452	5.481	10.190	10.541	10.161
1280	8.480	6.402	5.546	5.569	10.290	10.625	10.244
1281	8.627	6.513	5.642	5.659	10.391	10.709	10.327
1282	8.776	6.627	5.740	5.750	10.494	10.795	10.411
1283	8.929	6.742	5.839	5.843	10.597	10.881	10.496
1284	9.083	6.860	5.941	5.938	10.702	10.969	10.582
1285	9.241	6.981	6.045	6.034	10.808	11.057	10.669
1286	9.401	7.103	6.151	6.132	10.915	11.146	10.757
1287	9.564	7.228	6.259	6.231	11.024	11.235	10.846
1288	9.729	7.356	6.369	6.332	11.134	11.326	10.936
1289	9.898	7.486	6.482	6.435	11.245	11.417	11.027
1290	10.069	7.619	6.597	6.540	11.357	11.510	11.118
1291	10.243	7.754	6.714	6.646	11.471	11.603	11.211
1292	10.421	7.892	6.834	6.755	11.586	11.698	11.305
1293	10.601	8.032	6.956	6.865	11.702	11.793	11.400
1294	10.784	8.175	7.081	6.977	11.820	11.889	11.497
1295	10.971	8.321	7.208	7.091	11.939	11.986	11.594
1296	11.161	8.470	7.338	7.207	12.059	12.084	11.692
1297	11.354	8.622	7.471	7.325	12.181	12.184	11.792

Table 8.21 (continuation four)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1298	11.550	8.777	7.606	7.445	12.305	12.284	11.892
1299	11.749	8.935	7.744	7.567	12.430	12.385	11.994
1300	11.952	9.095	7.885	7.692	12.556	12.487	12.097
1301	12.159	9.259	8.029	7.818	12.684	12.591	12.202
1302	12.369	9.426	8.176	7.947	12.814	12.695	12.307
1303	12.582	9.596	8.325	8.078	12.945	12.800	12.414
1304	12.799	9.770	8.478	8.211	13.077	12.907	12.522
1305	13.020	9.946	8.634	8.346	13.212	13.015	12.631
1306	13.245	10.126	8.793	8.484	13.348	13.124	12.742
1307	13.473	10.310	8.956	8.625	13.485	13.234	12.854
1308	13.705	10.497	9.122	8.767	13.625	13.345	12.967
1309	13.941	10.687	9.291	8.913	13.766	13.458	13.082
1310	14.182	10.881	9.464	9.060	13.908	13.572	13.198
1311	14.426	11.078	9.640	9.211	14.053	13.687	13.316
1312	14.674	11.279	9.819	9.364	14.199	13.803	13.435
1313	14.927	11.484	10.003	9.519	14.348	13.920	13.555
1314	15.183	11.692	10.190	9.678	14.498	14.039	13.677
1315	15.444	11.905	10.381	9.839	14.650	14.160	13.801
1316	15.709	12.121	10.575	10.003	14.804	14.281	13.926
1317	15.979	12.341	10.774	10.170	14.960	14.404	14.053
1318	16.253	12.564	10.977	10.340	15.118	14.529	14.181
1319	16.532	12.792	11.183	10.513	15.278	14.655	14.311
1320	16.815	13.024	11.394	10.689	15.440	14.782	14.442
1321	17.103	13.259	11.609	10.867	15.604	14.911	14.575
1322	17.396	13.499	11.828	11.050	15.771	15.041	14.710
1323	17.693	13.743	12.051	11.235	15.939	15.173	14.847
1324	17.996	13.991	12.279	11.423	16.110	15.306	14.985
1325	18.303	14.244	12.511	11.615	16.283	15.441	15.125
1326	18.615	14.500	12.747	11.810	16.459	15.578	15.267
1327	18.931	14.761	12.989	12.009	16.637	15.716	15.411
1328	19.253	15.026	13.234	12.211	16.817	15.856	15.556
1329	19.580	15.295	13.485	12.416	16.999	15.997	15.704
1330	19.912	15.569	13.740	12.625	17.184	16.141	15.853
1331	20.250	15.847	14.000	12.838	17.372	16.286	16.004
1332	20.592	16.130	14.265	13.054	17.562	16.432	16.158
1333	20.939	16.417	14.534	13.275	17.755	16.581	16.313

Table 8.21 (continuation five)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1334	21.292	16.708	14.809	13.499	17.951	16.731	16.470
1335	21.650	17.004	15.088	13.726	18.149	16.884	16.630
1336	22.014	17.305	15.373	13.958	18.350	17.038	16.791
1337	22.382	17.610	15.663	14.194	18.554	17.194	16.955
1338	22.756	17.920	15.958	14.434	18.761	17.352	17.120
1339	23.135	18.234	16.258	14.678	18.970	17.512	17.288
1340	23.520	18.553	16.564	14.926	19.183	17.674	17.458
1341	23.910	18.877	16.875	15.178	19.398	17.838	17.631
1342	24.305	19.205	17.191	15.435	19.617	18.004	17.805
1343	24.706	19.538	17.513	15.696	19.839	18.172	17.982
1344	25.112	19.876	17.840	15.961	20.064	18.343	18.161
1345	25.523	20.218	18.173	16.231	20.292	18.515	18.343
1346	25.940	20.565	18.511	16.505	20.523	18.690	18.527
1347	26.362	20.917	18.855	16.784	20.758	18.867	18.714
1348	26.789	21.274	19.204	17.068	20.996	19.046	18.902
1349	27.221	21.635	19.559	17.357	21.238	19.228	19.094
1350	27.658	22.002	19.920	17.650	21.483	19.412	19.288
1351	28.101	22.373	20.286	17.948	21.732	19.598	19.484
1352	28.549	22.749	20.658	18.251	21.984	19.786	19.683
1353	29.001	23.129	21.036	18.559	22.240	19.977	19.885
1354	29.459	23.515	21.420	18.871	22.499	20.171	20.089
1355	29.922	23.905	21.810	19.189	22.763	20.367	20.296
1356	30.389	24.300	22.205	19.512	23.030	20.566	20.506
1357	30.861	24.701	22.606	19.841	23.300	20.767	20.718
1358	31.338	25.106	23.013	20.174	23.575	20.970	20.934
1359	31.820	25.515	23.426	20.513	23.854	21.177	21.152
1360	32.306	25.930	23.845	20.857	24.136	21.386	21.372
1361	32.796	26.350	24.270	21.206	24.422	21.598	21.596
1362	33.291	26.775	24.700	21.561	24.713	21.812	21.823
1363	33.790	27.205	25.136	21.922	25.007	22.029	22.052
1364	34.293	27.639	25.579	22.287	25.306	22.249	22.285
1365	34.800	28.079	26.027	22.659	25.608	22.472	22.520
1366	35.312	28.524	26.480	23.035	25.915	22.698	22.758
1367	35.827	28.974	26.940	23.418	26.226	22.927	23.000
1368	36.346	29.428	27.405	23.806	26.541	23.158	23.244
1369	36.869	29.888	27.876	24.199	26.860	23.393	23.492

Table 8.21 (continuation six)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1370	37.395	30.354	28.353	24.599	27.183	23.631	23.742
1371	37.925	30.824	28.835	25.004	27.510	23.871	23.996
1372	38.458	31.300	29.323	25.414	27.841	24.115	24.253
1373	38.994	31.781	29.817	25.830	28.177	24.362	24.513
1374	39.534	32.267	30.316	26.252	28.517	24.612	24.776
1375	40.077	32.758	30.820	26.679	28.860	24.866	25.043
1376	40.622	33.255	31.330	27.112	29.208	25.122	25.313
1377	41.171	33.758	31.845	27.550	29.560	25.382	25.586
1378	41.722	34.265	32.365	27.994	29.916	25.645	25.862
1379	42.276	34.779	32.890	28.444	30.276	25.911	26.141
1380	42.833	35.298	33.420	28.898	30.640	26.181	26.424
1381	43.392	35.822	33.955	29.358	31.007	26.454	26.710
1382	43.954	36.353	34.495	29.823	31.379	26.731	26.999
1383	44.518	36.889	35.039	30.293	31.754	27.011	27.292
1384	45.084	37.430	35.588	30.769	32.133	27.294	27.588
1385	45.652	37.978	36.141	31.249	32.516	27.581	27.887
1386	46.223	38.531	36.698	31.733	32.903	27.872	28.189
1387	46.795	39.090	37.260	32.223	33.292	28.166	28.495
1388	47.369	39.655	37.825	32.717	33.686	28.464	28.805
1389	47.945	40.226	38.393	33.215	34.082	28.765	29.117
1390	48.523	40.802	38.965	33.717	34.482	29.070	29.433
1391	49.102	41.384	39.541	34.224	34.885	29.378	29.752
1392	49.683	41.972	40.119	34.734	35.292	29.691	30.074
1393	50.264	42.565	40.700	35.247	35.701	30.007	30.400
1394	50.847	43.163	41.284	35.764	36.113	30.326	30.729
1395	51.431	43.766	41.870	36.285	36.528	30.650	31.061
1396	52.016	44.375	42.458	36.808	36.945	30.977	31.396
1397	52.601	44.988	43.048	37.333	37.365	31.307	31.734
1398	53.187	45.606	43.640	37.862	37.788	31.642	32.076
1399	53.774	46.227	44.233	38.392	38.213	31.980	32.421
1400	54.360	46.853	44.827	38.925	38.640	32.322	32.768
1401	54.946	47.483	45.422	39.459	39.069	32.668	33.119
1402	55.532	48.115	46.018	39.995	39.500	33.018	33.473
1403	56.117	48.750	46.614	40.533	39.932	33.371	33.830
1404	56.702	49.388	47.210	41.071	40.367	33.728	34.190
1405	57.285	50.028	47.806	41.610	40.803	34.088	34.552

Table 8.21 (continuation seven)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1406	57.867	50.669	48.401	42.150	41.241	34.453	34.918
1407	58.448	51.311	48.996	42.690	41.680	34.821	35.286
1408	59.027	51.954	49.590	43.231	42.120	35.192	35.657
1409	59.603	52.597	50.183	43.771	42.561	35.568	36.030
1410	60.177	53.240	50.775	44.311	43.003	35.946	36.406
1411	60.749	53.882	51.365	44.851	43.446	36.329	36.785
1412	61.318	54.522	51.953	45.390	43.889	36.715	37.166
1413	61.883	55.162	52.539	45.928	44.334	37.104	37.550
1414	62.445	55.799	53.123	46.465	44.778	37.497	37.936
1415	63.004	56.434	53.705	47.001	45.223	37.893	38.324
1416	63.558	57.066	54.284	47.535	45.668	38.292	38.714
1417	64.108	57.695	54.861	48.068	46.113	38.694	39.106
1418	64.654	58.320	55.434	48.599	46.558	39.100	39.500
1419	65.196	58.942	56.005	49.129	47.003	39.509	39.896
1420	65.733	59.561	56.572	49.656	47.447	39.921	40.294
1421	66.264	60.174	57.137	50.182	47.891	40.335	40.694
1422	66.791	60.784	57.697	50.705	48.334	40.753	41.095
1423	67.313	61.389	58.254	51.226	48.777	41.173	41.498
1424	67.829	61.989	58.808	51.744	49.219	41.596	41.902
1425	68.340	62.584	59.357	52.259	49.660	42.022	42.307
1426	68.846	63.174	59.903	52.772	50.100	42.450	42.714
1427	69.346	63.759	60.445	53.283	50.539	42.881	43.122
1428	69.840	64.338	60.983	53.790	50.976	43.313	43.531
1429	70.329	64.912	61.516	54.294	51.413	43.748	43.940
1430	70.812	65.480	62.045	54.796	51.848	44.185	44.351
1431	71.289	66.042	62.570	55.294	52.281	44.624	44.762
1432	71.761	66.599	63.091	55.789	52.713	45.065	45.174
1433	72.227	67.149	63.607	56.281	53.143	45.507	45.586
1434	72.687	67.693	64.118	56.769	53.571	45.951	45.999
1435	73.142	68.231	64.625	57.254	53.997	46.397	46.412
1436	73.590	68.763	65.128	57.736	54.422	46.843	46.825
1437	74.033	69.288	65.625	58.214	54.844	47.291	47.238
1438	74.471	69.807	66.118	58.689	55.264	47.740	47.651
1439	74.903	70.320	66.606	59.160	55.683	48.190	48.063
1440	75.329	70.826	67.089	59.628	56.098	48.640	48.476
1441	75.750	71.326	67.568	60.092	56.512	49.092	48.888

Table 8.21 (continuation eight)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1442	76.165	71.819	68.041	60.552	56.923	49.543	49.300
1443	76.575	72.305	68.510	61.009	57.332	49.995	49.711
1444	76.979	72.785	68.973	61.462	57.738	50.448	50.122
1445	77.378	73.258	69.432	61.911	58.142	50.900	50.531
1446	77.772	73.724	69.885	62.356	58.543	51.352	50.940
1447	78.160	74.183	70.334	62.798	58.941	51.804	51.348
1448	78.543	74.636	70.777	63.236	59.337	52.256	51.755
1449	78.921	75.083	71.215	63.670	59.730	52.707	52.160
1450	79.294	75.522	71.649	64.100	60.120	53.157	52.565
1451	79.661	75.956	72.077	64.527	60.507	53.607	52.968
1452	80.024	76.382	72.500	64.949	60.892	54.056	53.370
1453	80.382	76.802	72.918	65.368	61.273	54.503	53.770
1454	80.734	77.216	73.331	65.783	61.652	54.950	54.169
1455	81.082	77.623	73.738	66.194	62.028	55.395	54.566
1456	81.425	78.024	74.141	66.601	62.400	55.839	54.961
1457	81.763	78.419	74.539	67.004	62.770	56.281	55.355
1458	82.096	78.808	74.931	67.404	63.137	56.721	55.746
1459	82.424	79.190	75.319	67.800	63.501	57.160	56.136
1460	82.748	79.567	75.701	68.192	63.862	57.596	56.524
1461	83.067	79.938	76.078	68.580	64.220	58.031	56.910
1462	83.381	80.303	76.451	68.964	64.575	58.463	57.294
1463	83.691	80.662	76.818	69.345	64.926	58.893	57.675
1464	83.996	81.016	77.180	69.722	65.275	59.321	58.055
1465	84.296	81.364	77.538	70.095	65.621	59.746	58.432
1466	84.592	81.707	77.890	70.464	65.963	60.168	58.807
1467	84.883	82.045	78.238	70.830	66.303	60.588	59.180
1468	85.170	82.377	78.580	71.192	66.640	61.004	59.551
1469	85.452	82.704	78.918	71.550	66.973	61.418	59.919
1470	85.730	83.027	79.251	71.905	67.304	61.829	60.285
1471	86.004	83.344	79.579	72.256	67.631	62.237	60.648
1472	86.273	83.656	79.903	72.603	67.956	62.642	61.009
1473	86.538	83.964	80.222	72.947	68.277	63.043	61.367
1474	86.798	84.267	80.536	73.287	68.596	63.441	61.723
1475	87.054	84.565	80.845	73.624	68.911	63.836	62.077
1476	87.306	84.859	81.150	73.957	69.224	64.227	62.428
1477	87.553	85.148	81.451	74.287	69.533	64.615	62.776



Table 8.21 (continuation nine)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1478	87.796	85.432	81.747	74.613	69.840	64.999	63.122
1479	88.035	85.713	82.038	74.936	70.143	65.380	63.465
1480	88.269	85.988	82.325	75.255	70.444	65.757	63.806
1481	88.500	86.260	82.608	75.571	70.742	66.131	64.144
1482	88.726	86.527	82.886	75.884	71.037	66.500	64.479
1483	88.948	86.790	83.160	76.193	71.329	66.867	64.812
1484	89.166	87.049	83.430	76.499	71.619	67.229	65.142
1485	89.380	87.303	83.695	76.802	71.905	67.587	65.470
1486	89.590	87.553	83.957	77.102	72.189	67.942	65.795
1487	89.796	87.800	84.214	77.398	72.470	68.293	66.117
1488	89.998	88.042	84.467	77.691	72.748	68.640	66.437
1489	90.196	88.280	84.716	77.981	73.023	68.983	66.754
1490	90.389	88.513	84.962	78.268	73.296	69.323	67.069
1491	90.580	88.743	85.203	78.551	73.566	69.658	67.381
1492	90.766	88.969	85.441	78.832	73.833	69.990	67.690
1493	90.948	89.191	85.674	79.109	74.097	70.318	67.997
1494	91.127	89.409	85.904	79.384	74.359	70.642	68.301
1495	91.302	89.623	86.130	79.655	74.619	70.962	68.603
1496	91.473	89.834	86.353	79.924	74.875	71.278	68.902
1497	91.641	90.040	86.572	80.189	75.129	71.591	69.198
1498	91.805	90.243	86.787	80.451	75.381	71.900	69.492
1499	91.965	90.442	86.998	80.711	75.630	72.205	69.784
1500	92.123	90.637	87.207	80.968	75.876	72.506	70.072
1501	92.276	90.829	87.411	81.221	76.120	72.803	70.359
1502	92.427	91.017	87.613	81.472	76.361	73.097	70.642
1503	92.574	91.201	87.811	81.720	76.600	73.387	70.924
1504	92.718	91.382	88.005	81.965	76.837	73.674	71.202
1505	92.858	91.560	88.197	82.208	77.071	73.957	71.478
1506	92.996	91.734	88.385	82.447	77.302	74.236	71.752
1507	93.131	91.905	88.570	82.684	77.531	74.512	72.023
1508	93.262	92.072	88.752	82.918	77.758	74.784	72.292
1509	93.391	92.236	88.931	83.149	77.983	75.053	72.558
1510	93.517	92.398	89.106	83.378	78.205	75.318	72.822
1511	93.640	92.556	89.279	83.604	78.425	75.580	73.083
1512	93.760	92.711	89.449	83.827	78.642	75.838	73.342
1513	93.877	92.863	89.616	84.048	78.857	76.094	73.598

