

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



# Initial Alternate English Language Proficiency Assessments for California 2022–23 Technical Report

**Final Submitted May 8, 2024**

**By ETS**



Contract #CN220002

Table of Contents

[Chapter 1: Introduction 1](#_Toc165450872)

[1.1. Alternate ELPAC Overview 1](#_Toc165450873)

[1.2. Purposes of the Assessments 1](#_Toc165450874)

[1.3. Test Content and Design 2](#_Toc165450875)

[1.4. Intended Population 2](#_Toc165450876)

[1.5. Testing Windows and Estimated Testing Times 3](#_Toc165450877)

[1.6. Test Scores 3](#_Toc165450878)

[1.7. Groups and Organizations Involved with the Alternate ELPAC 4](#_Toc165450879)

[1.7.1. California State Board of Education 4](#_Toc165450880)

[1.7.2. California Department of Education 4](#_Toc165450881)

[1.7.3. California Educators 4](#_Toc165450882)

[1.7.4. Contractors 4](#_Toc165450883)

[1.8. Systems Overview and Functionality 6](#_Toc165450884)

[1.8.1. Test Operations Management System 6](#_Toc165450885)

[1.8.2. Test Delivery System 6](#_Toc165450886)

[1.8.3. Test Results for California’s Assessments Website 7](#_Toc165450887)

[1.9. Overview of the Technical Report 7](#_Toc165450888)

[References 9](#_Toc165450889)

[Chapter 2: Overview of Initial Alternate ELPAC Processes 10](#_Toc165450890)

[2.1. Item Development 10](#_Toc165450891)

[2.1.1. Selection of ELD Connectors for Development 10](#_Toc165450892)

[2.1.2. Item Format 11](#_Toc165450893)

[2.1.3. Task Type Specifications 11](#_Toc165450894)

[2.2. Test Assembly 11](#_Toc165450895)

[2.2.1. Test Blueprint 12](#_Toc165450896)

[2.2.2. Test Length 12](#_Toc165450897)

[2.3. Test Administration 12](#_Toc165450898)

[2.3.1. Test Security and Confidentiality 12](#_Toc165450899)

[2.3.2. Procedures to Maintain Standardization 13](#_Toc165450900)

[2.4. Fairness and Accessibility 13](#_Toc165450901)

[2.4.1. Overview 13](#_Toc165450902)

[2.4.2. Student Accessibility Resources 14](#_Toc165450903)

[2.5. Scores 14](#_Toc165450904)

[2.5.1. Score Reporting 14](#_Toc165450905)

[2.5.2. Aggregation Procedures 15](#_Toc165450906)

[2.6. Psychometric Analyses 15](#_Toc165450907)

[2.6.1. Description of the Classical Item Analyses 15](#_Toc165450908)

[2.6.2. Description of Item Response Theory Analyses 15](#_Toc165450909)

[References 16](#_Toc165450910)

[Chapter 3: Item Development and Review 17](#_Toc165450911)

[3.1. Overview 17](#_Toc165450912)

[3.1.1. Preparation 17](#_Toc165450913)

[3.1.2. Task Type Specifications 17](#_Toc165450914)

[3.1.3. Item Writing by Contractors 18](#_Toc165450915)

[3.1.4. Item Writer Training 18](#_Toc165450916)

[3.2. ETS Item Review Process 19](#_Toc165450917)

[3.2.1. Overview 19](#_Toc165450918)

[3.2.2. ETS Content Review 20](#_Toc165450919)

[3.2.3. ETS Accessibility Review 20](#_Toc165450920)

[3.2.4. ETS Editorial Review 20](#_Toc165450921)

[3.2.5. ETS Sensitivity and Fairness Review 20](#_Toc165450922)

[3.3. California Department of Education Review 21](#_Toc165450923)

[3.4. California Educator Review 21](#_Toc165450924)

[3.4.1. California Educators as Content Experts 22](#_Toc165450925)

[3.5. Data Review Meeting 22](#_Toc165450926)

[References 23](#_Toc165450927)

[Chapter 4: Test Assembly 24](#_Toc165450928)

[4.1. Overview 24](#_Toc165450929)

[4.2. Assessment Development 24](#_Toc165450930)

[4.2.1. High-Level Test Design 25](#_Toc165450931)

[4.2.2. Test Blueprint 26](#_Toc165450932)

[4.2.3. Form Assembly Specifications 27](#_Toc165450933)

[4.3. Test Production Process 27](#_Toc165450934)

[4.3.1. Selection of Task Types and Items 28](#_Toc165450935)

[4.3.2. Test Forms 28](#_Toc165450936)

[4.3.3. Psychometric Criteria and Review 28](#_Toc165450937)

[4.3.4. Content Review of Forms 29](#_Toc165450938)

[4.3.5. California Department of Education Forms Review 29](#_Toc165450939)

[4.3.6. Configuration of the Test Delivery System 29](#_Toc165450940)

[4.3.7. Test Form Delivery 30](#_Toc165450941)

[References 31](#_Toc165450942)

[Chapter 5: Test Administration 32](#_Toc165450943)

[5.1. Overview 32](#_Toc165450944)

[5.2. Administration and Scoring Training 32](#_Toc165450945)

[5.2.1. Overview 32](#_Toc165450946)

[5.2.2. Goals 33](#_Toc165450947)

[5.2.3. Local Educational Agency Training Requirement 34](#_Toc165450948)

[5.2.4. Materials on the Moodle Training Site 34](#_Toc165450949)

[5.3. User Roles and Standardization 36](#_Toc165450950)

[5.3.1. Local Educational Agency ELPAC Coordinator 36](#_Toc165450951)

[5.3.2. Site ELPAC Coordinator 37](#_Toc165450952)

[5.3.3. Test Examiner 37](#_Toc165450953)

[5.3.4. Instructions for Test Administration 38](#_Toc165450954)

[5.4. Local Educational Agency Training 39](#_Toc165450955)

[5.4.1. Synchronous and Asynchronous Training 40](#_Toc165450956)

[5.4.2. Videos and Guides 40](#_Toc165450957)

[5.4.3. Training for Proper Identification and Assignment of Designated Supports and Accommodations 41](#_Toc165450958)

[5.4.4. Feedback for Continuous Improvement Survey 42](#_Toc165450959)

[5.5. Accessibility Resources 42](#_Toc165450960)

[5.5.1. Accessibility Resource Categories 43](#_Toc165450961)

[5.5.2. Identification and Selection 46](#_Toc165450962)

[5.5.3. Assignment 46](#_Toc165450963)

[5.5.4. Delivery of Embedded and Non-Embedded Resources to Students 51](#_Toc165450964)

[5.5.5. Usage of Designated Supports and Accommodations 51](#_Toc165450965)

[5.6. Practice and Training Tests 51](#_Toc165450966)

[5.7. Test Security and Confidentiality 52](#_Toc165450967)

[5.7.1. ETS’ Office of Testing Integrity 52](#_Toc165450968)

[5.7.2. Procedures to Maintain Standardization of Test Security 53](#_Toc165450969)

[5.7.3. Test Security Monitoring 54](#_Toc165450970)

[5.7.4. Security of Electronic Files Using a Firewall 54](#_Toc165450971)

[5.7.5. Transfer of Scores via Secure Data Exchange 55](#_Toc165450972)

[5.7.6. Data Management in the Secure Database 55](#_Toc165450973)

[5.7.7. Statistical Analysis on Secure Servers 55](#_Toc165450974)

[5.7.8. Student Confidentiality 56](#_Toc165450975)

[5.7.9. Student Test Results 56](#_Toc165450976)

[5.7.10. Security and Test Administration Incident Reporting System Process 57](#_Toc165450977)

[5.7.11. Appeals 59](#_Toc165450978)

[References 61](#_Toc165450979)

[Chapter 6: Standard Setting 62](#_Toc165450980)

[6.1. Description 62](#_Toc165450981)

[Reference 63](#_Toc165450982)

[Chapter 7: Scoring and Reporting 64](#_Toc165450983)

[7.1. Student Test Scores 64](#_Toc165450984)

[7.1.1. Raw Scores 64](#_Toc165450985)

[7.1.2. Theta Scores 64](#_Toc165450986)

[7.1.3. Scale Scores for the Total Assessment 65](#_Toc165450987)

[7.1.4. Performance Levels 65](#_Toc165450988)

[7.2. Overview of Score Aggregation Procedures 66](#_Toc165450989)

[7.2.1. Student Score Distributions and Summary Statistics 66](#_Toc165450990)

[7.2.2. Demographic Student Group Summaries 68](#_Toc165450991)

[7.2.3. Student Group Distributions 69](#_Toc165450992)

[7.3. Reports Produced and Scores for Each Report 71](#_Toc165450993)

[7.3.1. Online Reporting 71](#_Toc165450994)

[7.3.2. Special Cases 71](#_Toc165450995)

[7.3.3. Types of Score Reports 71](#_Toc165450996)

[7.3.4. Score Report Applications 73](#_Toc165450997)

[7.3.5. Criteria for Interpreting Test Scores 73](#_Toc165450998)

[7.3.6. Criteria for Interpreting Score Reports 73](#_Toc165450999)

[References 74](#_Toc165451000)

[Appendix 7.A: Raw-Score-to-Scale-Score Distributions 75](#_Toc165451001)

[Appendix 7.B: Demographic Student Group Summaries 81](#_Toc165451002)

[Chapter 8: Psychometric Analyses 93](#_Toc165451003)

[8.1. Overview 93](#_Toc165451004)

[8.1.1. Summary of the Analyses 93](#_Toc165451005)

[8.1.2. Samples Used for the Analyses 93](#_Toc165451006)

[8.1.3. Test-Taking Rates 94](#_Toc165451007)

[8.2. Classical Item Analyses 94](#_Toc165451008)

[8.2.1. Classical Item Difficulty Indices (*p*-value and Average Item Score) 94](#_Toc165451009)

[8.2.2. Item-Total Correlation 95](#_Toc165451010)

[8.2.3. Distribution of Item Scores 96](#_Toc165451011)

[8.2.4. Omit Rates 96](#_Toc165451012)

[8.2.5. Classical Item Analyses Results 97](#_Toc165451013)

[8.3. Item Response Theory Analyses 98](#_Toc165451014)

[8.3.1. Item Response Theory Model 98](#_Toc165451015)

[8.3.2. Equating 99](#_Toc165451016)

[8.3.3. Scaling the Scores 99](#_Toc165451017)

[8.3.4. Parameter Estimates 101](#_Toc165451018)

[8.4. Reliability Analyses 103](#_Toc165451019)

[8.4.1. Reliability Measures 103](#_Toc165451020)

[8.4.2. Standard Error of Measurement 104](#_Toc165451021)

[8.4.3. Student Group Reliabilities 105](#_Toc165451022)

[8.4.4. Conditional Standard Errors of Measurement 107](#_Toc165451023)

[8.4.5. Decision Classification Analyses 109](#_Toc165451024)

[8.5. Validity Evidence 111](#_Toc165451025)

[8.5.1. Design of the Initial Alternate ELPAC 112](#_Toc165451026)

[8.5.2. Content 114](#_Toc165451027)

[8.5.3. Response Processes 114](#_Toc165451028)

[8.5.4. Internal Structure 115](#_Toc165451029)

[References 116](#_Toc165451030)

[Accessibility Information 118](#_Toc165451031)

[Alternative Text for Equation 8.1 118](#_Toc165451032)

[Alternative Text for Equation 8.2 118](#_Toc165451033)

[Alternative Text for Equation 8.3 118](#_Toc165451034)

[Alternative Text for Equation 8.4 118](#_Toc165451035)

[Alternative Text for Equation 8.5 118](#_Toc165451036)

[Alternative Text for Equation 8.6 118](#_Toc165451037)

[Alternative Text for Equation 8.7 118](#_Toc165451038)

[Alternative Text for Equation 8.8 118](#_Toc165451039)

[Alternative Text for Equation 8.9 119](#_Toc165451040)

[Alternative Text for Equation 8.10 119](#_Toc165451041)

[Alternative Text for Equation 8.11 119](#_Toc165451042)

[Alternative Text for Equation 8.12 119](#_Toc165451043)

[Alternative Text for Equation 8.13 119](#_Toc165451044)

[Alternative Text for Equation 8.14 119](#_Toc165451045)

[Alternative Text for Equation 8.15 119](#_Toc165451046)

[Alternative Text for Equation 8.16 119](#_Toc165451047)

[Alternative Text for Equation 8.17 119](#_Toc165451048)

[Appendix 8.A: Classical Item Analyses Results 120](#_Toc165451049)

[Appendix 8.B: Item Response Theory Results 126](#_Toc165451050)

[Chapter 9: Quality Control 132](#_Toc165451051)

[9.1. Quality Control of Item Development 132](#_Toc165451052)

[9.2. Quality Control of Test Assembly and Delivery 132](#_Toc165451053)

[9.2.1. Quality Control of Test Assignment 133](#_Toc165451054)

[9.3. Quality Control of Test Materials 134](#_Toc165451055)

[9.3.1. Test Administration Manuals 134](#_Toc165451056)

[9.3.2. Processing Test Materials 134](#_Toc165451057)

[9.4. Quality Control of Test Administration 134](#_Toc165451058)

[9.5. Quality Control of Scoring 135](#_Toc165451059)

[9.5.1. Machine-Scoring Procedures 135](#_Toc165451060)

[9.5.2. Rubric-Scored Item Scoring 135](#_Toc165451061)

[9.5.3. Development of Scoring Specifications 136](#_Toc165451062)

[9.6. Quality Control of Psychometric Processes 136](#_Toc165451063)

[9.6.1. Scoring Verification 136](#_Toc165451064)

[9.6.2. Psychometric Analyses 136](#_Toc165451065)

[9.7. Quality Control of Reporting 137](#_Toc165451066)

[9.8. Quality Control of End-to-End Testing 138](#_Toc165451067)

[9.8.1. Computer-based Assessments 138](#_Toc165451068)

[References 139](#_Toc165451069)

[Chapter 10: Threshold Score Validation Study 140](#_Toc165451070)

[Reference 141](#_Toc165451071)

[Chapter 11: Continuous and Systematic Improvement 142](#_Toc165451072)

[11.1. 2022–23 Feedback for Continuous Improvement Survey 142](#_Toc165451073)

[11.1.1. Recommendations for Improvement 142](#_Toc165451074)

[11.2. Test Design 142](#_Toc165451075)

[Test Delivery 143](#_Toc165451076)

[11.2.1. Changes to the Test Administrator Interface 143](#_Toc165451077)

[11.2.2. Changes to Ending the Assessment in the Test Delivery System 143](#_Toc165451078)

[11.3. Student Score Reports Redesign 143](#_Toc165451079)

[11.4. Accessibility Resources 143](#_Toc165451080)

[Reference 144](#_Toc165451081)

List of Tables

[Acronyms and Initialisms Used in the *Initial Alternate English Language Proficiency Assessments for California Technical Report* viii](#_Toc165451082)

[Table 1.1 Differences Between the Initial Alternate ELPAC and Summative Alternate   
ELPAC 1](#_Toc165451083)

[Table 4.1 Kindergarten Through Grade Twelve Initial Alternate ELPAC Forms Plan 27](#_Toc165451084)

[Table 5.1 Assignment of Accommodations and Designated Supports—Kindergarten Through Grade Two 48](#_Toc165451085)