Table 8.21 (continuation 10)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1514	93.992	93.012	89.780	84.266	79.070	76.345	73.852
1515	94.104	93.158	89.941	84.481	79.281	76.594	74.104
1516	94.214	93.301	90.100	84.694	79.490	76.839	74.353
1517	94.322	93.442	90.256	84.905	79.696	77.081	74.600
1518	94.427	93.580	90.409	85.112	79.900	77.320	74.844
1519	94.529	93.715	90.560	85.317	80.102	77.556	75.086
1520	94.630	93.847	90.708	85.520	80.302	77.788	75.326
1521	94.728	93.977	90.853	85.720	80.500	78.018	75.563
1522	94.824	94.105	90.996	85.918	80.696	78.244	75.799
1523	94.917	94.230	91.137	86.113	80.889	78.468	76.031
1524	95.009	94.352	91.275	86.306	81.081	78.688	76.262
1525	95.099	94.472	91.410	86.497	81.271	78.906	76.490
1526	95.187	94.590	91.544	86.685	81.458	79.121	76.716
1527	95.273	94.705	91.675	86.870	81.644	79.333	76.939
1528	95.357	94.818	91.804	87.054	81.827	79.542	77.161
1529	95.439	94.929	91.930	87.235	82.009	79.749	77.380
1530	95.520	95.038	92.055	87.413	82.188	79.953	77.597
1531	95.599	95.144	92.177	87.590	82.366	80.154	77.811
1532	95.676	95.248	92.297	87.764	82.542	80.352	78.024
1533	95.751	95.350	92.415	87.936	82.716	80.548	78.234
1534	95.825	95.451	92.532	88.105	82.888	80.742	78.442
1535	95.898	95.549	92.646	88.273	83.058	80.932	78.648
1536	95.969	95.645	92.758	88.438	83.226	81.121	78.852
1537	96.038	95.739	92.868	88.601	83.393	81.307	79.054
1538	96.106	95.831	92.977	88.762	83.558	81.490	79.253
1539	96.173	95.921	93.083	88.921	83.721	81.671	79.451
1540	96.238	96.009	93.188	89.077	83.882	81.850	79.646
1541	96.302	96.095	93.291	89.232	84.041	82.027	79.840
1542	96.365	96.180	93.392	89.384	84.199	82.201	80.031
1543	96.426	96.262	93.492	89.535	84.355	82.373	80.221
1544	96.487	96.343	93.589	89.683	84.510	82.543	80.408
1545	96.546	96.422	93.686	89.830	84.662	82.711	80.593
1546	96.604	96.499	93.780	89.974	84.813	82.876	80.777
1547	96.660	96.575	93.873	90.117	84.963	83.040	80.958
1548	96.716	96.649	93.964	90.257	85.111	83.201	81.138
1549	96.771	96.721	94.054	90.396	85.257	83.360	81.315

Table 8.21 (continuation 11)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1550	96.824	96.791	94.142	90.533	85.401	83.518	81.491
1551	96.877	96.860	94.229	90.667	85.544	83.673	81.665
1552	96.928	96.928	94.314	90.800	85.686	83.827	81.837
1553	96.979	96.993	94.398	90.932	85.826	83.978	82.007
1554	97.029	97.057	94.480	91.061	85.964	84.128	82.175
1555	97.077	97.120	94.561	91.188	86.101	84.276	82.342
1556	97.125	97.181	94.641	91.314	86.236	84.422	82.507
1557	97.172	97.241	94.719	91.438	86.370	84.567	82.670
1558	97.218	97.299	94.796	91.561	86.503	84.709	82.831
1559	97.264	97.356	94.872	91.681	86.634	84.850	82.990
1560	97.308	97.412	94.946	91.800	86.763	84.989	83.148
1561	97.352	97.466	95.019	91.917	86.892	85.127	83.304
1562	97.395	97.519	95.091	92.033	87.018	85.263	83.459
1563	97.437	97.570	95.161	92.146	87.144	85.397	83.612
1564	97.478	97.620	95.231	92.259	87.268	85.530	83.763
1565	97.519	97.669	95.299	92.369	87.390	85.661	83.912
1566	97.559	97.717	95.366	92.478	87.511	85.790	84.060
1567	97.598	97.764	95.432	92.586	87.631	85.919	84.207
1568	97.637	97.809	95.497	92.692	87.750	86.045	84.352
1569	97.675	97.854	95.561	92.796	87.867	86.170	84.495
1570	97.712	97.897	95.623	92.899	87.983	86.294	84.637
1571	97.748	97.939	95.685	93.000	88.098	86.416	84.777
1572	97.784	97.980	95.746	93.100	88.211	86.537	84.916
1573	97.820	98.020	95.805	93.199	88.323	86.657	85.053
1574	97.854	98.060	95.864	93.296	88.434	86.775	85.189
1575	97.889	98.098	95.921	93.391	88.544	86.892	85.323
1576	97.922	98.135	95.978	93.485	88.653	87.008	85.456
1577	97.955	98.171	96.033	93.578	88.760	87.122	85.588
1578	97.988	98.206	96.088	93.669	88.866	87.235	85.718
1579	98.020	98.241	96.142	93.759	88.971	87.347	85.847
1580	98.051	98.275	96.194	93.848	89.074	87.458	85.974
1581	98.082	98.307	96.246	93.935	89.177	87.568	86.100
1582	98.112	98.339	96.297	94.021	89.278	87.676	86.225
1583	98.142	98.371	96.347	94.106	89.378	87.783	86.349
1584	98.171	98.401	96.397	94.189	89.477	87.889	86.471
1585	98.200	98.431	96.445	94.271	89.575	87.994	86.592

Table 8.21 (continuation 12)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1586	98.229	98.460	96.493	94.352	89.672	88.098	86.711
1587	98.257	98.488	96.540	94.432	89.768	88.201	86.830
1588	98.284	98.515	96.586	94.510	89.863	88.302	86.947
1589	98.311	98.542	96.631	94.587	89.956	88.403	87.063
1590	98.338	98.569	96.675	94.663	90.049	88.502	87.178
1591	98.364	98.594	96.719	94.738	90.140	88.601	87.292
1592	98.390	98.619	96.762	94.812	90.230	88.699	87.404
1593	98.415	98.644	96.804	94.884	90.320	88.795	87.515
1594	98.440	98.668	96.846	94.956	90.408	88.891	87.625
1595	98.464	98.691	96.887	95.026	90.495	88.985	87.734
1596	98.488	98.714	96.927	95.095	90.582	89.079	87.842
1597	98.512	98.736	96.967	95.163	90.667	89.172	87.949
1598	98.535	98.758	97.005	95.231	90.751	89.264	88.054
1599	98.558	98.779	97.044	95.297	90.835	89.355	88.159
1600	98.581	98.800	97.081	95.362	90.917	89.445	88.263
1601	98.603	98.820	97.118	95.426	90.999	89.534	88.365
1602	98.625	98.840	97.154	95.488	91.079	89.622	88.466
1603	98.647	98.859	97.190	95.550	91.159	89.710	88.567
1604	98.668	98.878	97.225	95.611	91.237	89.796	88.666
1605	98.689	98.897	97.260	95.671	91.315	89.882	88.764
1606	98.709	98.915	97.294	95.731	91.392	89.967	88.861
1607	98.729	98.933	97.327	95.789	91.468	90.051	88.958
1608	98.749	98.950	97.360	95.846	91.542	90.135	89.053
1609	98.769	98.967	97.392	95.902	91.617	90.217	89.147
1610	98.788	98.984	97.424	95.958	91.690	90.299	89.241
1611	98.807	99.000	97.455	96.012	91.762	90.380	89.333
1612	98.825	99.016	97.486	96.066	91.834	90.461	89.424
1613	98.844	99.031	97.516	96.119	91.904	90.540	89.515
1614	98.862	99.046	97.546	96.171	91.974	90.619	89.604
1615	98.879	99.061	97.575	96.222	92.043	90.697	89.693
1616	98.897	99.076	97.604	96.273	92.112	90.775	89.781
1617	98.914	99.090	97.632	96.322	92.179	90.851	89.868
1618	98.931	99.104	97.660	96.371	92.246	90.927	89.954
1619	98.948	99.118	97.687	96.419	92.311	91.003	90.039
1620	98.964	99.132	97.714	96.467	92.376	91.077	90.123
1621	98.980	99.145	97.741	96.513	92.441	91.151	90.206
1622	98.996	99.158	97.767	96.559	92.504	91.225	90.289
1623	99.012	99.170	97.793	96.604	92.567	91.297	90.370

Table 8.21 (continuation 13)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1624	99.027	99.183	97.818	96.649	92.629	91.369	90.451
1625	99.042	99.195	97.843	96.692	92.690	91.441	90.531
1626	99.057	99.207	97.867	96.735	92.751	91.511	90.610
1627	99.072	99.219	97.892	96.778	92.811	91.582	90.689
1628	99.086	99.230	97.915	96.820	92.870	91.651	90.766
1629	99.100	99.242	97.939	96.861	92.928	91.720	90.843
1630	99.114	99.253	97.962	96.901	92.986	91.788	90.919
1631	99.128	99.264	97.984	96.941	93.043	91.856	90.994
1632	99.141	99.274	98.007	96.980	93.099	91.923	91.068
1633	99.155	99.285	98.029	97.019	93.155	91.989	91.142
1634	99.168	99.295	98.050	97.057	93.210	92.055	91.215
1635	99.181	99.305	98.071	97.094	93.264	92.121	91.287
1636	99.193	99.315	98.092	97.131	93.318	92.185	91.358
1637	99.206	99.325	98.113	97.167	93.371	92.249	91.429
1638	99.218	99.334	98.133	97.203	93.424	92.313	91.499
1639	99.230	99.344	98.153	97.238	93.476	92.376	91.568
1640	99.242	99.353	98.173	97.273	93.527	92.439	91.637
1641	99.254	99.362	98.193	97.307	93.577	92.501	91.704
1642	99.265	99.371	98.212	97.341	93.627	92.562	91.772
1643	99.277	99.380	98.231	97.374	93.677	92.623	91.838
1644	99.288	99.388	98.249	97.407	93.726	92.683	91.904
1645	99.299	99.397	98.268	97.439	93.774	92.743	91.969
1646	99.310	99.405	98.286	97.471	93.822	92.803	92.033
1647	99.320	99.413	98.303	97.502	93.869	92.861	92.097
1648	99.331	99.421	98.321	97.533	93.916	92.920	92.160
1649	99.341	99.429	98.338	97.563	93.962	92.978	92.222
1650	99.351	99.437	98.355	97.593	94.007	93.035	92.284
1651	99.361	99.444	98.372	97.622	94.052	93.092	92.345
1652	99.371	99.452	98.389	97.651	94.097	93.148	92.406
1653	99.381	99.459	98.405	97.680	94.141	93.204	92.466
1654	99.390	99.466	98.421	97.708	94.184	93.259	92.525
1655	99.400	99.474	98.437	97.736	94.227	93.314	92.584
1656	99.409	99.481	98.453	97.763	94.270	93.368	92.642
1657	99.418	99.488	98.468	97.790	94.312	93.422	92.700
1658	99.427	99.494	98.483	97.817	94.354	93.476	92.757
1659	99.436	99.501	98.498	97.843	94.395	93.529	92.813

Table 8.21 (continuation 14)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1660	99.445	99.508	98.513	97.869	94.435	93.581	92.869
1661	99.453	99.514	98.528	97.894	94.475	93.634	92.924
1662	99.461	99.520	98.542	97.919	94.515	93.685	92.979
1663	99.470	99.527	98.556	97.944	94.554	93.736	93.033
1664	99.478	99.533	98.570	97.968	94.593	93.787	93.087
1665	99.486	99.539	98.584	97.993	94.631	93.838	93.140
1666	99.494	99.545	98.598	98.016	94.669	93.888	93.192
1667	99.501	99.551	98.611	98.040	94.707	93.937	93.244
1668	99.509	99.557	98.625	98.063	94.744	93.986	93.296
1669	99.517	99.562	98.638	98.085	94.781	94.035	93.347
1670	99.524	99.568	98.651	98.108	94.817	94.083	93.397
1671	99.531	99.574	98.663	98.130	94.853	94.131	93.447
1672	99.538	99.579	98.676	98.152	94.888	94.178	93.497
1673	99.545	99.584	98.689	98.173	94.923	94.225	93.546
1674	99.552	99.590	98.701	98.195	94.958	94.272	93.594
1675	99.559	99.595	98.713	98.216	94.992	94.318	93.643
1676	99.566	99.600	98.725	98.236	95.026	94.363	93.690
1677	99.573	99.605	98.737	98.257	95.060	94.409	93.737
1678	99.579	99.610	98.748	98.277	95.093	94.454	93.784
1679	99.585	99.615	98.760	98.297	95.126	94.498	93.830
1680	99.592	99.620	98.771	98.316	95.158	94.543	93.876
1681	99.598	99.625	98.783	98.336	95.191	94.586	93.921
1682	99.604	99.629	98.794	98.355	95.222	94.630	93.966
1683	99.610	99.634	98.805	98.373	95.254	94.673	94.011
1684	99.616	99.639	98.816	98.392	95.285	94.715	94.055
1685	99.622	99.643	98.826	98.410	95.316	94.758	94.098
1686	99.628	99.648	98.837	98.428	95.346	94.800	94.142
1687	99.633	99.652	98.847	98.446	95.377	94.841	94.185
1688	99.639	99.656	98.858	98.464	95.406	94.882	94.227
1689	99.644	99.660	98.868	98.481	95.436	94.923	94.269
1690	99.650	99.665	98.878	98.498	95.465	94.964	94.311
1691	99.655	99.669	98.888	98.515	95.494	95.004	94.352
1692	99.660	99.673	98.898	98.532	95.523	95.044	94.393
1693	99.665	99.677	98.908	98.549	95.551	95.083	94.433
1694	99.670	99.681	98.917	98.565	95.579	95.122	94.473
1695	99.675	99.685	98.927	98.581	95.607	95.161	94.513
1696	99.680	99.689	98.936	98.597	95.635	95.199	94.552
1697	99.685	99.692	98.946	98.612	95.662	95.238	94.591
1698	99.690	99.696	98.955	98.628	95.689	95.275	94.630

Table 8.21 (continuation 15)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1699	99.694	99.700	98.964	98.643	95.715	95.313	94.668
1700	99.699	99.704	98.973	98.658	95.742	95.350	94.706
1701	N/A	N/A	N/A	98.673	95.768	95.387	94.744
1702	N/A	N/A	N/A	98.688	95.794	95.423	94.781
1703	N/A	N/A	N/A	98.702	95.820	95.459	94.818
1704	N/A	N/A	N/A	98.717	95.845	95.495	94.855
1705	N/A	N/A	N/A	98.731	95.870	95.530	94.891
1706	N/A	N/A	N/A	98.745	95.895	95.566	94.927
1707	N/A	N/A	N/A	98.759	95.920	95.600	94.962
1708	N/A	N/A	N/A	98.772	95.944	95.635	94.998
1709	N/A	N/A	N/A	98.786	95.968	95.669	95.033
1710	N/A	N/A	N/A	98.799	95.992	95.703	95.067
1711	N/A	N/A	N/A	98.812	96.016	95.737	95.102
1712	N/A	N/A	N/A	98.825	96.039	95.770	95.136
1713	N/A	N/A	N/A	98.838	96.063	95.803	95.170
1714	N/A	N/A	N/A	98.851	96.086	95.836	95.203
1715	N/A	N/A	N/A	98.863	96.109	95.869	95.236
1716	N/A	N/A	N/A	98.876	96.131	95.901	95.269
1717	N/A	N/A	N/A	98.888	96.154	95.933	95.302
1718	N/A	N/A	N/A	98.900	96.176	95.964	95.334
1719	N/A	N/A	N/A	98.912	96.198	95.996	95.366
1720	N/A	N/A	N/A	98.924	96.220	96.027	95.398
1721	N/A	N/A	N/A	98.935	96.241	96.058	95.430
1722	N/A	N/A	N/A	98.947	96.263	96.088	95.461
1723	N/A	N/A	N/A	98.958	96.284	96.118	95.492
1724	N/A	N/A	N/A	98.970	96.305	96.148	95.523
1725	N/A	N/A	N/A	98.981	96.326	96.178	95.553
1726	N/A	N/A	N/A	98.992	96.347	96.208	95.583
1727	N/A	N/A	N/A	99.002	96.367	96.237	95.613
1728	N/A	N/A	N/A	99.013	96.387	96.266	95.643
1729	N/A	N/A	N/A	99.024	96.407	96.295	95.673
1730	N/A	N/A	N/A	99.034	96.427	96.323	95.702
1731	N/A	N/A	N/A	99.045	96.447	96.351	95.731
1732	N/A	N/A	N/A	99.055	96.467	96.379	95.760
1733	N/A	N/A	N/A	99.065	96.486	96.407	95.788
1734	N/A	N/A	N/A	99.075	96.505	96.434	95.816
1735	N/A	N/A	N/A	99.085	96.525	96.462	95.844
1736	N/A	N/A	N/A	99.094	96.543	96.489	95.872
1737	N/A	N/A	N/A	99.104	96.562	96.515	95.900

Table 8.21 (continuation 16)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1738	N/A	N/A	N/A	99.114	96.581	96.542	95.927
1739	N/A	N/A	N/A	99.123	96.599	96.568	95.954
1740	N/A	N/A	N/A	99.132	96.618	96.594	95.981
1741	N/A	N/A	N/A	99.141	96.636	96.620	96.008
1742	N/A	N/A	N/A	99.151	96.654	96.646	96.035
1743	N/A	N/A	N/A	99.160	96.672	96.671	96.061
1744	N/A	N/A	N/A	99.168	96.689	96.696	96.087
1745	N/A	N/A	N/A	99.177	96.707	96.721	96.113
1746	N/A	N/A	N/A	99.186	96.724	96.746	96.138
1747	N/A	N/A	N/A	99.194	96.741	96.770	96.164
1748	N/A	N/A	N/A	99.203	96.759	96.794	96.189
1749	N/A	N/A	N/A	99.211	96.776	96.819	96.214
1750	N/A	N/A	N/A	99.220	96.792	96.842	96.239
1751	N/A	N/A	N/A	99.228	96.809	96.866	96.264
1752	N/A	N/A	N/A	99.236	96.826	96.889	96.288
1753	N/A	N/A	N/A	99.244	96.842	96.913	96.313
1754	N/A	N/A	N/A	99.252	96.858	96.936	96.337
1755	N/A	N/A	N/A	99.260	96.875	96.959	96.361
1756	N/A	N/A	N/A	99.267	96.891	96.981	96.385
1757	N/A	N/A	N/A	99.275	96.907	97.004	96.408
1758	N/A	N/A	N/A	99.283	96.922	97.026	96.431
1759	N/A	N/A	N/A	99.290	96.938	97.048	96.455
1760	N/A	N/A	N/A	99.298	96.954	97.070	96.478
1761	N/A	N/A	N/A	99.305	96.969	97.091	96.501
1762	N/A	N/A	N/A	99.312	96.984	97.113	96.523
1763	N/A	N/A	N/A	99.319	97.000	97.134	96.546
1764	N/A	N/A	N/A	99.326	97.015	97.155	96.568
1765	N/A	N/A	N/A	99.333	97.030	97.176	96.590
1766	N/A	N/A	N/A	99.340	97.044	97.197	96.612
1767	N/A	N/A	N/A	99.347	97.059	97.217	96.634
1768	N/A	N/A	N/A	99.354	97.074	97.238	96.656
1769	N/A	N/A	N/A	99.361	97.088	97.258	96.677
1770	N/A	N/A	N/A	99.367	97.103	97.278	96.699
1771	N/A	N/A	N/A	99.374	97.117	97.298	96.720
1772	N/A	N/A	N/A	99.380	97.131	97.318	96.741
1773	N/A	N/A	N/A	99.387	97.145	97.337	96.762
1774	N/A	N/A	N/A	99.393	97.159	97.356	96.782
1775	N/A	N/A	N/A	99.399	97.173	97.376	96.803
1776	N/A	N/A	N/A	99.405	97.187	97.395	96.823