[Table 5.2 Assignment of Accommodations and Designated Supports—Grade Three Through Grade Twelve 49](#_Toc165451086)

[Table 5.3 Types of Appeals 57](#_Toc165451087)

[Table 5.4 Number and Types of Incidents Submitted in STAIRS 59](#_Toc165451088)

[Table 5.5 Number of Appeals Approved in STAIRS in the 2022–23 Administration—All Grade Levels and Grade Spans 60](#_Toc165451089)

[Table 7.1 Mean and SD of Scale Scores by Grade Level 66](#_Toc165451090)

[Table 7.2 Percentage of Students in Each Performance Level by Grade Level 66](#_Toc165451091)

[Table 7.3 Demographic Student Groups to Be Reported 68](#_Toc165451092)

[Table 7.4 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Kindergarten 69](#_Toc165451093)

[Table 7.A.1 Raw-Score-to-Scale-Score Distribution for Kindergarten 75](#_Toc165451094)

[Table 7.A.2 Raw-Score-to-Scale-Score Distribution for Grade One 76](#_Toc165451095)

[Table 7.A.3 Raw-Score-to-Scale-Score Distribution for Grade Two 77](#_Toc165451096)

[Table 7.A.4 Raw-Score-to-Scale-Score Distribution for Grade Span Three Through Five 78](#_Toc165451097)

[Table 7.A.5 Raw-Score-to-Scale-Score Distribution for Grade Span Six Through Eight 79](#_Toc165451098)

[Table 7.A.6 Raw-Score-to-Scale-Score Distribution for Grade Span Nine Through   
Twelve 80](#_Toc165451099)

[Table 7.B.1 Demographic Summary for Students: Kindergarten 81](#_Toc165451100)

[Table 7.B.2 Demographic Summary for Students: Grade One 83](#_Toc165451101)

[Table 7.B.3 Demographic Summary for Students: Grade Two 85](#_Toc165451102)

[Table 7.B.4 Demographic Summary for Students: Grade Span Three Through Five 87](#_Toc165451103)

[Table 7.B.5 Demographic Summary for Students: Grade Span Six Through Eight 89](#_Toc165451104)

[Table 7.B.6 Demographic Summary for Students: Grade Span Nine Through Twelve 91](#_Toc165451105)

[Table 8.1 The Test-Taking Rates by Grade Level 94](#_Toc165451106)

[Table 8.2 Summary Statistics for Classical Item Analysis 97](#_Toc165451107)

[Table 8.3 Slopes and Intercepts That Convert Theta Scores to Reporting Scale Scores 101](#_Toc165451108)

[Table 8.4 Distribution of *b*-values 102](#_Toc165451109)

[Table 8.5 Reliability of Reporting Scale Scores 105](#_Toc165451110)

[Table 8.6 Student Group Reliabilities for Kindergarten 105](#_Toc165451111)

[Table 8.7 Scale Score CSEMs at the Score Reporting Range Thresholds 109](#_Toc165451112)

[Table 8.8 Decision Accuracy for Reaching a Performance Level 110](#_Toc165451113)

[Table 8.9 Decision Consistency for Reaching a Performance Level 110](#_Toc165451114)

[Table 8.10 Classification Accuracy and Consistency 111](#_Toc165451115)

[Table 8.A.1 Classical Item Statistics for Kindergarten 120](#_Toc165451116)

[Table 8.A.2 Classical Item Statistics for Grade One 121](#_Toc165451117)

[Table 8.A.3 Classical Item Statistics for Grade Two 122](#_Toc165451118)

[Table 8.A.4 Classical Item Statistics for Grade Span Three Through Five 123](#_Toc165451119)

[Table 8.A.5 Classical Item Statistics for Grade Span Six Through Eight 124](#_Toc165451120)

[Table 8.A.6 Classical Item Statistics for Grade Span Nine Through Twelve 125](#_Toc165451121)

[Table 8.B.1 IRT Item Statistics, Kindergarten 126](#_Toc165451122)

[Table 8.B.2 IRT Item Statistics, Grade One 127](#_Toc165451123)

[Table 8.B.3 IRT Item Statistics, Grade Two 128](#_Toc165451124)

[Table 8.B.4 IRT Item Statistics, Grade Span Three Through Five 129](#_Toc165451125)

[Table 8.B.5 IRT Item Statistics, Grade Span Six Through Eight 130](#_Toc165451126)

[Table 8.B.6 IRT Item Statistics, Grade Span Nine Through Twelve 131](#_Toc165451127)

List of Figures

[Figure 7.1 Percentage of students at each performance level 67](#_Toc157838681)

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

|  |  |
| --- | --- |
| Term | Definition |
| 1PL-IRT | one-parameter logistic item response theory |
| AERA | American Educational Research Association |
| AIS | average item score |
| ALTD | Assessment & Learning Technology Development |
| APA | American Psychological Association |
| AST | Administration and Scoring Training |
| CAASPP | California Assessment of Student Performance and Progress |
| CAI | Cambium Assessment, Inc. |
| CALPADS | California Longitudinal Pupil Achievement Data System |
| CalTAC | California Technical Assistance Center |
| *CCR* | *California Code of Regulations* |
| CCSSO | Council of Chief State School Officers |
| CDE | California Department of Education |
| CDS | county/district/school |
| CI | confidence interval |
| CR | constructed response |
| CSEM | conditional standard error of measurement |
| *DFA* | *Directions for Administration* |
| DIF | differential item functioning |
| DRM | data review meeting |
| *EC* | *Education Code* |
| EL | English learner |
| ELD | English language development |
| ELP | English language proficiency |
| ELPAC | English Language Proficiency Assessments for California |
| eSKM | Enterprise Score Key Management |
| GPCM | generalized partial credit model |
| HLS | home language survey |
| IDEA | Individuals with Disabilities Education Act |
| IEP | individualized education program |
| IFEP | initial fluent English proficient |
| IRM | item review meeting |
| IRT | item response theory |
| ISAAP | Individual Student Assessment Accessibility Profile |
| K | kindergarten |
| LEA | local educational agency |
| LOSS | lowest obtainable scale score |
| MCSS | multiple choice, single select |
| Moodle | Moodle Training Site |
| NCME | National Council on Measurement in Education |
| OTI | Office of Testing Integrity |
| PAR | Psychometric Analysis & Research |
| *PFA* | *Preparing for Administration* |
| PLD | performance level descriptor |
| QA | quality assurance |
| SBE | State Board of Education |
| SCOE | Sacramento County Office of Education |
| SD | standard deviation |
| SEM | standard error of measurement |
| SFTP | secure file transfer protocol |
| SSID | Statewide Student Identifier |
| SSR | Student Score Report |
| STAIRS | Security and Test Administration Incident Reporting System |
| TCC | test characteristic curve |
| TDS | test delivery system |
| TIP | Test Item Preview |
| TOMS | Test Operations Management System |
| UAT | user acceptance testing |
| UDL | Universal Design for Learning |
| *USC* | *United States Code* |

**This page is intentionally left blank.**

## Introduction

This technical report focuses on the development, administration, psychometric analyses, and results of the administration of the Initial Alternate English Language Proficiency Assessments for California (ELPAC) for the 2022–23 test administration. This chapter provides an introduction to the Initial Alternate ELPAC program, including background information, purpose of the assessment, test content, intended population, testing windows, organizations and systems involved, and an overview of the technical report.

### Alternate ELPAC Overview

The Alternate ELPAC “is the required state test for English language proficiency (ELP) given to students whose primary language is a language other than English” (California Department of Education [CDE], 2023a) and who have been found eligible for alternate assessments by their individualized education program (IEP) team. State and federal laws require that local educational agencies (LEAs) administer a state assessment of ELP to eligible students in kindergarten through grade twelve (K–12) (CDE, 2023a). California *Education Code (EC)* Section 313(a) requires that the assessment of ELP be conducted upon initial enrollment and annually thereafter until the LEA reclassifies the student as fluent English proficient.