Table 8.21 (continuation 17)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1777	N/A	N/A	N/A	99.411	97.201	97.413	96.844
1778	N/A	N/A	N/A	99.418	97.214	97.432	96.864
1779	N/A	N/A	N/A	99.424	97.228	97.451	96.884
1780	N/A	N/A	N/A	99.429	97.241	97.469	96.904
1781	N/A	N/A	N/A	99.435	97.254	97.487	96.923
1782	N/A	N/A	N/A	99.441	97.268	97.505	96.943
1783	N/A	N/A	N/A	99.447	97.281	97.523	96.962
1784	N/A	N/A	N/A	99.452	97.294	97.541	96.982
1785	N/A	N/A	N/A	99.458	97.307	97.558	97.001
1786	N/A	N/A	N/A	99.464	97.319	97.576	97.020
1787	N/A	N/A	N/A	99.469	97.332	97.593	97.038
1788	N/A	N/A	N/A	99.475	97.345	97.610	97.057
1789	N/A	N/A	N/A	99.480	97.357	97.627	97.076
1790	N/A	N/A	N/A	99.485	97.370	97.644	97.094
1791	N/A	N/A	N/A	99.490	97.382	97.661	97.113
1792	N/A	N/A	N/A	99.496	97.395	97.677	97.131
1793	N/A	N/A	N/A	99.501	97.407	97.694	97.149
1794	N/A	N/A	N/A	99.506	97.419	97.710	97.167
1795	N/A	N/A	N/A	99.511	97.431	97.726	97.184
1796	N/A	N/A	N/A	99.516	97.443	97.742	97.202
1797	N/A	N/A	N/A	99.521	97.455	97.758	97.220
1798	N/A	N/A	N/A	99.526	97.467	97.774	97.237
1799	N/A	N/A	N/A	99.530	97.478	97.789	97.254
1800	N/A	N/A	N/A	99.535	97.490	97.805	97.272
1801	N/A	N/A	N/A	N/A	97.502	97.820	97.289
1802	N/A	N/A	N/A	N/A	97.513	97.835	97.305
1803	N/A	N/A	N/A	N/A	97.524	97.851	97.322
1804	N/A	N/A	N/A	N/A	97.536	97.865	97.339
1805	N/A	N/A	N/A	N/A	97.547	97.880	97.356
1806	N/A	N/A	N/A	N/A	97.558	97.895	97.372
1807	N/A	N/A	N/A	N/A	97.569	97.910	97.388
1808	N/A	N/A	N/A	N/A	97.580	97.924	97.405
1809	N/A	N/A	N/A	N/A	97.591	97.938	97.421
1810	N/A	N/A	N/A	N/A	97.602	97.953	97.437
1811	N/A	N/A	N/A	N/A	97.613	97.967	97.453
1812	N/A	N/A	N/A	N/A	97.624	97.981	97.468
1813	N/A	N/A	N/A	N/A	97.634	97.995	97.484
1814	N/A	N/A	N/A	N/A	97.645	98.008	97.500
1815	N/A	N/A	N/A	N/A	97.656	98.022	97.515

Table 8.21 (continuation 18)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1816	N/A	N/A	N/A	N/A	97.666	98.036	97.530
1817	N/A	N/A	N/A	N/A	97.676	98.049	97.546
1818	N/A	N/A	N/A	N/A	97.687	98.062	97.561
1819	N/A	N/A	N/A	N/A	97.697	98.075	97.576
1820	N/A	N/A	N/A	N/A	97.707	98.089	97.591
1821	N/A	N/A	N/A	N/A	97.717	98.102	97.606
1822	N/A	N/A	N/A	N/A	97.728	98.114	97.620
1823	N/A	N/A	N/A	N/A	97.738	98.127	97.635
1824	N/A	N/A	N/A	N/A	97.747	98.140	97.649
1825	N/A	N/A	N/A	N/A	97.757	98.152	97.664
1826	N/A	N/A	N/A	N/A	97.767	98.165	97.678
1827	N/A	N/A	N/A	N/A	97.777	98.177	97.692
1828	N/A	N/A	N/A	N/A	97.787	98.189	97.707
1829	N/A	N/A	N/A	N/A	97.796	98.202	97.721
1830	N/A	N/A	N/A	N/A	97.806	98.214	97.734
1831	N/A	N/A	N/A	N/A	97.815	98.226	97.748
1832	N/A	N/A	N/A	N/A	97.825	98.237	97.762
1833	N/A	N/A	N/A	N/A	97.834	98.249	97.776
1834	N/A	N/A	N/A	N/A	97.844	98.261	97.789
1835	N/A	N/A	N/A	N/A	97.853	98.272	97.803
1836	N/A	N/A	N/A	N/A	97.862	98.284	97.816
1837	N/A	N/A	N/A	N/A	97.871	98.295	97.829
1838	N/A	N/A	N/A	N/A	97.881	98.307	97.843
1839	N/A	N/A	N/A	N/A	97.890	98.318	97.856
1840	N/A	N/A	N/A	N/A	97.899	98.329	97.869
1841	N/A	N/A	N/A	N/A	97.908	98.340	97.882
1842	N/A	N/A	N/A	N/A	97.916	98.351	97.895
1843	N/A	N/A	N/A	N/A	97.925	98.362	97.907
1844	N/A	N/A	N/A	N/A	97.934	98.372	97.920
1845	N/A	N/A	N/A	N/A	97.943	98.383	97.933
1846	N/A	N/A	N/A	N/A	97.952	98.394	97.945
1847	N/A	N/A	N/A	N/A	97.960	98.404	97.958
1848	N/A	N/A	N/A	N/A	97.969	98.415	97.970
1849	N/A	N/A	N/A	N/A	97.977	98.425	97.982
1850	N/A	N/A	N/A	N/A	97.986	98.435	97.994
1851	N/A	N/A	N/A	N/A	97.994	98.445	98.006
1852	N/A	N/A	N/A	N/A	98.003	98.455	98.018
1853	N/A	N/A	N/A	N/A	98.011	98.465	98.030
1854	N/A	N/A	N/A	N/A	98.019	98.475	98.042

Table 8.21 (continuation 19)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1855	N/A	N/A	N/A	N/A	98.028	98.485	98.054
1856	N/A	N/A	N/A	N/A	98.036	98.495	98.066
1857	N/A	N/A	N/A	N/A	98.044	98.505	98.077
1858	N/A	N/A	N/A	N/A	98.052	98.514	98.089
1859	N/A	N/A	N/A	N/A	98.060	98.524	98.100
1860	N/A	N/A	N/A	N/A	98.068	98.533	98.112
1861	N/A	N/A	N/A	N/A	98.076	98.543	98.123
1862	N/A	N/A	N/A	N/A	98.084	98.552	98.134
1863	N/A	N/A	N/A	N/A	98.092	98.561	98.145
1864	N/A	N/A	N/A	N/A	98.100	98.570	98.156
1865	N/A	N/A	N/A	N/A	98.108	98.580	98.167
1866	N/A	N/A	N/A	N/A	98.115	98.589	98.178
1867	N/A	N/A	N/A	N/A	98.123	98.598	98.189
1868	N/A	N/A	N/A	N/A	98.131	98.606	98.200
1869	N/A	N/A	N/A	N/A	98.138	98.615	98.211
1870	N/A	N/A	N/A	N/A	98.146	98.624	98.221
1871	N/A	N/A	N/A	N/A	98.153	98.633	98.232
1872	N/A	N/A	N/A	N/A	98.161	98.641	98.243
1873	N/A	N/A	N/A	N/A	98.168	98.650	98.253
1874	N/A	N/A	N/A	N/A	98.176	98.658	98.263
1875	N/A	N/A	N/A	N/A	98.183	98.667	98.274
1876	N/A	N/A	N/A	N/A	98.190	98.675	98.284
1877	N/A	N/A	N/A	N/A	98.198	98.683	98.294
1878	N/A	N/A	N/A	N/A	98.205	98.692	98.304
1879	N/A	N/A	N/A	N/A	98.212	98.700	98.314
1880	N/A	N/A	N/A	N/A	98.219	98.708	98.324
1881	N/A	N/A	N/A	N/A	98.226	98.716	98.334
1882	N/A	N/A	N/A	N/A	98.233	98.724	98.344
1883	N/A	N/A	N/A	N/A	98.240	98.732	98.354
1884	N/A	N/A	N/A	N/A	98.247	98.740	98.364
1885	N/A	N/A	N/A	N/A	98.254	98.748	98.373
1886	N/A	N/A	N/A	N/A	98.261	98.755	98.383
1887	N/A	N/A	N/A	N/A	98.268	98.763	98.392
1888	N/A	N/A	N/A	N/A	98.275	98.771	98.402
1889	N/A	N/A	N/A	N/A	98.282	98.778	98.411
1890	N/A	N/A	N/A	N/A	98.289	98.786	98.421
1891	N/A	N/A	N/A	N/A	98.295	98.793	98.430
1892	N/A	N/A	N/A	N/A	98.302	98.801	98.439
1893	N/A	N/A	N/A	N/A	98.309	98.808	98.448

Table 8.21 (continuation 20)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1894	N/A	N/A	N/A	N/A	98.315	98.815	98.457
1895	N/A	N/A	N/A	N/A	98.322	98.822	98.466
1896	N/A	N/A	N/A	N/A	98.329	98.830	98.475
1897	N/A	N/A	N/A	N/A	98.335	98.837	98.484
1898	N/A	N/A	N/A	N/A	98.342	98.844	98.493
1899	N/A	N/A	N/A	N/A	98.348	98.851	98.502
1900	N/A	N/A	N/A	N/A	98.355	98.858	98.511
1901	N/A	N/A	N/A	N/A	N/A	98.865	98.520
1902	N/A	N/A	N/A	N/A	N/A	98.872	98.528
1903	N/A	N/A	N/A	N/A	N/A	98.878	98.537
1904	N/A	N/A	N/A	N/A	N/A	98.885	98.545
1905	N/A	N/A	N/A	N/A	N/A	98.892	98.554
1906	N/A	N/A	N/A	N/A	N/A	98.899	98.562
1907	N/A	N/A	N/A	N/A	N/A	98.905	98.571
1908	N/A	N/A	N/A	N/A	N/A	98.912	98.579
1909	N/A	N/A	N/A	N/A	N/A	98.918	98.587
1910	N/A	N/A	N/A	N/A	N/A	98.925	98.595
1911	N/A	N/A	N/A	N/A	N/A	98.931	98.604
1912	N/A	N/A	N/A	N/A	N/A	98.937	98.612
1913	N/A	N/A	N/A	N/A	N/A	98.944	98.620
1914	N/A	N/A	N/A	N/A	N/A	98.950	98.628
1915	N/A	N/A	N/A	N/A	N/A	98.956	98.636
1916	N/A	N/A	N/A	N/A	N/A	98.963	98.644
1917	N/A	N/A	N/A	N/A	N/A	98.969	98.652
1918	N/A	N/A	N/A	N/A	N/A	98.975	98.659
1919	N/A	N/A	N/A	N/A	N/A	98.981	98.667
1920	N/A	N/A	N/A	N/A	N/A	98.987	98.675
1921	N/A	N/A	N/A	N/A	N/A	98.993	98.682
1922	N/A	N/A	N/A	N/A	N/A	98.999	98.690
1923	N/A	N/A	N/A	N/A	N/A	99.005	98.698
1924	N/A	N/A	N/A	N/A	N/A	99.010	98.705
1925	N/A	N/A	N/A	N/A	N/A	99.016	98.713
1926	N/A	N/A	N/A	N/A	N/A	99.022	98.720
1927	N/A	N/A	N/A	N/A	N/A	99.028	98.727
1928	N/A	N/A	N/A	N/A	N/A	99.033	98.735
1929	N/A	N/A	N/A	N/A	N/A	99.039	98.742
1930	N/A	N/A	N/A	N/A	N/A	99.045	98.749
1931	N/A	N/A	N/A	N/A	N/A	99.050	98.756
1932	N/A	N/A	N/A	N/A	N/A	99.056	98.764

Table 8.21 (continuation 21)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1933	N/A	N/A	N/A	N/A	N/A	99.061	98.771
1934	N/A	N/A	N/A	N/A	N/A	99.067	98.778
1935	N/A	N/A	N/A	N/A	N/A	99.072	98.785
1936	N/A	N/A	N/A	N/A	N/A	99.077	98.792
1937	N/A	N/A	N/A	N/A	N/A	99.083	98.799
1938	N/A	N/A	N/A	N/A	N/A	99.088	98.806
1939	N/A	N/A	N/A	N/A	N/A	99.093	98.812
1940	N/A	N/A	N/A	N/A	N/A	99.098	98.819
1941	N/A	N/A	N/A	N/A	N/A	99.104	98.826
1942	N/A	N/A	N/A	N/A	N/A	99.109	98.833
1943	N/A	N/A	N/A	N/A	N/A	99.114	98.839
1944	N/A	N/A	N/A	N/A	N/A	99.119	98.846
1945	N/A	N/A	N/A	N/A	N/A	99.124	98.852
1946	N/A	N/A	N/A	N/A	N/A	99.129	98.859
1947	N/A	N/A	N/A	N/A	N/A	99.134	98.865
1948	N/A	N/A	N/A	N/A	N/A	99.139	98.872
1949	N/A	N/A	N/A	N/A	N/A	99.144	98.878
1950	N/A	N/A	N/A	N/A	N/A	99.149	98.885

**Table 8.22 TCC Data for the Written Language Composite (Figure 8.2)**

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1150	2.135	0.034	0.091	0.333	0.338	0.751	1.411
1151	2.164	0.034	0.092	0.337	0.342	0.759	1.418
1152	2.193	0.035	0.094	0.341	0.346	0.768	1.425
1153	2.223	0.036	0.095	0.345	0.350	0.777	1.432
1154	2.253	0.037	0.097	0.349	0.354	0.786	1.439
1155	2.283	0.039	0.099	0.353	0.358	0.795	1.446
1156	2.314	0.040	0.100	0.357	0.362	0.804	1.453
1157	2.345	0.041	0.102	0.361	0.366	0.813	1.461
1158	2.377	0.042	0.104	0.365	0.370	0.823	1.468
1159	2.409	0.043	0.105	0.370	0.375	0.832	1.475
1160	2.441	0.044	0.107	0.374	0.379	0.842	1.483
1161	2.474	0.045	0.109	0.379	0.384	0.851	1.490
1162	2.508	0.047	0.111	0.383	0.388	0.861	1.498
1163	2.541	0.048	0.113	0.388	0.393	0.871	1.506
1164	2.576	0.049	0.115	0.393	0.397	0.881	1.514
1165	2.611	0.051	0.117	0.397	0.402	0.891	1.522
1166	2.646	0.052	0.119	0.402	0.407	0.902	1.530
1167	2.682	0.054	0.121	0.407	0.411	0.912	1.538
1168	2.718	0.055	0.123	0.412	0.416	0.922	1.546
1169	2.755	0.057	0.125	0.417	0.421	0.933	1.554
1170	2.792	0.058	0.127	0.422	0.426	0.944	1.563
1171	2.830	0.060	0.129	0.427	0.432	0.955	1.571
1172	2.868	0.062	0.132	0.433	0.437	0.966	1.580
1173	2.907	0.064	0.134	0.438	0.442	0.977	1.588
1174	2.946	0.065	0.136	0.443	0.447	0.988	1.597
1175	2.986	0.067	0.139	0.449	0.453	1.000	1.606
1176	3.027	0.069	0.141	0.455	0.458	1.011	1.615
1177	3.068	0.071	0.143	0.460	0.464	1.023	1.624
1178	3.109	0.073	0.146	0.466	0.469	1.035	1.633
1179	3.152	0.075	0.149	0.472	0.475	1.046	1.643
1180	3.194	0.077	0.151	0.478	0.481	1.059	1.652
1181	3.238	0.080	0.154	0.484	0.487	1.071	1.661
1182	3.282	0.082	0.157	0.490	0.493	1.083	1.671
1183	3.326	0.084	0.159	0.496	0.499	1.096	1.681
1184	3.372	0.087	0.162	0.502	0.505	1.108	1.691
1185	3.417	0.089	0.165	0.509	0.512	1.121	1.701
1186	3.464	0.092	0.168	0.515	0.518	1.134	1.711
1187	3.511	0.094	0.171	0.522	0.524	1.147	1.721
1188	3.559	0.097	0.174	0.529	0.531	1.160	1.731

Table 8.22 (continuation one)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1189	3.607	0.100	0.177	0.536	0.538	1.174	1.741
1190	3.656	0.103	0.180	0.542	0.544	1.187	1.752
1191	3.706	0.106	0.184	0.549	0.551	1.201	1.763
1192	3.757	0.109	0.187	0.557	0.558	1.215	1.773
1193	3.808	0.112	0.191	0.564	0.565	1.229	1.784
1194	3.860	0.115	0.194	0.571	0.572	1.243	1.795
1195	3.913	0.118	0.198	0.579	0.580	1.257	1.806
1196	3.966	0.122	0.201	0.586	0.587	1.272	1.818
1197	4.020	0.125	0.205	0.594	0.595	1.286	1.829
1198	4.075	0.129	0.209	0.602	0.602	1.301	1.841
1199	4.131	0.133	0.213	0.610	0.610	1.316	1.852
1200	4.187	0.136	0.216	0.618	0.618	1.331	1.864
1201	4.245	0.140	0.220	0.626	0.626	1.346	1.876
1202	4.303	0.144	0.225	0.634	0.634	1.362	1.888
1203	4.362	0.149	0.229	0.643	0.642	1.377	1.900
1204	4.421	0.153	0.233	0.651	0.650	1.393	1.913
1205	4.482	0.157	0.237	0.660	0.659	1.409	1.925
1206	4.543	0.162	0.242	0.669	0.668	1.425	1.938
1207	4.606	0.167	0.247	0.678	0.676	1.442	1.951
1208	4.669	0.171	0.251	0.687	0.685	1.458	1.964
1209	4.733	0.176	0.256	0.696	0.694	1.475	1.977
1210	4.798	0.182	0.261	0.706	0.703	1.492	1.990
1211	4.864	0.187	0.266	0.715	0.713	1.509	2.004
1212	4.931	0.192	0.271	0.725	0.722	1.526	2.017
1213	4.999	0.198	0.276	0.735	0.732	1.544	2.031
1214	5.068	0.204	0.281	0.745	0.741	1.561	2.045
1215	5.138	0.210	0.287	0.755	0.751	1.579	2.059
1216	5.209	0.216	0.292	0.765	0.761	1.597	2.074
1217	5.281	0.222	0.298	0.776	0.772	1.615	2.088
1218	5.354	0.229	0.304	0.786	0.782	1.634	2.103
1219	5.428	0.235	0.310	0.797	0.792	1.652	2.118
1220	5.503	0.242	0.316	0.808	0.803	1.671	2.133
1221	5.579	0.249	0.322	0.820	0.814	1.690	2.148
1222	5.657	0.256	0.329	0.831	0.825	1.709	2.163
1223	5.735	0.264	0.335	0.843	0.836	1.729	2.179
1224	5.815	0.272	0.342	0.854	0.848	1.748	2.195
1225	5.896	0.280	0.349	0.866	0.859	1.768	2.211
1226	5.978	0.288	0.356	0.878	0.871	1.788	2.227
1227	6.061	0.296	0.363	0.891	0.883	1.809	2.243

Table 8.22 (continuation two)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1228	6.145	0.305	0.370	0.903	0.895	1.829	2.260
1229	6.231	0.314	0.377	0.916	0.907	1.850	2.277
1230	6.318	0.323	0.385	0.929	0.920	1.871	2.294
1231	6.407	0.333	0.393	0.942	0.932	1.892	2.311
1232	6.496	0.343	0.401	0.955	0.945	1.913	2.328
1233	6.588	0.353	0.409	0.969	0.958	1.935	2.346
1234	6.680	0.363	0.418	0.983	0.972	1.957	2.364
1235	6.774	0.374	0.426	0.997	0.985	1.979	2.382
1236	6.869	0.385	0.435	1.011	0.999	2.001	2.400
1237	6.966	0.396	0.444	1.026	1.013	2.024	2.418
1238	7.064	0.408	0.453	1.040	1.027	2.046	2.437
1239	7.164	0.420	0.463	1.055	1.041	2.069	2.456
1240	7.266	0.433	0.472	1.071	1.056	2.093	2.475
1241	7.369	0.445	0.482	1.086	1.071	2.116	2.495
1242	7.473	0.459	0.493	1.102	1.086	2.140	2.515
1243	7.580	0.472	0.503	1.118	1.101	2.164	2.535
1244	7.688	0.486	0.514	1.134	1.117	2.188	2.555
1245	7.797	0.501	0.525	1.151	1.133	2.212	2.575
1246	7.909	0.515	0.536	1.167	1.149	2.237	2.596
1247	8.022	0.531	0.547	1.184	1.165	2.262	2.617
1248	8.137	0.547	0.559	1.202	1.182	2.287	2.638
1249	8.254	0.563	0.571	1.219	1.198	2.313	2.660
1250	8.373	0.580	0.584	1.237	1.216	2.339	2.682
1251	8.494	0.597	0.596	1.256	1.233	2.365	2.704
1252	8.616	0.615	0.609	1.274	1.251	2.391	2.726
1253	8.741	0.633	0.623	1.293	1.269	2.418	2.749
1254	8.868	0.652	0.637	1.312	1.287	2.444	2.771
1255	8.997	0.671	0.651	1.332	1.305	2.471	2.795
1256	9.128	0.691	0.665	1.351	1.324	2.499	2.818
1257	9.262	0.712	0.680	1.372	1.343	2.527	2.842
1258	9.398	0.733	0.695	1.392	1.363	2.554	2.866
1259	9.536	0.755	0.711	1.413	1.382	2.583	2.891
1260	9.676	0.778	0.727	1.434	1.402	2.611	2.915
1261	9.819	0.801	0.743	1.455	1.423	2.640	2.940
1262	9.964	0.825	0.760	1.477	1.443	2.669	2.966
1263	10.112	0.850	0.777	1.500	1.465	2.699	2.991
1264	10.262	0.875	0.795	1.522	1.486	2.728	3.017
1265	10.415	0.902	0.813	1.545	1.508	2.758	3.044
1266	10.571	0.929	0.832	1.569	1.530	2.789	3.070



Table 8.22 (continuation three)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1267	10.730	0.956	0.852	1.592	1.552	2.819	3.097
1268	10.891	0.985	0.871	1.617	1.575	2.850	3.125
1269	11.055	1.015	0.892	1.641	1.598	2.881	3.152
1270	11.222	1.045	0.913	1.666	1.621	2.913	3.181
1271	11.393	1.077	0.934	1.692	1.645	2.945	3.209
1272	11.566	1.109	0.956	1.717	1.669	2.977	3.238
1273	11.742	1.142	0.979	1.744	1.694	3.010	3.267
1274	11.922	1.177	1.002	1.771	1.719	3.042	3.296
1275	12.105	1.212	1.026	1.798	1.744	3.076	3.326
1276	12.291	1.249	1.050	1.825	1.770	3.109	3.357
1277	12.481	1.286	1.076	1.854	1.796	3.143	3.387
1278	12.674	1.325	1.102	1.882	1.823	3.177	3.419
1279	12.871	1.365	1.128	1.911	1.850	3.212	3.450
1280	13.071	1.406	1.156	1.941	1.877	3.247	3.482
1281	13.275	1.449	1.184	1.971	1.905	3.282	3.514
1282	13.483	1.493	1.213	2.002	1.933	3.318	3.547
1283	13.694	1.538	1.243	2.033	1.962	3.354	3.580
1284	13.910	1.584	1.274	2.064	1.991	3.390	3.614
1285	14.129	1.632	1.305	2.097	2.021	3.427	3.648
1286	14.353	1.681	1.338	2.129	2.051	3.464	3.682
1287	14.581	1.732	1.371	2.163	2.082	3.501	3.717
1288	14.813	1.784	1.406	2.197	2.113	3.539	3.752
1289	15.049	1.838	1.441	2.231	2.144	3.577	3.788
1290	15.289	1.894	1.478	2.266	2.176	3.616	3.824
1291	15.534	1.951	1.515	2.302	2.209	3.655	3.861
1292	15.783	2.010	1.554	2.338	2.242	3.694	3.898
1293	16.037	2.071	1.593	2.375	2.275	3.734	3.936
1294	16.296	2.134	1.634	2.413	2.309	3.774	3.974
1295	16.559	2.198	1.676	2.451	2.344	3.815	4.013
1296	16.827	2.265	1.720	2.490	2.379	3.856	4.052
1297	17.099	2.334	1.764	2.529	2.415	3.898	4.092
1298	17.377	2.404	1.810	2.570	2.451	3.939	4.132
1299	17.659	2.477	1.858	2.611	2.488	3.982	4.173
1300	17.946	2.552	1.907	2.652	2.525	4.025	4.214
1301	18.238	2.630	1.957	2.695	2.563	4.068	4.256
1302	18.535	2.709	2.008	2.738	2.601	4.112	4.298
1303	18.838	2.791	2.062	2.782	2.640	4.156	4.341
1304	19.145	2.876	2.117	2.826	2.680	4.200	4.385
1305	19.458	2.963	2.173	2.872	2.720	4.245	4.429

Table 8.22 (continuation four)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1306	19.776	3.053	2.231	2.918	2.761	4.291	4.473
1307	20.099	3.146	2.291	2.965	2.802	4.337	4.518
1308	20.427	3.241	2.353	3.012	2.845	4.384	4.564
1309	20.761	3.339	2.417	3.061	2.887	4.431	4.610
1310	21.100	3.440	2.483	3.111	2.931	4.478	4.657
1311	21.445	3.545	2.550	3.161	2.975	4.526	4.705
1312	21.796	3.652	2.620	3.212	3.019	4.575	4.753
1313	22.152	3.763	2.692	3.264	3.065	4.624	4.802
1314	22.513	3.877	2.766	3.317	3.111	4.673	4.851
1315	22.881	3.994	2.843	3.371	3.157	4.723	4.901
1316	23.254	4.115	2.921	3.426	3.205	4.774	4.952
1317	23.634	4.240	3.002	3.482	3.253	4.825	5.004
1318	24.019	4.369	3.086	3.539	3.301	4.877	5.056
1319	24.411	4.501	3.172	3.597	3.351	4.929	5.108
1320	24.809	4.637	3.261	3.656	3.401	4.982	5.162
1321	25.214	4.778	3.353	3.716	3.452	5.036	5.216
1322	25.625	4.922	3.448	3.777	3.504	5.090	5.271
1323	26.043	5.072	3.545	3.839	3.556	5.144	5.326
1324	26.468	5.225	3.645	3.902	3.609	5.200	5.383
1325	26.900	5.383	3.749	3.966	3.663	5.255	5.440
1326	27.340	5.546	3.856	4.031	3.718	5.312	5.497
1327	27.787	5.714	3.966	4.098	3.774	5.369	5.556
1328	28.243	5.887	4.080	4.166	3.830	5.427	5.615
1329	28.706	6.065	4.197	4.234	3.887	5.485	5.675
1330	29.178	6.248	4.317	4.305	3.945	5.544	5.736
1331	29.659	6.437	4.442	4.376	4.004	5.604	5.798
1332	30.149	6.631	4.570	4.448	4.064	5.665	5.860
1333	30.649	6.831	4.702	4.522	4.124	5.726	5.923
1334	31.158	7.037	4.839	4.598	4.185	5.787	5.987
1335	31.678	7.250	4.980	4.674	4.248	5.850	6.052
1336	32.208	7.468	5.125	4.752	4.311	5.913	6.118
1337	32.749	7.693	5.274	4.831	4.375	5.977	6.184
1338	33.301	7.925	5.428	4.912	4.440	6.042	6.252
1339	33.865	8.163	5.587	4.994	4.506	6.108	6.320
1340	34.441	8.409	5.751	5.078	4.573	6.174	6.389
1341	35.030	8.661	5.919	5.163	4.640	6.241	6.459
1342	35.631	8.921	6.093	5.249	4.709	6.309	6.530
1343	36.244	9.189	6.272	5.338	4.779	6.378	6.602
1344	36.871	9.464	6.457	5.427	4.849	6.447	6.675

Table 8.22 (continuation five)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1345	37.512	9.747	6.647	5.519	4.921	6.517	6.749
1346	38.166	10.039	6.843	5.611	4.994	6.589	6.823
1347	38.834	10.339	7.045	5.706	5.068	6.661	6.899
1348	39.516	10.647	7.253	5.802	5.142	6.734	6.976
1349	40.212	10.964	7.467	5.900	5.218	6.807	7.053
1350	40.921	11.289	7.687	6.000	5.295	6.882	7.132
1351	41.645	11.624	7.914	6.102	5.373	6.958	7.212
1352	42.381	11.968	8.148	6.205	5.452	7.034	7.292
1353	43.132	12.322	8.388	6.310	5.532	7.112	7.374
1354	43.895	12.685	8.636	6.417	5.613	7.191	7.457
1355	44.671	13.058	8.890	6.526	5.696	7.270	7.540
1356	45.459	13.441	9.152	6.637	5.779	7.351	7.625
1357	46.258	13.834	9.422	6.750	5.864	7.432	7.711
1358	47.069	14.237	9.698	6.865	5.950	7.515	7.798
1359	47.889	14.651	9.983	6.982	6.037	7.598	7.886
1360	48.719	15.075	10.276	7.101	6.125	7.683	7.975
1361	49.557	15.510	10.576	7.223	6.215	7.769	8.066
1362	50.403	15.955	10.885	7.346	6.306	7.856	8.157
1363	51.255	16.411	11.202	7.472	6.398	7.944	8.250
1364	52.113	16.877	11.528	7.600	6.491	8.033	8.344
1365	52.975	17.355	11.862	7.730	6.586	8.124	8.439
1366	53.840	17.843	12.205	7.863	6.682	8.215	8.535
1367	54.708	18.342	12.557	7.998	6.779	8.308	8.632
1368	55.576	18.852	12.917	8.135	6.878	8.402	8.731
1369	56.445	19.372	13.287	8.275	6.978	8.497	8.831
1370	57.312	19.903	13.665	8.417	7.080	8.594	8.932
1371	58.178	20.445	14.053	8.562	7.182	8.692	9.035
1372	59.039	20.997	14.450	8.710	7.287	8.791	9.138
1373	59.897	21.559	14.857	8.860	7.393	8.891	9.243
1374	60.749	22.132	15.272	9.013	7.500	8.993	9.350
1375	61.595	22.714	15.697	9.169	7.609	9.097	9.457
1376	62.433	23.306	16.132	9.328	7.720	9.201	9.566
1377	63.263	23.908	16.575	9.489	7.832	9.307	9.677
1378	64.084	24.520	17.029	9.653	7.945	9.415	9.789
1379	64.895	25.140	17.491	9.821	8.060	9.524	9.902
1380	65.696	25.770	17.963	9.991	8.177	9.635	10.016
1381	66.484	26.408	18.444	10.165	8.296	9.747	10.133
1382	67.261	27.055	18.935	10.341	8.416	9.861	10.250
1383	68.024	27.710	19.434	10.521	8.538	9.976	10.369

Table 8.22 (continuation six)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1384	68.774	28.373	19.943	10.704	8.662	10.093	10.490
1385	69.509	29.044	20.461	10.890	8.787	10.212	10.612
1386	70.230	29.722	20.987	11.080	8.915	10.332	10.735
1387	70.936	30.408	21.522	11.273	9.044	10.455	10.860
1388	71.627	31.101	22.066	11.469	9.175	10.578	10.987
1389	72.301	31.800	22.619	11.670	9.308	10.704	11.115
1390	72.960	32.507	23.179	11.873	9.443	10.832	11.245
1391	73.602	33.219	23.748	12.081	9.580	10.961	11.377
1392	74.228	33.937	24.324	12.292	9.719	11.092	11.510
1393	74.838	34.662	24.909	12.506	9.859	11.225	11.645
1394	75.432	35.391	25.500	12.725	10.002	11.360	11.781
1395	76.009	36.126	26.100	12.948	10.147	11.497	11.919
1396	76.570	36.866	26.706	13.174	10.294	11.637	12.060
1397	77.116	37.611	27.319	13.405	10.444	11.778	12.201
1398	77.646	38.360	27.939	13.640	10.595	11.921	12.345
1399	78.161	39.113	28.565	13.878	10.749	12.066	12.491
1400	78.661	39.870	29.197	14.122	10.905	12.214	12.638
1401	79.147	40.631	29.835	14.369	11.063	12.364	12.787
1402	79.618	41.395	30.479	14.621	11.224	12.516	12.938
1403	80.076	42.162	31.128	14.877	11.387	12.670	13.091
1404	80.521	42.931	31.783	15.137	11.552	12.826	13.246
1405	80.952	43.703	32.442	15.402	11.720	12.985	13.403
1406	81.372	44.477	33.106	15.672	11.890	13.147	13.562
1407	81.779	45.252	33.775	15.947	12.063	13.310	13.724
1408	82.175	46.029	34.447	16.226	12.238	13.476	13.887
1409	82.560	46.807	35.124	16.510	12.416	13.645	14.052
1410	82.934	47.586	35.804	16.798	12.597	13.816	14.219
1411	83.298	48.364	36.487	17.092	12.780	13.990	14.389
1412	83.652	49.143	37.174	17.391	12.966	14.167	14.561
1413	83.996	49.922	37.864	17.694	13.155	14.346	14.735
1414	84.332	50.699	38.556	18.003	13.346	14.527	14.911
1415	84.659	51.476	39.251	18.317	13.540	14.712	15.090
1416	84.977	52.251	39.948	18.636	13.737	14.899	15.271
1417	85.287	53.025	40.647	18.960	13.937	15.089	15.454
1418	85.590	53.796	41.348	19.289	14.140	15.282	15.640
1419	85.884	54.565	42.050	19.624	14.346	15.478	15.828
1420	86.172	55.331	42.754	19.964	14.555	15.677	16.018
1421	86.453	56.094	43.458	20.309	14.766	15.879	16.211
1422	86.727	56.854	44.163	20.660	14.981	16.084	16.407

Table 8.22 (continuation seven)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1423	86.994	57.609	44.869	21.016	15.199	16.292	16.605
1424	87.255	58.361	45.576	21.377	15.420	16.502	16.806
1425	87.510	59.109	46.282	21.744	15.644	16.716	17.009
1426	87.759	59.852	46.989	22.117	15.871	16.934	17.215
1427	88.002	60.589	47.695	22.495	16.101	17.154	17.424
1428	88.240	61.322	48.400	22.878	16.335	17.378	17.635
1429	88.472	62.049	49.105	23.267	16.571	17.604	17.849
1430	88.699	62.771	49.809	23.661	16.811	17.834	18.066
1431	88.921	63.487	50.512	24.061	17.054	18.068	18.285
1432	89.137	64.196	51.214	24.466	17.301	18.304	18.508
1433	89.349	64.899	51.914	24.876	17.550	18.544	18.733
1434	89.556	65.595	52.612	25.292	17.803	18.788	18.961
1435	89.759	66.285	53.308	25.714	18.059	19.035	19.192
1436	89.957	66.967	54.002	26.140	18.319	19.285	19.426
1437	90.150	67.642	54.694	26.572	18.581	19.538	19.662
1438	90.339	68.310	55.383	27.009	18.847	19.795	19.902
1439	90.524	68.970	56.069	27.451	19.117	20.056	20.144
1440	90.705	69.623	56.753	27.898	19.389	20.320	20.390
1441	90.882	70.267	57.433	28.350	19.665	20.587	20.638
1442	91.054	70.904	58.110	28.808	19.944	20.858	20.890
1443	91.223	71.533	58.784	29.269	20.227	21.132	21.144
1444	91.388	72.153	59.454	29.736	20.512	21.410	21.401
1445	91.550	72.765	60.120	30.207	20.801	21.691	21.661
1446	91.708	73.368	60.782	30.683	21.093	21.976	21.924
1447	91.862	73.963	61.439	31.162	21.388	22.264	22.190
1448	92.013	74.550	62.093	31.647	21.687	22.556	22.459
1449	92.160	75.127	62.742	32.135	21.988	22.851	22.731
1450	92.304	75.696	63.386	32.627	22.293	23.149	23.006
1451	92.445	76.257	64.026	33.123	22.601	23.451	23.284
1452	92.583	76.808	64.660	33.623	22.911	23.756	23.564
1453	92.717	77.350	65.290	34.126	23.225	24.064	23.848
1454	92.849	77.884	65.914	34.633	23.542	24.376	24.134
1455	92.978	78.409	66.533	35.143	23.861	24.691	24.423
1456	93.103	78.924	67.146	35.656	24.184	25.009	24.714
1457	93.226	79.431	67.754	36.171	24.509	25.331	25.009
1458	93.347	79.929	68.356	36.690	24.837	25.655	25.306
1459	93.464	80.418	68.952	37.211	25.168	25.983	25.606
1460	93.579	80.898	69.542	37.734	25.502	26.314	25.908
1461	93.692	81.369	70.126	38.260	25.838	26.648	26.213