### Purposes of the Assessments

The Alternate ELPAC consists of two assessments: the Initial Alternate ELPAC and the Summative Alternate ELPAC. Their purposes are as follows:

1. The Initial Alternate ELPAC provides information to determine a student’s initial classification as an English learner (EL), or as initial fluent English proficient (IFEP), for students with the most significant cognitive disabilities.
2. The Summative Alternate ELPAC provides information on annual student progress toward ELP and supports decisions on student reclassification as fluent English proficient for students with the most significant cognitive disabilities.

The contents of table 1.1 describe the differences between the Initial Alternate ELPAC and the Summative Alternate ELPAC.

Table 1.1 Differences Between the Initial Alternate ELPAC and Summative Alternate ELPAC

|  |  |
| --- | --- |
| Initial Alternate ELPAC | Summative Alternate ELPAC |
| This is an assessment used to identify a student as either an EL who needs support to learn English or as IFEP. | This is an assessment used to measure the ELP of EL students. The results will help the school or LEA determine whether the student is ready to be reclassified as proficient in English. |
| This assessment is administered to students with a home language survey (HLS) that lists a language other than English as the primary language within 30 days of when the student enrolls in a California public school for the first time. **Eligible students must have an IEP designating the use of alternate assessments.** | This assessment is administered to eligible students every spring, from February 1 to May 31. **Eligible students must have an IEP designating the use of alternate assessments.** |
| A student takes this assessment one time only. The Initial Alternate ELPAC is taken before the Summative Alternate ELPAC or Summative ELPAC, to determine whether the student is identified as an EL. | A student takes this assessment annually until the student is reclassified. |
| There is one test form. | There are two test forms that are refreshed annually. |
| There are six grade levels and grade spans: kindergarten, 1, 2, 3–5, 6–8, and 9–‍12. | There are seven grade levels and grade spans: kindergarten, 1, 2, 3–5, 6–8, 9–10, and 11–‍12. |

### Test Content and Design

The Initial Alternate ELPAC is designed to align with the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) via the English Language Development Connectors (ELD Connectors) (CDE, 2014), which reduce the depth, breadth, and complexity of the standards, as appropriate for students with the most significant cognitive disabilities. The ELD Connectors represent the highest level of expected performance in ELP for eligible students with the most significant cognitive disabilities at a given grade level or grade span. The ELD Connectors are not intended to represent the full range of performance in ELP that may be measured by a general ELP assessment.

The ELD Connectors were developed through collaboration among California educators, the CDE, and ETS’ research and assessment experts, as well as with guidance from the Alternate ELPAC Test Design Advisory Team of four nationally recognized experts on the assessment of EL students with the most significant cognitive disabilities.

### Intended Population

The Initial Alternate ELPAC is given to newly enrolled students in K–12 (up to age twenty-one) with the most significant cognitive disabilities whose primary language is a language other than English. Students who have been found eligible for alternate assessments by their IEP team and are enrolled in California schools for the first time, and who are potential EL students as indicated by the results of an HLS, must be administered the Initial Alternate ELPAC within 30 days of their enrollment.

### Testing Windows and Estimated Testing Times

The Initial Alternate ELPAC testing window runs from July 1 to June 30 annually. During this time period, any student whose primary or native language is a language other than English (determined by the HLS administered by the LEA), who has not previously been classified as an EL or IFEP student by a California public school, and who has an IEP that designates the use of alternate assessments is required to be administered the Initial Alternate ELPAC. The testing window for the administration of the Initial Alternate ELPAC was from July 5, 2022, through June 30, 2023.

During the Initial Alternate ELPAC testing window, LEAs were required to test and then provide the parent/guardian with the written results of the Initial Alternate ELPAC within 30 calendar days of the student’s initial date of enrollment in a California school. If the Initial Alternate ELPAC was administered prior to the student’s initial date of California enrollment, the written results of the Initial Alternate ELPAC could be provided to the parent/guardian up to 60 calendar days prior to enrollment, but not before July 1, 2022 (*California Code of Regulations,* Title 5,Section 11518.5[e]).

The Initial Alternate ELPAC is a computer-based, nonadaptive, untimed assessment that is delivered one-on-one by a test examiner. Students are allowed as much time as they need to complete their responses to each item. The assessment may be administered over multiple days. The estimated testing times for the Initial Alternate ELPAC task types were posted by task type in the *Initial Alternate ELPAC Test Administration Manual* (CDE, 2023b). Estimated testing times were provided for administration planning only.

### Test Scores

The Initial Alternate ELPAC measures the ELP of a student with the most significant cognitive disabilities and determines what English language support, if any, the student needs to succeed in school while receiving instruction in all school subjects. If a student is identified as IFEP after taking the Initial Alternate ELPAC, that student generally requires minimal support learning English and does not need EL programs and services. A student with overall performance levels of Novice EL or Intermediate EL requires substantial to moderate EL programs and services. Student test scores were used to identify their performance levels through a process called standard setting. Refer to [chapter 6](#_Appendix_4.A:_Demographic) for more information about this process.

The California State Board of Education (SBE) approved the reporting hierarchy of the Initial Alternate ELPAC in May 2019. Individual student scores for the Initial Alternate ELPAC for all grade levels and grade spans (i.e., K–12) included an overall performance level and scale score. The overall score placed the student within one of the three Initial ELPAC overall performance levels:

1. Novice EL, Level 1
2. Intermediate EL, Level 2
3. Initial fluent English proficient, Level 3

### Groups and Organizations Involved with the Alternate ELPAC

#### California State Board of Education

The SBE is the state agency that establishes educational policy for K–12 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 is also the state educational agency responsible for overseeing California’s compliance with programs that meet the requirements of the federal Every Student Succeeds Act as well as the state’s Public School Accountability Act that measures the academic performance and progress of schools on a variety of academic metrics (CDE, 2023d).

#### California Department of Education

The CDE oversees California’s public school system, which is responsible for the education of more than 5,800,000 children and young adults in more than 10,010 schools.[[1]](#footnote-2) 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 with educators, school staff, parents/guardians, and community partners which together, as a team, prepare students to live, work, and thrive in a highly connected world.

Within the CDE, it is the Instruction, Measurement, & Administration Branch that oversees programs promoting improved student achievement. Programs include oversight of statewide assessments and the collection and reporting of educational data (CDE, 2023c).

#### California Educators

A variety of California educators, including school administrators as well as those experienced in teaching EL students, students with the most significant cognitive disabilities, or both—who were selected on the basis of their qualifications, experiences, demographics, and geographic locations—were invited to participate in the Alternate ELPAC development process. In this process, California educators participated in tasks that included work related to defining the purpose and scope of the assessment, assessment design, item development, item reviews, standard setting, and score reporting.

#### Contractors

A number of organizations contribute to the success of the Initial Alternate ELPAC.

##### Primary Testing Contractor—ETS

The CDE and the SBE contract with ETS to develop, administer, and report the Initial Alternate ELPAC. As the primary testing contractor, ETS has overall responsibility for working with the CDE to implement and maintain an effective assessment system and coordinating ETS’ work with its subcontractors.