Table 8.22 (continuation eight)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1462	93.801	81.831	70.704	38.787	26.177	26.984	26.520
1463	93.909	82.284	71.276	39.317	26.518	27.324	26.830
1464	94.014	82.729	71.841	39.848	26.862	27.666	27.142
1465	94.117	83.164	72.400	40.380	27.208	28.011	27.456
1466	94.218	83.591	72.953	40.914	27.557	28.359	27.773
1467	94.316	84.010	73.499	41.449	27.908	28.709	28.092
1468	94.412	84.420	74.038	41.985	28.261	29.062	28.413
1469	94.506	84.821	74.570	42.522	28.616	29.418	28.736
1470	94.599	85.214	75.096	43.059	28.973	29.776	29.061
1471	94.689	85.598	75.614	43.596	29.333	30.136	29.389
1472	94.777	85.974	76.126	44.134	29.694	30.499	29.718
1473	94.863	86.342	76.631	44.672	30.058	30.864	30.048
1474	94.948	86.702	77.128	45.210	30.423	31.231	30.381
1475	95.031	87.053	77.619	45.748	30.790	31.600	30.715
1476	95.112	87.397	78.102	46.286	31.159	31.971	31.051
1477	95.191	87.733	78.578	46.823	31.530	32.344	31.389
1478	95.269	88.061	79.047	47.359	31.902	32.718	31.728
1479	95.345	88.381	79.508	47.895	32.276	33.095	32.069
1480	95.420	88.694	79.962	48.429	32.652	33.473	32.411
1481	95.493	89.000	80.409	48.963	33.029	33.853	32.754
1482	95.565	89.298	80.848	49.495	33.407	34.234	33.099
1483	95.635	89.589	81.280	50.027	33.787	34.617	33.445
1484	95.703	89.873	81.704	50.556	34.169	35.002	33.791
1485	95.771	90.150	82.121	51.085	34.551	35.387	34.140
1486	95.837	90.420	82.531	51.611	34.935	35.774	34.489
1487	95.902	90.683	82.933	52.136	35.320	36.162	34.839
1488	95.965	90.940	83.328	52.659	35.706	36.551	35.190
1489	96.027	91.190	83.715	53.181	36.093	36.941	35.542
1490	96.089	91.434	84.095	53.700	36.481	37.332	35.895
1491	96.148	91.672	84.467	54.217	36.870	37.724	36.248
1492	96.207	91.903	84.832	54.732	37.260	38.117	36.603
1493	96.265	92.129	85.190	55.245	37.651	38.511	36.958
1494	96.321	92.348	85.541	55.756	38.042	38.905	37.314
1495	96.377	92.562	85.884	56.264	38.435	39.300	37.670
1496	96.431	92.770	86.220	56.770	38.828	39.696	38.027
1497	96.485	92.973	86.549	57.273	39.222	40.092	38.385
1498	96.537	93.170	86.871	57.774	39.617	40.489	38.743
1499	96.589	93.362	87.186	58.272	40.012	40.886	39.102
1500	96.640	93.549	87.494	58.767	40.407	41.284	39.461

Table 8.22 (continuation nine)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1501	96.689	93.731	87.795	59.259	40.804	41.682	39.820
1502	96.738	93.908	88.090	59.749	41.200	42.080	40.180
1503	96.786	94.080	88.377	60.236	41.598	42.478	40.541
1504	96.833	94.247	88.658	60.720	41.995	42.877	40.901
1505	96.879	94.410	88.933	61.201	42.393	43.276	41.262
1506	96.925	94.568	89.201	61.679	42.791	43.675	41.624
1507	96.970	94.722	89.463	62.154	43.189	44.073	41.985
1508	97.014	94.872	89.719	62.626	43.588	44.472	42.347
1509	97.057	95.017	89.968	63.095	43.987	44.871	42.709
1510	97.099	95.159	90.212	63.560	44.386	45.270	43.071
1511	97.141	95.296	90.449	64.023	44.785	45.669	43.433
1512	97.182	95.430	90.681	64.482	45.184	46.068	43.795
1513	97.222	95.560	90.907	64.938	45.583	46.466	44.158
1514	97.262	95.686	91.127	65.390	45.982	46.864	44.521
1515	97.301	95.809	91.342	65.839	46.381	47.262	44.883
1516	97.339	95.928	91.552	66.285	46.780	47.660	45.246
1517	97.377	96.044	91.756	66.727	47.178	48.058	45.608
1518	97.414	96.157	91.955	67.166	47.577	48.455	45.971
1519	97.450	96.266	92.149	67.602	47.975	48.851	46.334
1520	97.486	96.372	92.339	68.034	48.373	49.247	46.696
1521	97.521	96.476	92.523	68.462	48.770	49.643	47.059
1522	97.556	96.576	92.702	68.887	49.168	50.039	47.421
1523	97.590	96.674	92.877	69.308	49.564	50.433	47.783
1524	97.624	96.768	93.048	69.726	49.961	50.828	48.146
1525	97.657	96.860	93.214	70.140	50.357	51.221	48.507
1526	97.689	96.950	93.376	70.550	50.752	51.614	48.869
1527	97.722	97.036	93.533	70.957	51.147	52.007	49.230
1528	97.753	97.121	93.687	71.360	51.541	52.398	49.592
1529	97.784	97.203	93.836	71.760	51.934	52.789	49.953
1530	97.815	97.282	93.982	72.155	52.327	53.180	50.313
1531	97.845	97.359	94.124	72.547	52.719	53.569	50.673
1532	97.874	97.434	94.262	72.935	53.111	53.958	51.033
1533	97.904	97.507	94.396	73.319	53.501	54.346	51.393
1534	97.932	97.578	94.527	73.700	53.891	54.733	51.751
1535	97.961	97.647	94.655	74.077	54.279	55.119	52.110
1536	97.988	97.714	94.779	74.450	54.667	55.504	52.468
1537	98.016	97.778	94.900	74.819	55.054	55.888	52.826
1538	98.043	97.841	95.018	75.184	55.440	56.271	53.182
1539	98.070	97.903	95.133	75.546	55.825	56.653	53.539

Table 8.22 (continuation 10)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1540	98.096	97.962	95.245	75.904	56.208	57.034	53.895
1541	98.122	98.020	95.354	76.258	56.591	57.414	54.250
1542	98.147	98.076	95.460	76.608	56.972	57.793	54.604
1543	98.172	98.130	95.563	76.954	57.352	58.170	54.958
1544	98.197	98.183	95.664	77.296	57.731	58.547	55.311
1545	98.221	98.234	95.762	77.635	58.109	58.922	55.663
1546	98.245	98.284	95.858	77.969	58.485	59.296	56.014
1547	98.269	98.332	95.951	78.300	58.860	59.669	56.365
1548	98.292	98.380	96.041	78.627	59.234	60.040	56.714
1549	98.315	98.425	96.130	78.950	59.606	60.410	57.063
1550	98.337	98.469	96.216	79.270	59.977	60.779	57.411
1551	98.360	98.513	96.300	79.585	60.346	61.146	57.758
1552	98.382	98.554	96.382	79.897	60.714	61.512	58.103
1553	98.403	98.595	96.461	80.205	61.080	61.876	58.448
1554	98.425	98.634	96.539	80.509	61.444	62.239	58.792
1555	98.446	98.673	96.615	80.810	61.807	62.600	59.135
1556	98.466	98.710	96.688	81.106	62.168	62.960	59.476
1557	98.487	98.746	96.760	81.399	62.528	63.318	59.816
1558	98.507	98.781	96.830	81.689	62.886	63.675	60.155
1559	98.527	98.815	96.899	81.974	63.242	64.030	60.493
1560	98.546	98.848	96.965	82.256	63.596	64.383	60.830
1561	98.565	98.881	97.030	82.534	63.948	64.735	61.165
1562	98.584	98.912	97.094	82.808	64.299	65.084	61.499
1563	98.603	98.942	97.156	83.079	64.648	65.432	61.831
1564	98.622	98.972	97.216	83.346	64.994	65.779	62.163
1565	98.640	99.000	97.275	83.609	65.339	66.123	62.492
1566	98.658	99.028	97.332	83.869	65.682	66.466	62.821
1567	98.675	99.055	97.388	84.126	66.023	66.807	63.148
1568	98.693	99.081	97.443	84.378	66.362	67.146	63.473
1569	98.710	99.107	97.496	84.628	66.699	67.483	63.797
1570	98.727	99.132	97.548	84.873	67.034	67.818	64.119
1571	98.744	99.156	97.598	85.116	67.366	68.152	64.440
1572	98.760	99.179	97.648	85.355	67.697	68.483	64.759
1573	98.776	99.202	97.696	85.590	68.025	68.813	65.076
1574	98.792	99.224	97.743	85.822	68.352	69.140	65.392
1575	98.808	99.245	97.789	86.050	68.676	69.466	65.706
1576	98.824	99.266	97.834	86.276	68.998	69.789	66.018
1577	98.839	99.287	97.878	86.498	69.317	70.110	66.329
1578	98.854	99.306	97.921	86.716	69.635	70.430	66.638



Table 8.22 (continuation 11)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1579	98.869	99.325	97.963	86.932	69.950	70.747	66.945
1580	98.884	99.344	98.004	87.144	70.263	71.062	67.250
1581	98.899	99.362	98.044	87.353	70.573	71.375	67.553
1582	98.913	99.380	98.083	87.559	70.882	71.686	67.855
1583	98.927	99.397	98.121	87.761	71.188	71.995	68.155
1584	98.941	99.413	98.158	87.961	71.491	72.302	68.452
1585	98.955	99.429	98.194	88.157	71.792	72.607	68.748
1586	98.968	99.445	98.230	88.351	72.091	72.909	69.042
1587	98.982	99.460	98.264	88.541	72.388	73.210	69.335
1588	98.995	99.475	98.298	88.728	72.682	73.508	69.625
1589	99.008	99.489	98.331	88.913	72.974	73.804	69.913
1590	99.021	99.503	98.364	89.094	73.263	74.098	70.199
1591	99.034	99.517	98.395	89.273	73.550	74.389	70.483
1592	99.046	99.530	98.426	89.449	73.834	74.679	70.766
1593	99.059	99.543	98.457	89.622	74.116	74.966	71.046
1594	99.071	99.555	98.486	89.792	74.396	75.251	71.324
1595	99.083	99.567	98.515	89.959	74.673	75.534	71.600
1596	99.095	99.579	98.543	90.124	74.948	75.814	71.874
1597	99.106	99.590	98.571	90.286	75.220	76.093	72.147
1598	99.118	99.602	98.598	90.445	75.490	76.369	72.417
1599	99.129	99.612	98.625	90.602	75.757	76.643	72.685
1600	99.140	99.623	98.650	90.756	76.022	76.914	72.951
1601	99.151	99.633	98.676	90.908	76.284	77.184	73.214
1602	99.162	99.643	98.701	91.057	76.544	77.451	73.476
1603	99.173	99.653	98.725	91.203	76.801	77.716	73.736
1604	99.184	99.662	98.749	91.348	77.056	77.979	73.994
1605	99.194	99.671	98.772	91.489	77.309	78.239	74.249
1606	99.205	99.680	98.794	91.629	77.559	78.498	74.503
1607	99.215	99.689	98.817	91.766	77.806	78.754	74.754
1608	99.225	99.697	98.838	91.900	78.051	79.007	75.003
1609	99.235	99.705	98.860	92.033	78.294	79.259	75.251
1610	99.245	99.713	98.881	92.163	78.534	79.508	75.496
1611	99.254	99.721	98.901	92.291	78.772	79.756	75.739
1612	99.264	99.728	98.921	92.417	79.007	80.000	75.980
1613	99.273	99.736	98.941	92.540	79.240	80.243	76.218
1614	99.283	99.743	98.960	92.662	79.470	80.484	76.455
1615	99.292	99.750	98.979	92.781	79.698	80.722	76.690
1616	99.301	99.756	98.997	92.899	79.924	80.958	76.922
1617	99.310	99.763	99.015	93.014	80.147	81.192	77.153

Table 8.22 (continuation 12)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1618	99.319	99.769	99.033	93.127	80.368	81.423	77.381
1619	99.327	99.775	99.050	93.239	80.586	81.653	77.608
1620	99.336	99.781	99.067	93.348	80.802	81.880	77.832
1621	99.344	99.787	99.083	93.456	81.016	82.105	78.054
1622	99.353	99.793	99.100	93.561	81.227	82.328	78.275
1623	99.361	99.798	99.116	93.665	81.436	82.549	78.493
1624	99.369	99.804	99.131	93.767	81.643	82.767	78.709
1625	99.377	99.809	99.147	93.868	81.848	82.984	78.923
1626	99.385	99.814	99.162	93.966	82.050	83.198	79.135
1627	99.393	99.819	99.176	94.063	82.249	83.410	79.345
1628	99.401	99.824	99.191	94.158	82.447	83.620	79.553
1629	99.408	99.828	99.205	94.252	82.642	83.828	79.759
1630	99.416	99.833	99.219	94.343	82.835	84.034	79.963
1631	99.423	99.837	99.232	94.434	83.026	84.238	80.165
1632	99.431	99.842	99.245	94.522	83.215	84.439	80.365
1633	99.438	99.846	99.258	94.609	83.401	84.639	80.563
1634	99.445	99.850	99.271	94.695	83.586	84.836	80.759
1635	99.452	99.854	99.284	94.779	83.768	85.031	80.954
1636	99.459	99.858	99.296	94.861	83.948	85.225	81.146
1637	99.466	99.861	99.308	94.943	84.126	85.416	81.336
1638	99.473	99.865	99.320	95.022	84.301	85.605	81.525
1639	99.479	99.869	99.332	95.101	84.475	85.792	81.711
1640	99.486	99.872	99.343	95.178	84.647	85.977	81.896
1641	99.492	99.875	99.354	95.253	84.816	86.160	82.078
1642	99.499	99.879	99.365	95.327	84.984	86.341	82.259
1643	99.505	99.882	99.376	95.400	85.150	86.520	82.438
1644	99.512	99.885	99.386	95.472	85.313	86.697	82.615
1645	99.518	99.888	99.397	95.542	85.475	86.872	82.790
1646	99.524	99.891	99.407	95.612	85.634	87.045	82.963
1647	99.530	99.894	99.417	95.680	85.792	87.216	83.135
1648	99.536	99.897	99.427	95.746	85.948	87.386	83.304
1649	99.542	99.899	99.436	95.812	86.102	87.553	83.472
1650	99.547	99.902	99.446	95.876	86.254	87.718	83.638
1651	99.553	99.905	99.455	95.940	86.404	87.882	83.802
1652	99.559	99.907	99.464	96.002	86.552	88.043	83.965
1653	99.564	99.909	99.473	96.063	86.699	88.203	84.125
1654	99.570	99.912	99.482	96.123	86.844	88.360	84.284
1655	99.575	99.914	99.490	96.182	86.987	88.516	84.441
1656	99.581	99.916	99.499	96.240	87.128	88.670	84.597

Table 8.22 (continuation 13)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1657	99.586	99.919	99.507	96.297	87.267	88.823	84.751
1658	99.591	99.921	99.515	96.353	87.405	88.973	84.903
1659	99.596	99.923	99.523	96.408	87.541	89.121	85.053
1660	99.601	99.925	99.531	96.462	87.675	89.268	85.201
1661	99.606	99.927	99.539	96.515	87.808	89.413	85.348
1662	99.611	99.929	99.546	96.568	87.939	89.556	85.493
1663	99.616	99.930	99.553	96.619	88.068	89.698	85.637
1664	99.621	99.932	99.561	96.669	88.196	89.837	85.779
1665	99.626	99.934	99.568	96.719	88.322	89.975	85.919
1666	99.631	99.936	99.575	96.768	88.447	90.111	86.058
1667	99.635	99.937	99.582	96.815	88.570	90.246	86.195
1668	99.640	99.939	99.589	96.862	88.692	90.379	86.330
1669	99.644	99.941	99.595	96.909	88.812	90.510	86.464
1670	99.649	99.942	99.602	96.954	88.930	90.639	86.596
1671	99.653	99.944	99.608	96.999	89.048	90.767	86.727
1672	99.658	99.945	99.615	97.043	89.163	90.893	86.856
1673	99.662	99.946	99.621	97.086	89.277	91.018	86.984
1674	99.666	99.948	99.627	97.128	89.390	91.141	87.110
1675	99.670	99.949	99.633	97.170	89.502	91.262	87.234
1676	99.674	99.950	99.639	97.211	89.612	91.382	87.357
1677	99.678	99.952	99.644	97.251	89.720	91.500	87.479
1678	99.682	99.953	99.650	97.291	89.828	91.616	87.599
1679	99.686	99.954	99.656	97.329	89.933	91.731	87.718
1680	99.690	99.955	99.661	97.368	90.038	91.845	87.835
1681	99.694	99.957	99.667	97.405	90.141	91.957	87.951
1682	99.698	99.958	99.672	97.442	90.244	92.068	88.065
1683	99.702	99.959	99.677	97.479	90.344	92.177	88.178
1684	99.706	99.960	99.682	97.514	90.444	92.284	88.290
1685	99.709	99.961	99.687	97.549	90.542	92.390	88.400
1686	99.713	99.962	99.692	97.584	90.639	92.495	88.509
1687	99.717	99.963	99.697	97.618	90.735	92.598	88.616
1688	99.720	99.964	99.702	97.651	90.830	92.700	88.723
1689	99.724	99.965	99.706	97.684	90.924	92.801	88.827
1690	99.727	99.966	99.711	97.716	91.016	92.900	88.931
1691	99.730	99.966	99.716	97.748	91.107	92.997	89.033
1692	99.734	99.967	99.720	97.779	91.197	93.094	89.134
1693	99.737	99.968	99.724	97.810	91.286	93.189	89.234
1694	99.740	99.969	99.729	97.840	91.374	93.283	89.332
1695	99.744	99.970	99.733	97.870	91.461	93.375	89.429