Activities conducted directly by ETS include, but are not limited to, the following:

* Providing management of the program activities
* Developing high-quality items that are aligned to the 2012 ELD Standards via the ELD Connectors
* Supporting and training county offices of education, LEAs, and direct funded charter schools
* Constructing, producing, and controlling the quality of Initial Alternate ELPAC test forms and related test materials, including grade- and content-specific *Directions for Administration (DFAs)*
* Hosting and maintaining a website with resources for LEA ELPAC coordinators
* Developing, hosting, and providing support for the Test Operations Management System (TOMS)
* Supporting the California Educator Reporting System
* Processing student test assignments
* Processing orders and shipment of test materials
* Producing and distributing score reports electronically
* Developing a summary score reporting website that can be viewed by the public
* Completing all psychometric procedures
* Providing a tiered help desk support system for LEAs

##### Subcontractor—Cambium Assessment, Inc

ETS also monitors and manages the work of Cambium Assessment, Inc. (CAI), subcontractor to ETS for the ELPAC System of computer-based assessments. Activities conducted by CAI include

* providing the CAI proprietary test delivery system (TDS), including the Student Testing Interface, Test Administrator Interface, secure browser, and practice and training tests;
* hosting and providing support for its TDS, a component of the overall ELPAC Assessment Delivery System;
* scoring machine-scorable items; and
* providing high-level technology help desk support to LEAs for technology issues directly related to the TDS.

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

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

* Supporting and training county offices of education, LEAs, and charter schools
* Developing informational materials
* Recruiting and providing logistics for educator meetings
* Producing Administration and Scoring Training materials and videos, including an online training site for LEA coordinators and test examiners
* Producing *DFA*s

### Systems Overview and Functionality

#### Test Operations Management System

TOMS is the password-protected, web-based system used by LEAs to manage all aspects of ELPAC testing. TOMS serves various functions, including, but not limited to, the following:

* Managing test administration windows
* Assigning and managing ELPAC online user roles
* Managing student test assignments and accessibility resources
* Ordering test materials
* Viewing and downloading reports
* Reporting security incidents
* Providing a platform for authorized user access to secure materials, such as ELPAC *DFAs,* student data and results, ELPAC user information, and access to the ELPAC Security and Test Administration Incident Reporting System/Appeals process

TOMS receives student enrollment data and LEA and school hierarchy data from the California Longitudinal Pupil Achievement Data System (CALPADS) via daily feed. CALPADS is “a longitudinal data system used to maintain individual-level data including student demographics, course data, discipline, assessments, staff assignments, and other data for state and federal reporting.”[[2]](#footnote-3)

LEA staff involved in the administration of the ELPAC—such as LEA ELPAC coordinators, site ELPAC coordinators, and test examiners—are assigned varying levels of access to TOMS. For example, only an LEA ELPAC coordinator is given permission to assign and manage user roles; a test administrator or test examiner cannot download student reports. A description of user roles is explained more extensively in the *2022–23 Initial Alternate ELPAC Test Administration Manual* (CDE, 2023b).

#### Test Delivery System

The TDS is the means by which the statewide computer-based assessments are delivered to students. Components of the TDS include

* the Test Administrator Interface, the web browser–based application that allows test examiners to activate student assessments and monitor student testing;
* the Student Testing Interface, on which students take the assessment using the secure browser; and
* the secure browser, the computer-based application through which the Student Testing Interface may be accessed. (The secure browser prevents students from accessing other applications during testing.)

#### Test Results for California’s Assessments Website

The Test Results for California’s Assessments website is used by educators, families, researchers, and interested members of the public to view aggregated results from the Initial Alternate ELPAC. The primary purpose of the Test Results for California’s Assessments website is to provide users with access to results data for groups of students and to allow comparison of test result data for various student groups. Test scores for a given grade level are aggregated at the school, LEA or direct funded charter school, county, and state levels. The aggregated scores are generated for selected student groups of interest (e.g., gender, ethnicity, economic status, migrant status, and disability status) and for the total population.

### Overview of the Technical Report

This technical report addresses the characteristics of the 2022–23 administration of the Initial Alternate ELPAC and contains 10 additional chapters, as follows:

* [Chapter 2](#_Overview_of_Alternate) presents an overview of the processes involved in a testing cycle for the Initial Alternate ELPAC. This includes item development, test assembly, test administration, fairness and accessibility, generation of test scores, and score reports.
* [Chapter 3](#_Item_Development_and_1) 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](#_Toc122102494) describes the processes involved in test assembly, including test forms, psychometric criteria, and CDE forms review.
* [Chapter 5](#_Test_Administration) details the processes involved in the 2022–23 administration, with emphasis on efforts made to ensure the standardization of Initial Alternate ELPAC computer-based testing. It also describes the procedures followed to maintain test security throughout the test administration process.
* [Chapter 6](#_Standard_Setting) summarizes the standard setting process that established the base-year performance level scores.
* [Chapter 7](#_Scoring_and_Reporting) provides information on the scoring processes and summarizes the types of scores and score reports.
* [Chapter 8](#_Appendix_6.B:_Means) summarizes the statistical procedures conducted for the 2022–23 operational administration. These analyses include
* classical item analyses, and
* item response theory analyses.

This chapter also discusses the procedures designed to support the reliability and validity of score use and interpretations.

* [Chapter 9](#_Quality_Control) highlights the quality-control processes used at various stages of the administration of the Initial Alternate ELPAC, including item development, test form development, test administration, scoring procedures, psychometric analysis, and score reporting.
* [Chapter 10](#_In-Test_Survey_3) discusses where to find more information on the threshold score validation study.
* [Chapter 11](#_Continuous_and_Systematic) describes analysis and administration processes and features targeted for improvement during future test administrations.

### References

*California* *Code of Regulations,* Title 1, General *Education Code* Provisions, Division 1, Part 1, Chapter 3, Article 3.5, Section 313(a). (n.d.).

*California* *Code of Regulations,* Title 5, Education, Division 1, Chapter 11, Subchapter 7.6, Article 2. (n.d.).

California Department of Education. (2014). *California English language development standards: Kindergarten through grade 12.* California Department of Education website.

California Department of Education. (2023a, October). *English Language Proficiency Assessments for California (ELPAC).* California Department of Education website.

California Department of Education. (2023b). *Initial Alternate ELPAC test administration manual.* Sacramento, CA: California Department of Education.

California Department of Education. (2023c, August). *Organization.* California Department of Education website.

California Department of Education. (2023d, October). *State Board of Education responsibilities.* California Department of Education website.

## Overview of Initial Alternate ELPAC Processes

This chapter provides an overview of the processes implemented by ETS during a typical, full testing cycle for the Initial Alternate English Language Proficiency Assessments for California (ELPAC), including item development, test design, test administration, and scoring. The details on each step in the process will be presented in the subsequent chapters.

### Item Development

As part of the adaptation and alignment process, ETS developed test items for the Initial Alternate ELPAC in accordance with the *ETS* *Standards for Quality and Fairness* (ETS, 2014).

Refer to chapter 3 of the *2021–22 Alternate English Language Proficiency Assessments for California Operational Field Test Technical Report* for the scope of work usually completed that incorporates the full range of assessment development activities (California Department of Education [CDE], 2023a).

#### Selection of ELD Connectors for Development

The Initial Alternate ELPAC is designed to align with the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) via the English Language Development Connectors (ELD Connectors), which reduce the depth, breadth, and complexity of the standards, as appropriate for students with the most significant cognitive disabilities (CDE, 2014, 2019).

The development of the ELD Connectors began with a review of the Council of Chief State School Officers (CCSSO) English Language Proficiency Standards for English Learners (ELs) with Significant Cognitive Disabilities (CCSSO, 2019). A crosswalk was developed to link the CCSSO ELD Standards to the 2012 ELD Standards. The results of the development effort were ELD Connectors for each of the 2012 ELD Standards assessed on the ELPAC at each of the Initial Alternate ELPAC grade levels and grade spans (i.e., kindergarten, grade one, grade two, grade span three through five, grade span six through eight, and grade span nine through twelve). The development of the ELD Connectors was consistent with the approach used for the California Alternate Assessments for English language arts/‌literacy, mathematics, and science and was a necessary foundational step in the development of the Alternate ELPAC.