Table 8.22 (continuation 14)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1696	99.747	99.971	99.737	97.899	91.547	93.466	89.525
1697	99.750	99.971	99.741	97.928	91.632	93.556	89.620
1698	99.753	99.972	99.745	97.956	91.715	93.645	89.713
1699	99.756	99.973	99.749	97.984	91.798	93.732	89.806
1700	99.759	99.973	99.753	98.011	91.880	93.818	89.897
1701	N/A	N/A	N/A	98.038	91.961	93.903	89.987
1702	N/A	N/A	N/A	98.064	92.040	93.987	90.076
1703	N/A	N/A	N/A	98.090	92.119	94.070	90.164
1704	N/A	N/A	N/A	98.116	92.197	94.151	90.250
1705	N/A	N/A	N/A	98.141	92.274	94.231	90.336
1706	N/A	N/A	N/A	98.166	92.350	94.310	90.420
1707	N/A	N/A	N/A	98.190	92.425	94.388	90.503
1708	N/A	N/A	N/A	98.214	92.499	94.465	90.586
1709	N/A	N/A	N/A	98.238	92.572	94.541	90.667
1710	N/A	N/A	N/A	98.261	92.645	94.616	90.747
1711	N/A	N/A	N/A	98.284	92.716	94.689	90.826
1712	N/A	N/A	N/A	98.306	92.787	94.762	90.904
1713	N/A	N/A	N/A	98.328	92.856	94.833	90.982
1714	N/A	N/A	N/A	98.350	92.925	94.904	91.058
1715	N/A	N/A	N/A	98.372	92.994	94.973	91.133
1716	N/A	N/A	N/A	98.393	93.061	95.042	91.207
1717	N/A	N/A	N/A	98.413	93.127	95.109	91.280
1718	N/A	N/A	N/A	98.434	93.193	95.176	91.353
1719	N/A	N/A	N/A	98.454	93.258	95.242	91.424
1720	N/A	N/A	N/A	98.474	93.322	95.306	91.495
1721	N/A	N/A	N/A	98.493	93.386	95.370	91.564
1722	N/A	N/A	N/A	98.512	93.449	95.433	91.633
1723	N/A	N/A	N/A	98.531	93.511	95.495	91.701
1724	N/A	N/A	N/A	98.550	93.572	95.555	91.768
1725	N/A	N/A	N/A	98.568	93.633	95.616	91.834
1726	N/A	N/A	N/A	98.586	93.692	95.675	91.899
1727	N/A	N/A	N/A	98.604	93.752	95.733	91.964
1728	N/A	N/A	N/A	98.621	93.810	95.791	92.027
1729	N/A	N/A	N/A	98.638	93.868	95.847	92.090
1730	N/A	N/A	N/A	98.655	93.925	95.903	92.152
1731	N/A	N/A	N/A	98.672	93.981	95.958	92.213
1732	N/A	N/A	N/A	98.688	94.037	96.012	92.274
1733	N/A	N/A	N/A	98.704	94.092	96.066	92.333
1734	N/A	N/A	N/A	98.720	94.147	96.119	92.392

Table 8.22 (continuation 15)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1735	N/A	N/A	N/A	98.736	94.201	96.170	92.451
1736	N/A	N/A	N/A	98.751	94.254	96.222	92.508
1737	N/A	N/A	N/A	98.766	94.307	96.272	92.565
1738	N/A	N/A	N/A	98.781	94.359	96.322	92.621
1739	N/A	N/A	N/A	98.796	94.411	96.371	92.676
1740	N/A	N/A	N/A	98.810	94.462	96.419	92.731
1741	N/A	N/A	N/A	98.825	94.512	96.466	92.785
1742	N/A	N/A	N/A	98.839	94.562	96.513	92.838
1743	N/A	N/A	N/A	98.853	94.611	96.559	92.891
1744	N/A	N/A	N/A	98.866	94.660	96.605	92.943
1745	N/A	N/A	N/A	98.880	94.708	96.650	92.994
1746	N/A	N/A	N/A	98.893	94.756	96.694	93.045
1747	N/A	N/A	N/A	98.906	94.803	96.738	93.095
1748	N/A	N/A	N/A	98.919	94.849	96.780	93.144
1749	N/A	N/A	N/A	98.932	94.896	96.823	93.193
1750	N/A	N/A	N/A	98.944	94.941	96.864	93.241
1751	N/A	N/A	N/A	98.956	94.986	96.906	93.289
1752	N/A	N/A	N/A	98.968	95.031	96.946	93.336
1753	N/A	N/A	N/A	98.980	95.075	96.986	93.382
1754	N/A	N/A	N/A	98.992	95.118	97.025	93.428
1755	N/A	N/A	N/A	99.004	95.162	97.064	93.473
1756	N/A	N/A	N/A	99.015	95.204	97.102	93.518
1757	N/A	N/A	N/A	99.026	95.246	97.140	93.563
1758	N/A	N/A	N/A	99.037	95.288	97.177	93.606
1759	N/A	N/A	N/A	99.048	95.329	97.213	93.649
1760	N/A	N/A	N/A	99.059	95.370	97.249	93.692
1761	N/A	N/A	N/A	99.070	95.411	97.285	93.734
1762	N/A	N/A	N/A	99.080	95.451	97.320	93.776
1763	N/A	N/A	N/A	99.091	95.490	97.355	93.817
1764	N/A	N/A	N/A	99.101	95.529	97.389	93.858
1765	N/A	N/A	N/A	99.111	95.568	97.422	93.898
1766	N/A	N/A	N/A	99.121	95.607	97.455	93.938
1767	N/A	N/A	N/A	99.131	95.644	97.488	93.977
1768	N/A	N/A	N/A	99.140	95.682	97.520	94.015
1769	N/A	N/A	N/A	99.150	95.719	97.551	94.054
1770	N/A	N/A	N/A	99.159	95.756	97.583	94.092
1771	N/A	N/A	N/A	99.168	95.792	97.613	94.129
1772	N/A	N/A	N/A	99.177	95.828	97.644	94.166
1773	N/A	N/A	N/A	99.186	95.864	97.673	94.202

Table 8.22 (continuation 16)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1774	N/A	N/A	N/A	99.195	95.899	97.703	94.239
1775	N/A	N/A	N/A	99.204	95.934	97.732	94.274
1776	N/A	N/A	N/A	99.212	95.968	97.760	94.309
1777	N/A	N/A	N/A	99.221	96.002	97.789	94.344
1778	N/A	N/A	N/A	99.229	96.036	97.816	94.379
1779	N/A	N/A	N/A	99.238	96.070	97.844	94.413
1780	N/A	N/A	N/A	99.246	96.103	97.871	94.446
1781	N/A	N/A	N/A	99.254	96.136	97.898	94.480
1782	N/A	N/A	N/A	99.262	96.168	97.924	94.513
1783	N/A	N/A	N/A	99.269	96.200	97.950	94.545
1784	N/A	N/A	N/A	99.277	96.232	97.975	94.577
1785	N/A	N/A	N/A	99.285	96.263	98.000	94.609
1786	N/A	N/A	N/A	99.292	96.294	98.025	94.640
1787	N/A	N/A	N/A	99.300	96.325	98.050	94.671
1788	N/A	N/A	N/A	99.307	96.356	98.074	94.702
1789	N/A	N/A	N/A	99.314	96.386	98.098	94.732
1790	N/A	N/A	N/A	99.321	96.416	98.121	94.762
1791	N/A	N/A	N/A	99.328	96.446	98.144	94.792
1792	N/A	N/A	N/A	99.335	96.475	98.167	94.821
1793	N/A	N/A	N/A	99.342	96.504	98.189	94.850
1794	N/A	N/A	N/A	99.349	96.533	98.212	94.879
1795	N/A	N/A	N/A	99.355	96.561	98.233	94.907
1796	N/A	N/A	N/A	99.362	96.589	98.255	94.935
1797	N/A	N/A	N/A	99.368	96.617	98.276	94.963
1798	N/A	N/A	N/A	99.375	96.645	98.297	94.991
1799	N/A	N/A	N/A	99.381	96.672	98.318	95.018
1800	N/A	N/A	N/A	99.387	96.699	98.338	95.044
1801	N/A	N/A	N/A	N/A	96.726	98.358	95.071
1802	N/A	N/A	N/A	N/A	96.753	98.378	95.097
1803	N/A	N/A	N/A	N/A	96.779	98.398	95.123
1804	N/A	N/A	N/A	N/A	96.805	98.417	95.149
1805	N/A	N/A	N/A	N/A	96.831	98.436	95.174
1806	N/A	N/A	N/A	N/A	96.856	98.455	95.199
1807	N/A	N/A	N/A	N/A	96.882	98.473	95.224
1808	N/A	N/A	N/A	N/A	96.907	98.491	95.249
1809	N/A	N/A	N/A	N/A	96.932	98.509	95.273
1810	N/A	N/A	N/A	N/A	96.956	98.527	95.297
1811	N/A	N/A	N/A	N/A	96.981	98.545	95.321
1812	N/A	N/A	N/A	N/A	97.005	98.562	95.344

Table 8.22 (continuation 17)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1813	N/A	N/A	N/A	N/A	97.029	98.579	95.368
1814	N/A	N/A	N/A	N/A	97.052	98.596	95.391
1815	N/A	N/A	N/A	N/A	97.076	98.612	95.413
1816	N/A	N/A	N/A	N/A	97.099	98.629	95.436
1817	N/A	N/A	N/A	N/A	97.122	98.645	95.458
1818	N/A	N/A	N/A	N/A	97.145	98.661	95.480
1819	N/A	N/A	N/A	N/A	97.167	98.676	95.502
1820	N/A	N/A	N/A	N/A	97.190	98.692	95.524
1821	N/A	N/A	N/A	N/A	97.212	98.707	95.545
1822	N/A	N/A	N/A	N/A	97.234	98.722	95.566
1823	N/A	N/A	N/A	N/A	97.256	98.737	95.587
1824	N/A	N/A	N/A	N/A	97.277	98.752	95.608
1825	N/A	N/A	N/A	N/A	97.299	98.766	95.628
1826	N/A	N/A	N/A	N/A	97.320	98.780	95.648
1827	N/A	N/A	N/A	N/A	97.341	98.794	95.669
1828	N/A	N/A	N/A	N/A	97.362	98.808	95.688
1829	N/A	N/A	N/A	N/A	97.382	98.822	95.708
1830	N/A	N/A	N/A	N/A	97.403	98.835	95.728
1831	N/A	N/A	N/A	N/A	97.423	98.849	95.747
1832	N/A	N/A	N/A	N/A	97.443	98.862	95.766
1833	N/A	N/A	N/A	N/A	97.463	98.875	95.785
1834	N/A	N/A	N/A	N/A	97.483	98.888	95.803
1835	N/A	N/A	N/A	N/A	97.502	98.900	95.822
1836	N/A	N/A	N/A	N/A	97.521	98.913	95.840
1837	N/A	N/A	N/A	N/A	97.541	98.925	95.858
1838	N/A	N/A	N/A	N/A	97.560	98.937	95.876
1839	N/A	N/A	N/A	N/A	97.578	98.949	95.894
1840	N/A	N/A	N/A	N/A	97.597	98.961	95.911
1841	N/A	N/A	N/A	N/A	97.615	98.973	95.929
1842	N/A	N/A	N/A	N/A	97.634	98.984	95.946
1843	N/A	N/A	N/A	N/A	97.652	98.996	95.963
1844	N/A	N/A	N/A	N/A	97.670	99.007	95.980
1845	N/A	N/A	N/A	N/A	97.688	99.018	95.997
1846	N/A	N/A	N/A	N/A	97.706	99.029	96.013
1847	N/A	N/A	N/A	N/A	97.723	99.040	96.030
1848	N/A	N/A	N/A	N/A	97.740	99.050	96.046
1849	N/A	N/A	N/A	N/A	97.758	99.061	96.062
1850	N/A	N/A	N/A	N/A	97.775	99.071	96.078
1851	N/A	N/A	N/A	N/A	97.792	99.082	96.094

Table 8.22 (continuation 18)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1852	N/A	N/A	N/A	N/A	97.808	99.092	96.109
1853	N/A	N/A	N/A	N/A	97.825	99.102	96.125
1854	N/A	N/A	N/A	N/A	97.842	99.112	96.140
1855	N/A	N/A	N/A	N/A	97.858	99.121	96.155
1856	N/A	N/A	N/A	N/A	97.874	99.131	96.170
1857	N/A	N/A	N/A	N/A	97.890	99.140	96.185
1858	N/A	N/A	N/A	N/A	97.906	99.150	96.200
1859	N/A	N/A	N/A	N/A	97.922	99.159	96.214
1860	N/A	N/A	N/A	N/A	97.938	99.168	96.229
1861	N/A	N/A	N/A	N/A	97.953	99.177	96.243
1862	N/A	N/A	N/A	N/A	97.968	99.186	96.257
1863	N/A	N/A	N/A	N/A	97.984	99.195	96.271
1864	N/A	N/A	N/A	N/A	97.999	99.204	96.285
1865	N/A	N/A	N/A	N/A	98.014	99.212	96.299
1866	N/A	N/A	N/A	N/A	98.029	99.221	96.312
1867	N/A	N/A	N/A	N/A	98.043	99.229	96.326
1868	N/A	N/A	N/A	N/A	98.058	99.237	96.339
1869	N/A	N/A	N/A	N/A	98.073	99.245	96.353
1870	N/A	N/A	N/A	N/A	98.087	99.253	96.366
1871	N/A	N/A	N/A	N/A	98.101	99.261	96.379
1872	N/A	N/A	N/A	N/A	98.115	99.269	96.392
1873	N/A	N/A	N/A	N/A	98.129	99.277	96.404
1874	N/A	N/A	N/A	N/A	98.143	99.284	96.417
1875	N/A	N/A	N/A	N/A	98.157	99.292	96.430
1876	N/A	N/A	N/A	N/A	98.171	99.299	96.442
1877	N/A	N/A	N/A	N/A	98.184	99.307	96.454
1878	N/A	N/A	N/A	N/A	98.198	99.314	96.467
1879	N/A	N/A	N/A	N/A	98.211	99.321	96.479
1880	N/A	N/A	N/A	N/A	98.224	99.328	96.491
1881	N/A	N/A	N/A	N/A	98.237	99.335	96.502
1882	N/A	N/A	N/A	N/A	98.251	99.342	96.514
1883	N/A	N/A	N/A	N/A	98.263	99.349	96.526
1884	N/A	N/A	N/A	N/A	98.276	99.356	96.537
1885	N/A	N/A	N/A	N/A	98.289	99.363	96.549
1886	N/A	N/A	N/A	N/A	98.301	99.369	96.560
1887	N/A	N/A	N/A	N/A	98.314	99.376	96.572
1888	N/A	N/A	N/A	N/A	98.326	99.382	96.583
1889	N/A	N/A	N/A	N/A	98.339	99.388	96.594
1890	N/A	N/A	N/A	N/A	98.351	99.395	96.605



Table 8.22 (continuation 19)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1891	N/A	N/A	N/A	N/A	98.363	99.401	96.616
1892	N/A	N/A	N/A	N/A	98.375	99.407	96.626
1893	N/A	N/A	N/A	N/A	98.387	99.413	96.637
1894	N/A	N/A	N/A	N/A	98.399	99.419	96.648
1895	N/A	N/A	N/A	N/A	98.410	99.425	96.658
1896	N/A	N/A	N/A	N/A	98.422	99.431	96.668
1897	N/A	N/A	N/A	N/A	98.433	99.437	96.679
1898	N/A	N/A	N/A	N/A	98.445	99.442	96.689
1899	N/A	N/A	N/A	N/A	98.456	99.448	96.699
1900	N/A	N/A	N/A	N/A	98.467	99.453	96.709
1901	N/A	N/A	N/A	N/A	N/A	99.459	96.719
1902	N/A	N/A	N/A	N/A	N/A	99.464	96.729
1903	N/A	N/A	N/A	N/A	N/A	99.470	96.739
1904	N/A	N/A	N/A	N/A	N/A	99.475	96.748
1905	N/A	N/A	N/A	N/A	N/A	99.480	96.758
1906	N/A	N/A	N/A	N/A	N/A	99.485	96.768
1907	N/A	N/A	N/A	N/A	N/A	99.490	96.777
1908	N/A	N/A	N/A	N/A	N/A	99.496	96.786
1909	N/A	N/A	N/A	N/A	N/A	99.501	96.796
1910	N/A	N/A	N/A	N/A	N/A	99.505	96.805
1911	N/A	N/A	N/A	N/A	N/A	99.510	96.814
1912	N/A	N/A	N/A	N/A	N/A	99.515	96.823
1913	N/A	N/A	N/A	N/A	N/A	99.520	96.832
1914	N/A	N/A	N/A	N/A	N/A	99.525	96.841
1915	N/A	N/A	N/A	N/A	N/A	99.529	96.850
1916	N/A	N/A	N/A	N/A	N/A	99.534	96.859
1917	N/A	N/A	N/A	N/A	N/A	99.538	96.867
1918	N/A	N/A	N/A	N/A	N/A	99.543	96.876
1919	N/A	N/A	N/A	N/A	N/A	99.547	96.885
1920	N/A	N/A	N/A	N/A	N/A	99.552	96.893
1921	N/A	N/A	N/A	N/A	N/A	99.556	96.902
1922	N/A	N/A	N/A	N/A	N/A	99.560	96.910
1923	N/A	N/A	N/A	N/A	N/A	99.565	96.918
1924	N/A	N/A	N/A	N/A	N/A	99.569	96.926
1925	N/A	N/A	N/A	N/A	N/A	99.573	96.935
1926	N/A	N/A	N/A	N/A	N/A	99.577	96.943
1927	N/A	N/A	N/A	N/A	N/A	99.581	96.951
1928	N/A	N/A	N/A	N/A	N/A	99.585	96.959
1929	N/A	N/A	N/A	N/A	N/A	99.589	96.967

Table 8.22 (continuation 20)

<b>Scale Score</b>	<b>Kindergarten</b>	<b>Grade 1</b>	<b>Grade 2</b>	<b>Grades 3–5</b>	<b>Grades 6–8</b>	<b>Grades 9–10</b>	<b>Grades 11–12</b>
1930	N/A	N/A	N/A	N/A	N/A	99.593	96.974
1931	N/A	N/A	N/A	N/A	N/A	99.597	96.982
1932	N/A	N/A	N/A	N/A	N/A	99.601	96.990
1933	N/A	N/A	N/A	N/A	N/A	99.604	96.998
1934	N/A	N/A	N/A	N/A	N/A	99.608	97.005
1935	N/A	N/A	N/A	N/A	N/A	99.612	97.013
1936	N/A	N/A	N/A	N/A	N/A	99.616	97.020
1937	N/A	N/A	N/A	N/A	N/A	99.619	97.028
1938	N/A	N/A	N/A	N/A	N/A	99.623	97.035
1939	N/A	N/A	N/A	N/A	N/A	99.626	97.042
1940	N/A	N/A	N/A	N/A	N/A	99.630	97.050
1941	N/A	N/A	N/A	N/A	N/A	99.633	97.057
1942	N/A	N/A	N/A	N/A	N/A	99.637	97.064
1943	N/A	N/A	N/A	N/A	N/A	99.640	97.071
1944	N/A	N/A	N/A	N/A	N/A	99.643	97.078
1945	N/A	N/A	N/A	N/A	N/A	99.647	97.085
1946	N/A	N/A	N/A	N/A	N/A	99.650	97.092
1947	N/A	N/A	N/A	N/A	N/A	99.653	97.099
1948	N/A	N/A	N/A	N/A	N/A	99.656	97.106
1949	N/A	N/A	N/A	N/A	N/A	99.659	97.113
1950	N/A	N/A	N/A	N/A	N/A	99.663	97.119

## Chapter 9: Validity

---

This chapter summarizes validity evidence to support the use and score interpretation of the Summative English Language Proficiency Assessments for California (ELPAC). It includes material on content validity and evidence of fairness and reliability.

### 9.1. Validity of the ELPAC Test Design

The Summative ELPAC was developed in accordance with the criteria for test development, administration, and use described in the *Standards for Educational and Psychological Testing* (2014) adopted by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME).