The Connectors development plan began with the creation of sample ELD Connectors by ETS, followed by reviews of the sample Connectors by the CDE and the Alternate ELPAC Test Design Advisory Team mentioned in section [*1.3 Test Content and Design*](#_Test_Content_and). The sample ELD Connectors were revised on the basis of the review feedback, and then ETS developed the remaining Connectors. The CDE reviewed the full range of Connectors, revisions were made as necessary, and an in-depth review of the ELD Connectors was made by California educators. The ELD Connectors were approved by the CDE in March 2019 and the California State Board of Education (SBE)—through the approval of the Alternate ELPAC blueprints—in May 2020.

#### Item Format

The Initial Alternate ELPAC includes the following primary computer-based item formats:

* **Selected-response items—**Students are instructed to select one choice. Most Initial Alternate ELPAC items have two or three response options. The items are assigned one point and are machine-scored.
* **Constructed-response items—**Students are instructed to respond to the test question, which is scored by the test examiner according to a rubric. Students receive scores of either two points, one point, or zero points on each item.

#### Task Type Specifications

The *Alternate ELPAC Task Type Specifications* describe the tasks contained in the assessment. Each task type is intended to measure specific ELD Connectors consistently (CDE, 2020a). The task type specifications were developed collaboratively by the CDE, ETS, and California educators and approved by the CDE in 2020. The task type specifications are reviewed annually so that refinements can be made to the development of new items based on requirements.

During item development, item developers were provided with the *Alternate ELPAC Task Type Specifications* and a style guide that contained detailed information about the specifications for the various components of item development. Refer to subsection [*3.1.2 Task Type Specifications*](#_Task_Type_Specifications) for detailed information about the task type specifications.

### Test Assembly

The 2022–23 operational assessment was assembled in accordance with the Alternate ELPAC blueprint, approved by the SBE in May 2020 (CDE, 2020b). This document provides guidance for both the Summative Alternate ELPAC and the Initial Alternate ELPAC.

The assembly began with the selection of seven task types, each with two or four items, for 24 operational items. The task types on the assessment appeared in sequential order according to the blueprint.

After the initial assembly, assessment developers reviewed the assembled forms using comprehensive checklists to evaluate blueprint alignment, item content, clueing and content overlap, and overall balance of content regarding disability, gender, and ethnicity representation; variety of item types; and so forth.

After assessment developers assembled and reviewed the draft test forms, the forms were submitted for psychometric review for consistency with the blueprint and form assembly specifications and received subsequent approval. Approved forms then received additional content and editorial reviews, including key checks and a review of related *Directions for Administration (DFAs),* before being submitted to the CDE for review and feedback. ETS worked with the CDE to make test form revisions until the CDE approved them.

#### Test Blueprint

The Initial Alternate ELPAC and Summative Alternate ELPAC follow a single blueprint, as described in the high-level test design for the Alternate ELPAC and as approved by the SBE in May 2020 (CDE, 2020b). The Alternate ELPAC test blueprint defines grade levels or grade spans, differentiates the task types based on the ELD Connectors, and outlines the linear sequence the task types appear on the assessment as they progress from low linguistic complexity to high linguistic complexity. The Alternate ELPAC test blueprint is unique to each grade level or grade span based on the combination of Connectors within each task type. The blueprint specifies the number of items and points for each task type, the total number of items, and the total number of points on each assessment according to standards (CDE, 2020b).

#### Test Length

There are 24 operational items on the Initial Alternate ELPAC at each grade level and grade span. Refer to [*Chapter 4: Test Assembly*](#_Toc122102494) for more details on test form assembly.

### Test Administration

The Initial Alternate ELPAC was administered using the secure browser and test delivery system, ensuring a secure, confidential, standardized, consistent, and appropriate administration for students. Additional information about the administration of the Initial Alternate ELPAC can be found in [*Chapter 5: Test Administration*](#_Test_Administration).

#### Test Security and Confidentiality

All operational assessments within the ELPAC System are secure. For the Initial Alternate ELPAC administration, every person having access to test materials maintained the security and confidentiality of the assessments. ETS’ internal Code of Ethics requires that all test information, including tangible materials (such as test items and test results), confidential files, processes, and activities were kept secure. To ensure security for all assessments that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI). A detailed description of the OTI and its mission is presented in subsection[*5.7.1 ETS’ Office of Testing Integrity*](#_ETS’_Office_of_2) in [*Chapter 5: Test Administration*](#_Test_Administration).

In the pursuit of enforcing secure practices, ETS strives to safeguard the various processes involved in an assessment development and administration cycle. Those processes are listed next. The practices related to each of the following security processes are discussed in detail in section [*5.7 Test Security and Confidentiality*](#_Test_Security_and):

* Procedures to maintain standardization of test security
* Test security monitoring
* Security of electronic files using a firewall
* Transfer of scores via secure data exchange
* Data management in the secure database
* Statistical analysis on secure servers
* Student confidentiality
* Student test results

#### Procedures to Maintain Standardization

ETS takes all necessary measures to ensure the standardization of administration of the Initial Alternate ELPAC.

The Initial Alternate ELPAC is administered in conjunction with the other assessments that compose the ELPAC System. ETS employs processes to ensure the standardization of an administration cycle; these processes are discussed in more detail in section [*5.3 User Roles and Standardization*](#_Toc120783965).

Staff at local educational agencies (LEAs) involved in the ELPAC administration include LEA ELPAC coordinators, site ELPAC coordinators, and test examiners. The responsibilities of each of the staff members are described in the *Initial Alternate ELPAC Test Administration Manual* (CDE, 2023d).

Several series of instructions regarding the ELPAC administration are compiled in detailed manuals and provided to the LEA staff. Such documents include, but are not limited to, the following:

* ***Initial Alternate ELPAC Test Administration Manual*—**This web-based manual provides test administration procedures and guidelines for LEA ELPAC coordinators and site ELPAC coordinators (CDE, 2023d). (Refer to [*5.3.4.3 Initial Alternate ELPAC Test Administration Manual*](#_Summative_Alternate_ELPAC) in [chapter 5](#_Test_Administration) for more information.)
* ***California Assessment of Student Performance and Progress (CAASPP) and ELPAC Test Operations Management System (TOMS) User Guide*—**This web-based manual provides instructions for TOMS, allowing LEA staff, including LEA ELPAC coordinators and site ELPAC coordinators, to perform several tasks, including adding and managing users, assigning assessments, and configuring computer-based student test settings (CDE, 2023b). (Refer to [*5.3.4.4 CAASPP and ELPAC Test Operations Management System User Guide*](#_CAASPP_and_ELPAC_2) in [chapter 5](#_Test_Administration) for more information.)
* ***Preparing for Administration—***This document includes planning and preparation content to assist test examiners with test preparation (CDE, 2023c). (Refer to [*5.3.4.1 Preparing for Administration*](#_Preparing_for_Administration) in [chapter 5](#_Test_Administration) for more information.)
* ***DFA*s—**These directions include test examiner directions and scripts for administering the assessments. They contain grade-specific and form-specific information needed by the test examiners during test sessions. (Refer to [*5.3.4.2 Directions for Administration*](#_Directions_for_Administration) in [chapter 5](#_Test_Administration) for more information.)

### Fairness and Accessibility

Several procedures are in place to ensure that the Initial Alternate ELPAC is fair and accessible to all students. This section provides information on the available accessibility resources.

#### Overview

All eligible students enrolled in a California public school participate in the ELPAC System of assessments, including students with disabilities. Additional resources are sometimes needed for these students. The CDE provides a full range of assessment resources for all students, including those who are students with disabilities.

#### Student Accessibility Resources

There are four different categories of student accessibility resources in the California assessment accessibility system, including universal tools, designated supports, accommodations, and unlisted resources that are permitted for use in ELPAC computer-based assessments. These are listed in the CDE California Assessment Accessibility Resources Matrix (Accessibility Matrix) (CDE, 2022).

**Universal tools** are available to all students. These resources may be turned on and off when embedded as part of the technology platform for the computer-based ELPAC on the basis of student preference and selection.

**Designated supports** are available to all students when determined as needed by an educator or team of educators, with parent/guardian and student input as appropriate, or when specified in the student’s individualized education program (IEP) or Section 504 plan.

**Accommodations** must be permitted on the ELPAC for all eligible students when specified in the student’s IEP or Section 504 plan.

**Unlisted resources** are non-embedded and made available if specified in the eligible student’s IEP or Section 504 plan and do not jeopardize test security, and only on approval by the CDE. An unlisted resource may change the construct being measured.

While most of the resources presented for the ELPAC computer-based assessments are available for the Initial Alternate ELPAC, there are a few resources that are not applicable because the Initial Alternate ELPAC is designed to be given one-on-one in the student’s language of instruction, using the student’s identified instructional resources. For example, the speech-to-text accommodation is not available for an alternate assessment.

Table 5.1 and table 5.2 present counts and percentages of students assigned designated supports, accommodations, and unlisted resources for the 2022–23 Initial Alternate ELPAC administration. Table 5.1 and table 5.2 were created using student demographic data in version 2 of the production data file (“P2”) updated on October 11, 2023.

The majority of students did not use any designated supports, accommodations, or unlisted resources.

### Scores

Individual student scores were reported for the 2022–23 Initial Alternate ELPAC administration. Student performance on the reporting scale was designated into one of the three performance levels described in subsection [*7.1.4 Performance Levels*](#_Performance_Levels). For information regarding score specifications and score reports, refer to [*Chapter 7: Scoring and Reporting*](#_Scoring_and_Reporting).

#### Score Reporting

TOMS is a secure website hosted by ETS that permits LEA users to manage aspects of ELPAC test administration such as test assignment and the assignment of test settings. TOMS also provides a secure means for LEA ELPAC coordinators to download Student Score Reports as PDF files.

#### Aggregation Procedures

To provide meaningful results to interested educators, Initial Alternate ELPAC scores for a given grade-level assessment were aggregated at the school, LEA or direct funded charter school, county, and state levels. State-level results are available on the Test Results for California’s Assessments website. The aggregated scores were presented for all students or selected demographic student groups.

Aggregated scores were generated by combining student scores at the state, LEA or direct funded charter school, or school level; combining student scores for all students; or by combining student scores for students who represent selected demographic student groups.

The aggregation procedures used to present Initial Alternate ELPAC results are described in section [*7.2 Overview of Score Aggregation Procedures*](#_Overview_of_Score). Aggregated results by demographic variables are presented in table 7.4 for kindergarten students. Demographic variables include (but are not limited to) gender, ethnicity, primary disability type, and economic status. There is not sufficient data to make student group comparisons meaningful at higher grade levels or grade spans.

The table shows the numbers of students with valid scores in each group, scale score means and standard deviations, and the percentage of students in each performance level. To protect student privacy, statistics are presented in the table as “N/A” when the number of students in the sample is 10 or fewer. Definitions for the demographic student groups included in this table are provided in table 7.3.

### Psychometric Analyses

Psychometric analyses were conducted on the data from the Initial Alternate ELPAC, including classical item analyses, item response theory (IRT) score scaling, and reliability analyses. The results of these analyses support understanding of item performance and internal structure of the assessment and provide validity evidence for both response processes and scoring. Detailed descriptions of these analyses are presented in [*Chapter 8: Psychometric Analyses*](#_Analyses_and_Results).

#### Description of the Classical Item Analyses

The psychometric analyses for the Initial Alternate ELPAC data included classical item analyses to evaluate the performance of the items. The classical item analyses included the computation of item difficulty indices, the item-total correlation indices, the omission rate of each item, and the proportion of test takers obtaining each score point for polytomous items. CDE-approved flagging rules based on these statistics identified items that were not performing as expected. A description of the classical item analyses procedure is provided in section [*8.2 Classical Item Analyses*](#_Demographic_Student_Group)*.*

#### Description of Item Response Theory Analyses

IRT is used to calibrate items, link item parameter estimates, scale or equate test scores across different forms or test administrations, evaluate item performance, build an item bank, and assemble test forms. Detailed information on the IRT models is included in section [*8.3 Item Response Theory Analyses*](#_Item_Response_Theory). As all items on this assessment came from the 2021–22 Alternate ELPAC operational field test administration, more information about the calibration and linking can be found in the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023a).

### References

California Department of Education. (2014). *California English language development standards: Kindergarten through grade 12.* California Department of Education website.

California Department of Education. (2019). *California English Language Development connectors for the Alternate English Language Proficiency Assessments for California.* California Department of Education website.

California Department of Education. (2020a). *Alternate English Language Proficiency Assessments for California task type specifications* [Unpublished manuscript]. Sacramento, CA: California Department of Education.

California Department of Education. (2020b). *Alternate English Language Proficiency Assessments for California test blueprint.* Approved by the California State Board of Education in May 2020. California Department of Education website.

California Department of Education. (2022). *California assessment accessibility resources matrix*. California Department of Education website.

California Department of Education. (2023a). *Alternate English Language Proficiency Assessments for California 2021–22 operational field test technical report*. Sacramento, CA: California Department of Education.

California Department of Education. (2023b). *CAASPP and ELPAC Test Operations Management System user guide*. Sacramento, CA: California Department of Education.

California Department of Education. (2023c). *Initial Alternate ELPAC preparing for administration*. Sacramento, CA: California Department of Education.

California Department of Education. (2023d). *Initial Alternate ELPAC test administration manual*. Sacramento, CA: California Department of Education.

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

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.

## Item Development and Review

This chapter discusses the detailed procedures of item development for the 2022–23 Initial Alternate English Language Proficiency Assessments for California (ELPAC) administration.

Assessment development activities described in this chapter are those typical during the item development cycle. Refer to chapter 3 of the *2021–22 Alternate English Language Proficiency Assessments for California Operational Field Test Technical Report* for the scope of work usually completed that incorporates the full range of assessment development activities (California Department of Education [CDE], 2023).

### Overview

The Initial Alternate ELPAC and the Summative Alternate ELPAC share the same item pool. The items used in the Initial Alternate ELPAC 2022–23 forms are from the items developed and field-tested in the 2021–22 Alternate ELPAC operational field test administration. The details of the item development activities for the Alternate ELPAC item pool such as the item writer workshops, item review meetings (IRMs), and data review meetings (DRMs) can be found in Chapter 3: Item Development and Review in the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023), which relates to the development of the item pool used by both the Initial Alternate ELPAC and Summative Alternate ELPAC.

In partnership with the Sacramento County Office of Education, ETS convened Alternate ELPAC item writer workshops and IRMs to develop test items for the Alternate ELPAC. In addition, ETS trained a small group of experienced contractors and California educators to draft Alternate ELPAC items. After the items went through ETS internal and CDE reviews, California educators reviewed the items during IRMs.

This section describes how California educators were selected and the process used to develop new items. IRMs were conducted in June 2019 and June 2020; some of those items were field-tested on the 2021–22 Alternate ELPAC operational field test. Of those items, some were used on the 2022–23 Initial Alternate ELPAC, and some are available for eventual operational use on the Summative Alternate ELPAC.

#### Preparation

Several test design tasks were conducted to prepare for the 2022–23 Initial Alternate ELPAC. These tasks occurred prior to item development and assessment development tasks. After a review of the draft Alternate ELPAC test blueprint, a high-level test design was developed, followed by a pilot study using cognitive lab methodology, and, finally, the task type specifications were created. Refer to subsection [*3.2.5.2 Cognitive Laboratory*](#_Cognitive_Laboratory) for information about this study. Additionally, the *Alternate ELPAC Pilot Using Cognitive Lab Methodology Study* report contains more details on how that study informed early item development efforts (CDE, 2020a).