Test validation is an ongoing process, beginning at initial conceptualization and continuing throughout the lifetime of the assessment. Every aspect of an assessment provides evidence in support of its validity—or evidence to the contrary—including design, content requirements, item development, and psychometric characteristics. “Validity refers to the degree to which evidence and theory support the interpretations made from test scores. Validity is, therefore, the most fundamental consideration in developing and evaluating tests. The process of validation involves accumulating evidence to provide a sound, scientific basis for the proposed score interpretations” (AERA, APA, & NCME, 2014, p. 9).

#### 9.1.1. Purpose of the ELPAC

The Summative ELPAC was designed and developed to provide scores representing English language proficiency (ELP) performance levels for required educational decision making as defined by the test purposes in the California *Education Code (EC)* Section 313. The primary inferences from the test results, in general, include (a) the proficiency level of individual students and (b) English language development (ELD) program effectiveness based on the results of groups of students.

Progress can be tracked over years and grades. The results can be used to analyze the strengths and weaknesses of students’ growth in the four domains measured and to report progress to parents. The results can also be used as one body of evidence in making administrative decisions about ELD program effectiveness, class grouping, needs assessment, and placement in English learner (EL) programs.

#### 9.1.2. The Constructs to Be Measured

The Summative ELPAC is designed to show how well students perform relative to the 2012 *California English Language Development Standards, Kindergarten Through Grade 12* (2012 ELD Standards) (California Department of Education [CDE], 2014). The standards describe the ELP knowledge, skills, and abilities that students are expected to acquire at each grade. The Summative ELPAC test blueprints describe the assessment task types that the students perform, the number of items per task type, and the alignment of the items to the 2012 ELD Standards (CDE, 2017).

*EC* Section 60810 specifies that the state ELP assessment shall measure the language domains of Listening, Speaking, Reading, and Writing. The test blueprints describe the assessment task types and the number of items that are used to assess students’ ELP in each language domain.

The Summative ELPAC provides three scale scores and placement within one of four levels. The oral language composite scale score and reporting level are drawn from the Listening and Speaking results. The written language composite scale score and reporting level are drawn from the Reading and Writing results. The overall scale score and reporting level are derived from the oral language composite and written language composite.

At grades one through twelve, the overall scale score is derived from the equal weighting of the oral language composite and the written language composite. At kindergarten, where students are developing foundational literacy skills, the overall scale score is derived from differential weighting in which 70 percent of the overall scale score comes from the oral language composite and 30 percent of the overall scale score comes from the written language composite.

In addition, assessment results are used to place students within one of three levels in each of the four domains of Listening, Speaking, Reading, and Writing.

### **9.1.3. The Interpretations and Uses of the Scores**

Student scores were delivered to local educational agencies and used as one criterion for considering whether a student will be reclassified as fluent English proficient. *EC* Section 313(f) describes four criteria that are used to establish reclassification policies and procedures:

1. Assessment of language proficiency using an objective assessment instrument, including, but not limited to, the ELPAC
2. Teacher evaluation, including, but not limited to, a review of the student's curriculum mastery
3. Parental opinion and consultation
4. Comparison of student performance in basic skills against an empirically established range of performance in basic skills based upon the performance of English proficient students of the same age

Summative ELPAC data also is used to calculate the English Language Progress Indicator for the California School Dashboard. ELPAC results are now used for federal accountability as required by Title I.

### **9.1.4. The Intended Population**

The ELPAC is the required state test for ELP that must be given to students whose primary language is a language other than English. The test-taking population for the Summative ELPAC includes students who have been formally identified as ELs in kindergarten through grade twelve based upon the results from the initial assessment, which was the Initial ELPAC during the 2018–2019 academic year. ELs continue to take the Summative ELPAC each year to monitor their ELP until they are reclassified as fluent English proficient.

Students with disabilities who cannot take one or more domains of the ELPAC with allowed universal tools, designated supports, or accommodations take a locally determined alternate assessment(s), as noted in their individualized education program.

## 9.2. Validity of the ELPAC Test Content

### 9.2.1. Description of the State Standards

The 2012 ELD Standards were developed and approved by the California State Board of Education in 2012 and then published in 2014. The 2012 ELD Standards describe the key knowledge, skills, and abilities that students who are learning English need to access, engage with, and achieve in grade-level academic content. The 2012 ELD Standards provide a framework to guide the development of ELD assessment systems that help California educators ensure that all ELs make progress in the English language knowledge, skills, and abilities needed to become college- and career-ready.

### 9.2.2. Item Writing Guidelines

Item writing guidelines were developed to define the task types and content of the items. They were used as a key reference document during item writer training to provide guidance to item writers and drive consistency and efficiency in item development (CDE, 2016). The *Item Writing Guidelines for the ELPAC* were intended to facilitate the development of comparable items that measure appropriate skills and content aligned with the 2012 ELD Standards.

### 9.2.3. Test Blueprints

Test blueprints describe the content of the Summative ELPAC and include four tables with information about the task types in each of the four language domains of Listening, Speaking, Reading, and Writing. Task types are individual items or sets of items that require a student to perform an activity to elicit information about the student's ELP.

The test blueprints provide information about the number of items and points that were administered per task type within each grade level and domain. The test blueprints also provide the alignment of task types with the 2012 ELD Standards (CDE, 2017).

### 9.2.4. Form Assembly Process

The form assembly process began with the creation of test development specifications, which described the content characteristics, psychometric characteristics, and quantity of items to be used in the 2018–2019 Summative ELPAC. Educational Testing Service (ETS) created the test development specifications that the CDE then reviewed and approved.

After the test development specifications were approved, ETS assessment specialists assembled the tests in the ETS Item Banking Information System (IBIS) according to the specifications. IBIS was then used to generate form planners, which are spreadsheets that contain essential item information such as the number of items, alignment of items according to the 2012 ELD Standards, and keys to multiple-choice items. ETS assessment specialists and psychometricians reviewed the form planners before they were delivered to the CDE. CDE staff reviewed the form planners. After ETS made any necessary edits, the CDE approved the form planners. After approval, the form planners were used as the basis for developing the test materials needed to administer the Summative ELPAC: *Examiner's Manuals*, Test Books, Answer Books, and audio recordings.

## 9.3. Validity of the ELPAC's Internal Structure

Internal structure evidence evaluates the strength or salience of the major dimensions underlying an assessment using indices of measurement precision such as fairness and differential item functioning (DIF) analysis, test reliability, and reliability of performance classifications.

### 9.3.1. Fairness and Differential Item Functioning

#### 9.3.1.1. Bias and Sensitivity Reviews

To develop test materials that are fair and unbiased to all students, ELPAC test items underwent reviews by Bias and Sensitivity Review Panels. The first set of approximately 2,000 ELPAC test items was reviewed by a Bias and Sensitivity Review Panel from August 3 through August 5, 2016, and the second set of about 200 items was reviewed from February 21 through February 22, 2018. California educators reviewed the text and artwork of the newly developed items, and each item was either approved as is, approved with revisions, or rejected. As described in section [3.2.6 Item Review Panels](#), the educators added value to the item pool by revising items to make them fair and unbiased measures of ELP.

#### 9.3.1.2. Differential Item Functioning

DIF analyses were conducted to identify differences in item performance by student gender (male and female) and ethnicity, for students matched on ability. There was one item identified as having significant levels of DIF for any domain. Refer to subsection [8.4 Differential Item Functioning \(DIF\)](#) for a description the DIF analyses and for the results of the DIF analyses performed on Summative ELPAC items.

### 9.3.2. Reliability

#### 9.3.2.1. Overall Reliability Estimates

The results of reliability analyses on the four domains and two composite scores are presented in [table 8.16](#). The results indicate that the reliability estimates for each domain of the test were moderately high, ranging from 0.79 to 0.95 across grade level or grade span. For the oral and written composite scores, the reliability estimates were high, ranging from 0.86 to 0.94 across grade level or grade span. Overall reliability estimates were high, ranging from 0.94 to 0.96.

#### 9.3.2.2. Reliability Estimates for Student Groups

The reliabilities are also computed for various student groups. The student groups considered were based on gender (male and female), ethnicity, economic status, migrant status, and special education service status. Reliability estimates for each domain and composite scores are reported for each student group in table 8.B.1 through table 8.B.7. The reliability estimates for each student group showed a similar pattern as the reliability estimates for overall students.

#### 9.3.2.3. Interrater Reliability

Interrater reliability is evaluated by computing the percentage of exact agreement between two raters. Refer to subsection [8.5.4 Writing Score Reliability](#) for a description of agreement analysis and to [appendix 8.E](#), where the results of the analyses are reported.

#### 9.3.2.4. Reliability of Performance Classifications

The methodology used for estimating the reliability of classification decisions is evaluated with the decision classification analyses in subsection [8.5.5 Decision Classification Analyses](#). The results of these analyses are presented in [appendix 8.F](#).

### 9.3.3. Other Validity Evidence

Convergent and discriminant validity evidence can also be established through a pattern of high correlations among scales that purport to measure domains that are known to be closely related and lower correlations among scales that are intended to measure dissimilar

domains. The pattern of correlations within the Summative ELPAC provides preliminary evidence of validity by showing that the correlations among oral and written language skills are positive and reasonably high, ranging from 0.636 for kindergarten to 0.809 for grade span eleven and twelve. Correlations in this range suggest that oral and written scores measure common and unique aspects of English language proficiency.

Correlations between each composite and the overall scores were larger, ranging from 0.842 to 0.958 for kindergarten and grade span eleven and twelve, respectively. These strong relationships are expected, given that each composite is a major contributor to the overall score. These larger correlations suggest that the overall score captures a substantial portion of the information found in the composite scores.

[Table 9.1](#) Correlation Among Two Composites and the Overall Score provides the correlations between composite scores and overall scores.

**Table 9.1 Correlation Among Two Composites and the Overall Score for the Summative ELPAC**

<b>Grade Level or Grade Span</b>	<b>Composite</b>	<b>Written</b>	<b>Overall</b>
Kindergarten	Oral	0.636	0.952
Kindergarten	Written	1.000	0.842
Grade 1	Oral	0.622	0.879
Grade 1	Written	1.000	0.921
Grade 2	Oral	0.638	0.899
Grade 2	Written	1.000	0.910
Grade span 3–5	Oral	0.744	0.943
Grade span 3–5	Written	1.000	0.924
Grade span 6–8	Oral	0.729	0.948
Grade span 6–8	Written	1.000	0.909
Grade span 9–10	Oral	0.759	0.958
Grade span 9–10	Written	1.000	0.914
Grade span 11–12	Oral	0.809	0.958
Grade span 11–12	Written	1.000	0.944

Another source of validity evidence can be found by comparing the pattern of correlations for the composite and overall scores from 2018–2019 Initial ELPAC test takers with those obtained from the Summative ELPAC test takers. A correlation table for the Initial ELPAC is presented in [table 9.2](#). A similar pattern of correlations for initial and summative scores shows that the relationship between each composite and the overall score is similar for both assessments.

**Table 9.2 Correlation Among Composites and the Overall Score for the Initial ELPAC**

<b>Grade Level or Grade Span</b>	<b>Composite</b>	<b>Written</b>	<b>Overall</b>
Kindergarten	Oral	0.645	0.998
Kindergarten	Written	1.000	0.687
Grade 1	Oral	0.856	0.990

<b>Grade Level or Grade Span</b>	<b>Composite</b>	<b>Written</b>	<b>Overall</b>
Grade 1	Written	1.000	0.921

Table 9.2 (continuation)

<b>Grade Level or Grade Span</b>	<b>Composite</b>	<b>Written</b>	<b>Overall</b>
Grade 2	Oral	0.847	0.966
Grade 2	Written	1.000	0.956
Grade span 3–5	Oral	0.900	0.980
Grade span 3–5	Written	1.000	0.968
Grade span 6–8	Oral	0.881	0.981
Grade span 6–8	Written	1.000	0.956
Grade span 9–12	Oral	0.887	0.980
Grade span 9–12	Written	1.000	0.960

A correlation table between the overall Summative ELPAC scores and Smarter Balanced English language arts/literacy (ELA) scores for students who completed both assessments during the 2018–2019 administration can be found in [table 9.3](#). The correlations are moderately correlated, which indicates that these assessments measure unique aspects of the English language.

**Table 9.3 Correlation of Summative ELPAC Overall and Smarter Balanced for ELA Scores**

<b>Grade Level</b>	<b>N</b>	<b>Correlation</b>
3	94,632	0.68
4	96,470	0.67
5	85,298	0.64
6	71,771	0.61
7	65,384	0.60
8	54,409	0.59
11	37,127	0.56



## References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- California Department of Education. (2014). *California English language development standards: Kindergarten through grade 12*. Sacramento, CA: California Department of Education. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndpublication14.pdf>
- California Department of Education. (2016). *Item writing guidelines for the English Language Proficiency Assessments for California*. [Unpublished report]. Sacramento, CA: California Department of Education.
- California Department of Education. (2017). Summative assessment test blueprints for the English Language Proficiency Assessments for California. In *California State Board of Education September 2017 agenda; Subject English Language Proficiency Assessments for California: Approve the Revised Test Blueprints, the Revised General Performance Level Descriptors, and the Reporting Hierarchy* (pp. 9–21). Sacramento, CA: California Department of Education. Retrieved from <https://www.cde.ca.gov/be/aq/aq/yr17/documents/sep17item18.doc>

## Chapter 10: Quality Control Procedures

---

This chapter highlights the quality control processes used at various stages of administration of the Summative English Language Proficiency Assessments for California (ELPAC).

### 10.1. Quality Control of Item and Test Development

The California Department of Education (CDE) and Educational Testing Service (ETS) implemented rigorous quality control procedures throughout the test development, administration, scoring, analyses, and reporting processes for the Summative ELPAC. As part of this effort, ETS staff worked with the ETS Office of Professional Standards Compliance, which publishes and maintains the *ETS Standards for Quality and Fairness* (ETS, 2014). These *Standards* support the goals of delivering technically sound, fair, and useful products and services; and assisting the public and auditors in evaluating those products and services. Quality control procedures are outlined in this chapter.

#### 10.1.1. Quality Control of Item Writing

After the CDE approved newly developed items for field testing, ETS performed a final review of the items in the ETS Item Banking System called final content review. During this review, an assessment specialist who was familiar with the Summative ELPAC task types performed an independent review of each item to ensure that the item content, metadata, graphics, and audio files were all accurate. The assessment specialist also reviewed comments that were made during previous reviews to ensure that they were implemented. All items were reviewed and approved at final content review before field testing.

#### 10.1.2. Quality Control of Item Selection

Both ETS Assessment & Learning Technology Development (ALTD) staff and Psychometric Analysis & Research staff checked the prior use of items to ensure that items of the appropriate status were used as equating items, operational items, and field test items. Cross checks were also performed to ensure that none of the items placed on an operational form appeared in a public-facing document, such as a practice test.

### 10.2. Quality Control of Test Materials

After the CDE approved all printed test materials (i.e., *Examiner's Manuals*, Test Books, and Answer Books), ETS performed a final certification check of the test materials. For each test form, ETS staff ensured that the various test materials worked together and that all cross references regarding page numbers and question numbers were accurate. The test-length audio files were also checked as part of this process. In each case, the final certification check was completed and any needed revisions were applied before the *Examiner's Manuals*, Test Books, and Answer Books were delivered to the printers for reproduction, and before the test-length audio files were uploaded into the ELPAC Test Operations Management System (TOMS).

#### 10.2.1. Test Administration Manuals

ETS staff consulted with internal subject-matter experts and conducted validation checks to verify that test instruction manuals accurately matched the test booklets and testing processes. Copy editors and content editors reviewed each document for spelling, grammar, accuracy, and adherence to CDE style. Manuals such as the *2018–19 Summative ELPAC Test Administration Manual* were approved by the CDE before they were published

to the ELPAC website at <https://www.elpac.org/>. Only nonsecure documents were posted to this website.

### **10.2.2. Processing Test Materials**

Upon receipt of the test materials, ETS personnel examined each shipment for a number of conditions, including physical damage, shipping errors, and omission of materials. The number of students recorded on the Group Information Sheet (GIS)—the precoded identification sheet that accompanied the grade-level test materials for a school—was compared to the number of Answer Books returned to ETS.

ETS' image-scanning process, which captured security information electronically and compared scorable material quantities reported on the GIS to actual documents scanned, was used when processing returned Answer Books. Local educational agencies (LEAs) were contacted by phone if there were any missing shipments or the quantity of materials returned was less than expected.

### **10.3. Quality Control of Test Delivery**

ETS used several methods to manage and monitor the security of the ELPAC paper-based test materials. All secure test materials were coded with an individual label that identified the item and the number of materials being packed in a shipment, thus allowing ETS to track materials from the time they left the warehouse until they were returned for scoring.

Materials were shipped using United Parcel Service or, for larger orders, via freight. In either case, tracking numbers were used to track these shipments until they were securely delivered at the LEA's warehouse.

### **10.4. Quality Control of Test Assignment**

State and federal law (California *Education Code* sections 313 and 60810 and federal law Titles I and III of the Every Student Succeeds Act) require that all students whose primary language is other than English or Sign Language be assessed for English language proficiency.

LEAs have a role in ensuring students identified as English learners (ELs) are administered the Summative ELPAC annually. Students who were ELs in spring 2019 as a result of taking the Initial ELPAC were required to take the Summative ELPAC. Those ELs must be administered the Summative ELPAC annually until they are reclassified as fluent English proficient.

Proficiency classifications (e.g., initial fluent English proficient, EL) are found in the California Longitudinal Pupil Achievement Data System (CALPADS)—the data system used to maintain student data—in the English Language Acquisition Status (ELAS) field. Students who take the Summative ELPAC have an ELAS of EL.

### **10.5. Quality Control of Test Administration**

During the Summative ELPAC administration, every person who either worked with the assessments, communicated test results, or received testing information was responsible for maintaining the security and confidentiality of the tests, including CDE staff, ETS staff, ETS subcontractors, LEA ELPAC coordinators, site ELPAC coordinators, and teachers.

ETS' Code of Ethics requires that all test information, including tangible materials (e.g., test items and test books), confidential files (e.g., those containing personally identifiable student

information), and processes related to test administration (e.g., the packing and delivery of test materials) are kept secure. For the 2018–2019 Summative ELPAC, ETS had systems in place that maintained tight security for test items, test booklets, and test results, as well as for student data.

To ensure security for all the tests that ETS develops or handles, ETS maintains an Office of Testing Integrity.

## 10.6. Quality Control of Machine-Scoring Procedures

The quality control of paper-pencil tests is ensured by an independent group at ETS that signs in to the ETS Enterprise Score Key Management (eSKM) system and checks scoring keys. This group must sign off and approve the keys before scoring for the administration can begin. This team also creates scoring stencils to be used during the administration to overlay on top of a student's Answer Book to verify the score computed by eSKM is accurate. These quality control procedures were followed during the 2018–2019 Summative ELPAC administration.

## 10.7. Quality Control of Hand Scoring Procedures for Writing

Rater qualifications, rater certifications, and daily rater calibrations are all processes used to control the reliability of constructed-response scoring. For the Summative ELPAC, raters were led through a training period by trained ALTD staff, content scoring leaders, group scoring leaders, and scoring leaders for an assigned grade level and specific prompt types prior to the scoring period. In the training period, raters were trained to appropriately apply the rubrics by using the ELPAC benchmark sample papers.

Trained raters were scheduled to score in four- or eight-hour shifts. Scoring leaders were qualified raters who have the responsibility of providing feedback to raters to provide additional content support and offer corrective mentoring for struggling raters.

Each rater was assigned a secure user ID and password to log on to the scoring system and was required to sign a confidentiality agreement. System access for the rater was restricted to the scheduled work hours. Prior to scoring, raters passed calibration tests that demonstrated sufficient training in ELPAC scoring criteria and an ability to score accurately.

Refer to subsection [7.3 Constructed Response Scoring for Writing](#) for details about constructed response scoring processes.

### 10.7.1. Second Scores

Ten percent of responses were scored twice to check agreement among the raters. The second scores were used for statistical analysis to obtain interrater reliability. A rater did not know whether the response being scored was for a first or a second score. Only the first rating counted towards a student's final score, except in the case of a discrepancy. In the event of a discrepancy of more than one score point between raters, a scoring leader provided a third score, which counted towards the student's final score in place of the first rating.

The second reading sample was not a stratified random sample. The selection of a second reading response also was not based on the first reading score or any demographic information associated with the response. Instead, responses flagged for second reading were flagged at random by the scoring system for each item identification number.

### 10.7.2. Read-Behinds

Scoring leaders performed read-behinds by reviewing selected responses after raters submitted scores. Leaders reviewed rater scoring statistics (i.e., interrater reliability, score point distributions, and validity performance) to determine priorities for monitoring via read-behinds and assign additional training when needed. Responses determined to be scored incorrectly during read-behind review could be rescored by leadership and used to inform and instruct raters as a performance-improvement strategy.

### 10.7.3. Validity Responses

Validity responses were provided randomly as part of the set of “live” responses being scored, so a rater did not know that the response being scored was for validity. These responses were selected from “live” responses by scoring leaders prior to the scoring of the item. Leadership staff identified the response to be used for validity and the system added the response to the validity pool for use during scoring.

## 10.8. Quality Control of Psychometric Processes

### 10.8.1. Development of Scoring Specifications

A number of measures were taken to ascertain that the scoring keys were applied to the student responses as intended and the student scores were computed accurately. ETS built and reviewed the scoring system models based on scoring specifications developed by ETS and approved by the CDE. Machine-scored item responses and demographic information were collected from the Answer Books by ETS. Human-scored item responses were sent electronically to the ETS Online Network for Evaluation for scoring by trained, qualified raters. Record counts were verified against the counts obtained during security check-in from the document processing staff to ensure all students were accounted for in the file.

Once the record counts were reviewed, the machine-scored item responses were scored against the appropriate answer key. In addition, the student’s original response string was stored for data verification and auditing purposes.

The scoring specifications contained detailed scoring procedures, along with the procedures for determining whether a student attempted a test and whether that student response data should be included in the statistical analyses and calculations for computing summary data. Standard quality inspections were performed on all data files, including the evaluation of each student data record for correctness and completeness. Student results were kept confidential and secure at all times.

### 10.8.2. Development of Scoring Procedures

The ETS eSKM scoring system utilizes scoring procedures specified by psychometricians and provides scoring services. Following scoring, a series of quality control checks were carried out by ETS psychometricians to ensure the accuracy of each score.

#### 10.8.2.1. Enterprise Score Key Management System Processing

ETS developed two independent and parallel scoring structures to produce students’ scores: the eSKM<sup>5</sup> scoring system, which collected, scored, and delivered individual students’ scores to the ETS reporting system; and the parallel scoring system developed by ETS Technology and Information Processing Services (TIPS), which scored individual

---

<sup>5</sup> The eSKM system produced the ETS scores of record.

students' responses. The two scoring systems independently applied the same scoring algorithms and specifications.

ETS psychometricians verified the eSKM scoring by comparing all individual student scores from TIPS and resolving any discrepancies. This parallel processing redundancy is an internal quality control step and is in place to verify the accuracy of scoring. Students' scores were reported only when the two parallel systems produce identical results.

If scores did not match, the mismatch was investigated by ETS' Psychometrics, Statistics, and Data Science and eSKM teams and resolved. The mismatch could be a result of a CDE decision not to score an item because a problem was identified in a particular item or rubric. In cases of a mismatch, ETS applied the problem item notification (PIN) not to score the item through the systematic process in eSKM; the mismatch would be possible if TIPS was still in the process of applying the PIN in the parallel system when the student score was being compared. This real-time scoring check was designed to continually detect mismatches and track remediation.

Finally, data extracts were sent to ETS' Data Quality Services for data validation. Following validation, the student response statistical extracts were made available to the psychometricians.

#### **10.8.2.2. Psychometric Processing**

Psychometricians verified the eSKM scoring by comparing the parallel scoring programs, conducting extensive analyses to resolve any discrepancies, and verifying the accuracy of all student scores and reported results. In particular, psychometricians checked variables such as overall scale scores and performance levels, composite scale scores and performance levels, domain performance levels, and number of scored items.

All scores complied with the ETS scoring specifications and the parallel scoring process to ensure the quality and accuracy of scoring and to support the transfer of scores into the database of the student records scoring system before student reports were generated.

### **10.9. Quality Control of Reporting**

To ensure the quality of Summative ELPAC test results, for both individual student and summary reports, three general areas were evaluated:

1. Report formats were compared with input sources from the CDE-approved samples.
2. Report data was validated through quality-control checks performed by ETS' Data Quality Services and Resolutions teams. Additionally, all Student Score Reports (SSRs) were run through ETS' patented quality control (QC) Integrator software.
3. Quality check and production reports were proofread by the CDE and ETS prior to making the score reports available to the LEA for download in TOMS.

All reports were required to include a single, accurate LEA code, a charter school number (if applicable), an LEA name, and a school name. All elements conformed to the CDE's official county/district/school (CDS) code and naming records. From the start of processing through scoring and reporting, the CDS Master File was used to verify and confirm accurate codes and names. CALPADS provided a revised LEA Master File to ETS throughout the year as updates became available.

After the reports were validated against the CDE's requirements, a set of reports for QC LEAs were provided to the CDE and ETS for review and approval. Electronic reports were

sent to the CDE and organized as they were expected to look in production. The CDE and ETS reviewed and approved the report package after a thorough examination.

Upon the CDE's approval of the reports generated for the QC LEAs, ETS proceeded with the report production. The QC districts incorporated CDE-selected LEAs to validate a subset of LEAs that contained key reporting characteristics and demographics representative of the state and provided the final check prior to generating the reports and making them available to the LEAs for download from TOMS.

### **10.9.1. Student Scores Excluded from Summary Reports**

Students who were marked as exempt on the student's Answer Book for both the Listening and Speaking domains, both the Reading and Writing domains, or for all four domains are excluded from summary reports.

### **10.9.2. End-to-End Testing for Operational Administration**

ETS conducted end-to-end testing prior to the start of the test administration. The purpose of this testing was to verify that all systems, processes, and resources were ready for the operational administration.

To begin the quality control process for paper-pencil test administration, the ETS resolutions team completed response documents by marking responses on Answer Books for fictitious students in selected schools and across several LEAs. They marked Answer Books with answers that were all correct, all incorrect, and other test response combinations. These response combinations were the expected results across levels and score ranges. The response booklets were sent for processing, batching, and scanning. Once released from scanning, the test results were sent through the system for scoring and reporting. SSRs were created, along with data files for subject-matter experts in the teams to review and verify.

Individual SSRs were generated based on the fictitious students when 100 percent quality control was demonstrated by ETS' Resolution staff.

## Reference

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>



## Chapter 11: Historical Results

In November 2018, the State Board of Education adopted revised threshold scores for the Summative English Language Proficiency Assessments for California (ELPAC). These revisions were applied across the seven grades and grade spans for the overall scale score, oral language scale score, and written language scale score for the 2018–2019 administration, as shown in [table 7.6](#). Hence the percentage of students at proficiency levels from the 2018–2019 Summative ELPAC are not comparable with the results of the 2017–2018 Summative ELPAC proficiency levels. Consequently, there will be no historical comparison report of the proficiency levels between 2017–2018 and 2018–2019. However, the reporting scale score *is* comparable between the 2017–2018 and 2018–2019 administrations because the reporting scale has not been changed.

Historical comparisons of the Summative ELPAC scale scores are performed to identify trends in student performance and test characteristics over time. Such comparisons were performed for all grades and grade spans for two operational administrations: 2017–2018 and 2018–2019. The cross-sectional comparisons for the same grades in different years (different students) are presented in this chapter.

The indicators of student performance include the mean and standard deviation of overall scale scores, oral language scale scores, and written language scale scores. Test characteristics are compared by examining the reliability and standard error of measurement (SEM) for each test.

### 11.1. Base-Year Comparison

In cross-sectional comparisons, the results from cohorts of students from the 2017–2018 Summative ELPAC administration are compared to student results in the same grades from the 2018–2019 administration. For example, students enrolled in grade three for the 2017–2018 administration are compared with students enrolled in grade three for the 2018–2019 administration.

#### 11.1.1. Examinee Characteristics

[Table 11.1](#) shows the number of students with valid scores for 2017–2018 and 2018–2019.

**Table 11.1 Number of Students, by Grade, Tested Across Years**

Grade	Number of Students with Valid Scores in 2017–2018	Number of Students with Valid Scores in 2018–2019
Kindergarten	175,789	154,118
Grade 1	145,762	127,703
Grade 2	127,304	115,895
Grade 3	113,495	100,010
Grade 4	103,695	101,702
Grade 5	90,517	90,236
Grade 6	78,158	76,325
Grade 7	66,055	70,245
Grade 8	54,843	58,729

Table 11.1 (continuation)

Grade	Number of Students with Valid Scores in 2017–2018	Number of Students with Valid Scores in 2018–2019
Grade 9	51,515	54,965
Grade 10	48,327	47,296
Grade 11	41,769	41,958
Grade 12	33,239	36,600

The number of test takers shows a decrease from the 2017–2018 administration for kindergarten through grade six while the number of test takers for higher grades (with the exception of grade ten) show a slight increase. [Appendix 11.A](#) provides a summary of test takers for each student group.

**11.1.1.1. Summary Statistics of Overall Students**

[Table 11.2](#) shows the mean and standard deviation for overall scale scores for the number of students with valid scores in 2017–2018 and 2018–2019.

**Table 11.2 Overall Scale Score Mean and Standard Deviation Across Administrations**

Grade	2017–2018		2018–2019	
	Mean	Standard Deviation	Mean	Standard Deviation
Kindergarten	1431	62	1426	67
Grade 1	1467	62	1453	64
Grade 2	1487	55	1483	64
Grade 3	1489	49	1485	61
Grade 4	1504	51	1509	65
Grade 5	1517	55	1521	69
Grade 6	1520	58	1517	73
Grade 7	1526	62	1526	80
Grade 8	1532	68	1533	85
Grade 9	1529	85	1524	99
Grade 10	1538	87	1537	102
Grade 11	1539	85	1531	98
Grade 12	1535	95	1516	128

[Table 11.3](#) shows the mean and standard deviation for oral scale scores, for the number of students with valid scores in 2017–2018 and 2018–2019.

**Table 11.3 Oral Scale Score Mean and Standard Deviation Across Administrations**

Grade	2017–2018		2018–2019	
	Mean	Standard Deviation	Mean	Standard Deviation
Kindergarten	1439	62	1434	67
Grade 1	1468	61	1460	64
Grade 2	1488	58	1486	69
Grade 3	1486	59	1482	69
Grade 4	1500	62	1506	75
Grade 5	1512	67	1517	80

Table 11.3 (continuation)

Grade	2017–2018	2017–2018	2018–2019	2018–2019
	Mean	Standard Deviation	Mean	Standard Deviation
Grade 6	1514	73	1514	89
Grade 7	1519	79	1523	98
Grade 8	1523	85	1529	104
Grade 9	1523	111	1519	123
Grade 10	1532	114	1534	127
Grade 11	1530	106	1519	114
Grade 12	1527	112	1507	140

[Table 11.4](#) shows the mean and standard deviation for written scale scores, for the number of students with valid scores in 2017–2018 and 2018–2019.

**Table 11.4 Written Scale Score Mean and Standard Deviation Across Administrations**

Grade	2017–2018	2017–2018	2018–2019	2018–2019
	Mean	Standard Deviation	Mean	Standard Deviation
Kindergarten	1412	84	1405	89
Grade 1	1465	81	1446	78
Grade 2	1486	64	1479	72
Grade 3	1493	50	1488	62
Grade 4	1509	51	1512	64
Grade 5	1523	55	1524	68
Grade 6	1525	54	1519	68
Grade 7	1532	58	1529	73
Grade 8	1541	64	1536	78
Grade 9	1535	73	1529	87
Grade 10	1544	75	1539	90
Grade 11	1548	79	1542	94
Grade 12	1543	91	1524	127

Beginning in 2019–2020, proficiency level comparisons will be provided based on the revised threshold scores introduced in 2018–2019.

### 11.1.2. Scale Score Distributions

Table 11.A.1 through table 11.A.5 in [appendix 11.A](#) present N counts overall and for student groups of interest. Table 11.B.1 through table 11.B.12 provide scale score distributions observed for the 2017–2018 and 2018–2019 administrations for each grade level. All scale score distributions and corresponding frequency counts are presented in intervals of 25 scale score points.

The scale score ranges for each grade are defined in [table 7.6](#).

## 11.2. Test-Form Characteristics

The item and test analysis results of the Summative ELPAC for the comparison years are described in this section. Tables and figures are found in [appendix 11.C](#).

Table 11.C.1 and table 11.C.2 present the average proportion of correct values for the operational items in each domain of the Summative ELPAC. The mean proportion correct is affected by both the difficulty of the items and the ability of the students administered the items.

Table 11.C.3 through table 11.C.6 present the mean item response theory *a*- and *b*-parameter estimates for the two composite language skills of the Summative ELPAC based on the Summative ELPAC base scale that was created following the field test administration.

The average point-biserial correlations for the Summative ELPAC operational items are presented in table 11.C.7 and table 11.C.8. The reliability estimates and SEMs, expressed in raw score units, appear in table 11.C.9 through table 11.C.12. Like the average proportion correct, point-biserial correlations and reliability estimates for the operational items are affected by both item characteristics and student characteristics.

The test characteristic curves (TCCs) of the two composite language skills are presented in figure 11.C.1 through figure 11.C.14; the data on which these TCCs are based is found in table 11.C.13 through table 11.C.26.

The 2018–2019 kindergarten oral language test is harder than the 2017–2018 test form, while the 2018–2019 grade span six through eight oral language test is easier than the 2017–2018 test form. The oral language test forms for other grades are comparable. Most 2018–2019 written test forms are comparable with the 2017–2018 written test form, except grade span six through eight, which was slightly harder than the 2017–2018 test form.

# Chapter 12: Continuous Program Improvement

---

This chapter summarizes the completed and ongoing improvements for the Summative ELPAC in the areas of threshold validation, test development, and test delivery.

## 12.1. Test Development

During the administration of the 2017–2018 Summative ELPAC, local educational agencies (LEAs) informed the California Department of Education (CDE) that they wanted to conduct local scoring of the 2018–2019 Summative English Language Proficiency Assessments for California (ELPAC) to make preliminary plans for the following school year. As a result, the CDE requested that Educational Testing Service (ETS) develop materials that would allow LEAs to optionally perform local scoring to generate unofficial, preliminary results. ETS continued to produce official scores.

The goal of the secure *2018–2019 Summative ELPAC Scoring Guide* was to provide a resource that could be used to perform local scoring after the assessment has been administered. The *Scoring Guide* included all of the following materials:

- Writing rubrics and anchor samples
- List of items and scoring keys
- Guidance for local scoring: calculation of raw scores and scaled scores
- Raw score conversion tables
- Performance level descriptors
- Student score sheet

The *2018–2019 Summative ELPAC Scoring Guide* was produced and posted in the Test Operations Management System, so that any LEA that wished to perform local scoring had access to this document.

## 12.2. Test Delivery

### 12.2.1. Move to Online Testing

The 2019–2020 Summative ELPAC will be an online assessment.

### 12.2.2. Postadministration Survey

During the 2018–2019 Summative ELPAC administration, ETS administered a post-test survey to LEAs. The survey focused on gathering information on the test materials delivery; clarity of the test administration, *Examiner's Manuals*, and return instructions; and overall administration experience.

In response to the LEA feedback, ETS is implementing the following improvements for the 2019–2020 administration:

- Clarifying the use of universal tools, designated supports, and accommodations by promoting the existing short California Assessment of Student Performance and Progress (CAASPP) demonstration videos to ELPAC stakeholders new to online testing
- Creating and sending pre-identification labels for the K–2 Writing materials without requiring LEAs to order them

- Creating manuals and user guides that have a similar look and feel to the CAASPP online testing resources for consistency and familiarity

### **12.2.3. Training and Communication**

As ETS continues work on the Summative ELPAC, recruitment, training, and communication will be a focal point moving forward. ETS will continue to provide timely communications for each critical component of the ELPAC administration, including material order dates and deadlines and training schedules. ETS will continue to work with the Sacramento County Office of Education to emphasize the importance and necessity of training, along with providing statewide training to LEA staff so they are prepared to administer the test. Training will continue to focus on local scoring of the Speaking domain.

ETS will continue to support familiarizing students with the ELPAC items using practice and training tests and informational videos. Parent engagement continues to be an important factor for student engagement and familiarization. To that end, ETS will work with the CDE to increase communication and information targeted at parents. Communications will also encourage LEAs to use the practice and training tests to prepare students to become more familiar with the ELPAC.

## **12.3. Constructed-Response Scoring**

During scoring of Writing responses for the 2018–2019 Summative ELPAC, ETS implemented an automated system in the Online Network for Evaluation (ONE) to monitor rater accuracy and provide remediation. ONE monitored rater performance over time. If a rater was not performing at the expected level for a certain prompt, the rater was redirected to complete a mandatory training (i.e., remediation) set before resuming operational scoring. Performance was measured by monitoring raters' scoring accuracy against prescored validity papers that were inserted into each rater's scoring queue. The scoring agreement rates required for each score scale were determined by analysis of past data. Remediation sets were adjusted to target common trouble areas for raters. If rater performance failed to improve after remediation, the rater was restricted from scoring that prompt.

Plans have been put in place to “frontload” double-reads of Writing responses during future Summative ELPAC scoring windows. The current process is to set the double-read percentage at a fixed rate (10 percent) for the entire scoring period. This process will be modified to set the percentage at a higher rate in the beginning of the scoring window, to receive early feedback about potential scoring issues. The percentage double-read will be set to produce approximately 500 double-reads, to allow early analysis of the data. The percentage will then be decreased to a fixed level for the rest of the Summative ELPAC scoring window.