#### Task Type Specifications

Task type specifications describe the types of tasks that are used on the Initial Alternate ELPAC by

* guiding item writers during item development,
* maintaining consistency and efficiency in item development,
* enabling accessibility considerations, and
* providing the CDE with a reference guide to use while reviewing items.

The 2022–23 Initial Alternate ELPAC contains seven task types (CDE, 2020b). Each task type required a student to answer questions about a passage or image to elicit information about the student’s English language proficiency (ELP). Each task type consists of two or more items that align with the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014) via the English Language Development Connectors (ELD Connectors). The test items within a task type are aligned to one or more primary ELD Connectors and, in some cases, secondary ELD Connectors.

The Initial Alternate ELPAC is designed to assess ELP, including the language domains of Listening, Speaking, Reading, and Writing, in an integrated manner; and report on ELP as a whole, not by individual language domains. Because of the intent to assign an overall score and to provide students with the flexibility to use their individually preferred communication modes, the Initial Alternate ELPAC test items are coded as either “receptive” or “expressive.”

##### Receptive (Listening and Reading)

Receptive test items require a student to demonstrate comprehension of a stimulus by selecting a response from two or three options; the student is not required to generate any language. Receptive items are selected-response items.

##### Expressive (Speaking and Writing)

Expressive test items require a student to communicate to others their understandings and ideas related to the stimulus, using an individually preferred expressive mode of communication. Expressive items can be multiple choice or constructed response.

#### Item Writing by Contractors

In 2019, ETS’ assessment specialists worked with a small group of contractors (i.e., outside item writers) who were fully trained, experienced item writers with a record of developing quality items for other ETS English language assessments. These contractors developed items in accordance with the *Alternate English Language Proficiency Assessments for California Task Type Specifications* (CDE, 2020b).

#### Item Writer Training

Item writer training is a vital part of establishing the validity chain for item and task development. In addition to relying on internal item writing experts for the Initial Alternate ELPAC, ETS recruited and trained educators in 2012 ELD Standards.

The three primary goals for the training were to

1. provide teachers with knowledge, via professional development on writing items and *Directions for Administration (DFA)* scripts, that they can use to help develop or refine their own classroom teaching and assessments;
2. ensure that teachers who successfully completed the training were ready to develop high-quality items for the Alternate ELPAC; and
3. leverage the experiences, perspectives, and expertise of the teachers in writing items for the Alternate ELPAC.

ETS held item writer training workshops to provide prospective item writers with professional development in several areas. A review of the general assessment development process gave trainees a sense of the total life cycle of an item.

Participants learned best practices in item writing to provide clarity within the item and avoid bias or sensitivity concerns, learned how to review a passage for item opportunities, and were introduced to how the new, innovative item types work.

Given that the trainees were California educators and educational leaders, ETS also emphasized incorporation of current effective teaching practices and instructional activities. Small-group and individual work generated sample items that the ETS facilitators then used in a large-group discussion to analyze and ascertain overall item quality. The ETS team also provided post hoc feedback via email and phone calls to trained item writers on further item samples and ideas submitted ahead of contractual item submissions.

For detailed information regarding item writer recruitment, refer to the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023).

### ETS Item Review Process

After items were drafted, ETS placed items and *DFA* scripts developed for the Alternate ELPAC through an extensive internal item review process designed to provide the best standards-based assessments possible. This section summarizes the item review process that confirmed the quality of Alternate ELPAC items.

#### Overview

Once an item was accepted for authoring, ETS employed a series of internal reviews. These reviews used established criteria to judge the quality of item content and to ensure that each item measured what it was intended to measure. These internal reviews also examined the overall quality of the items ahead of their being reviewed by the CDE and by educators at IRMs, which are described in more detail in section [*3.4 California Educator Review*](#_California_Educator_Review_1).

All items were entered into the Item Banking Information System (IBIS) with corresponding artwork and metadata. Within IBIS, items received content reviews by ETS’ assessment specialists and fairness and editorial reviews by ETS’ editors and fairness reviewers.

The CDE reviewed proposed changes to items in response to reviews by the participants of the IRMs to ensure the quality of the item pool. The CDE then gained access to Alternate ELPAC items and conducted reviews in IBIS. ETS revised items in response to comments from the CDE prior to using them in the assessment forms.

The ETS review process for the Alternate ELPAC includes the following; these tasks are described in the next subsections:

1. Content review
2. Accessibility review
3. Editorial review
4. Sensitivity and fairness review

Throughout this multistep item review process, the lead assessment specialists and development team members at ETS continually evaluated the activities and items for adherence to the rules for item development.

#### ETS Content Review

On all items ETS developed, assessment specialists conducted three reviews on items and stimuli. These assessment specialists verified thatthe items, *DFA* scripts, and stimuli were in compliance with ETS’ written guidelines for clarity, style, accuracy, and appropriateness for California students and were also in compliance with the approved item specifications, the *California Assessment of Student Performance and Progress (CAASPP) and ELPAC Item Review Acceptance Criteria* (ETS, 2019), and other ETS-produced procedures such as the ETS guidelines for fair tests and communications (2016). Assessment specialists reviewed each item in terms of the following characteristics:

* Relevance to the purpose of the assessment
* Match of each item to the item specifications, including the tier of item complexity
* Match of each item to the principles of quality item writing
* Match of each item to the identified standard or standards
* Difficulty of the item
* Accuracy of the content of the item
* Readability of the item or passage
* Grade-level and grade-span appropriateness of the item
* Appropriateness of any illustrations, graphs, or figures

Assessment specialists verified the classification of each item, both to evaluate the correctness of the classification and to confirm that the task posed by the item was relevant to the outcome it was intended to measure. The reviewers could accept the item and classification as written, suggest revisions, or recommend that the item be discarded. These steps occurred prior to the CDE’s review.

#### ETS Accessibility Review

The ETS Accessible Content & Inclusive Solutions team advised on accessibility of items and item types during the ETS content review. These experts on alternate test formats reviewed all items, with a focus on accessibility for all student populations, and provided potential refinement solutions to improve the accessibility in items and assessments.

#### ETS Editorial Review

After assessment specialists and researchers reviewed each item, a group of specially trained editors also reviewed each item in preparation for consideration by the CDE and the item review panelists. The editors checked items for clarity, correctness of language, appropriateness of language for the grade level or grade span assessed, adherence to the style guidelines, and conformity with accepted item-writing practices.

#### ETS Sensitivity and Fairness Review

##### Review

ETS’ assessment specialists who were specially trained to identify and edit or eliminate items that contained content or wording that could be construed to be offensive to, or biased against, members of specific student groups (e.g., ethnicity, race, or gender) conducted the next level of review (ETS, 2014, 2016). These trained staff members reviewed every item before the CDE and IRMs. Newly developed items were then submitted to the CDE for review prior to educator reviews.

The review process promoted a general responsiveness to the following:

* Cultural diversity
* Diversity of background, cultural tradition, and viewpoints to be found in the test-taking populations
* Changing roles and attitudes toward various groups
* Role of language in setting and changing attitudes toward various groups
* Topics that may be unsettling or otherwise distract the student from the content being measured, such as natural disasters, disease, or family discord
* Contributions of diverse groups (including ethnic and minority groups, individuals with disabilities, and women) to the history and culture of the United States and the achievements of individuals within these groups
* Item accessibility for language learners of diverse backgrounds
* Item accessibility for English learner students with the most significant cognitive disabilities

##### Cognitive Laboratory

Under the direction of the CDE, ETS conducted a pilot study using cognitive laboratory methodology across California in January 2020 (CDE, 2020a). Based on results of the study, minor adjustments to the task types were made, and further guidance for optional individualization was developed. For further information regarding the cognitive laboratory, which relates to the development of both the Summative Alternate ELPAC and Initial Alternate ELPAC, refer to the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023).

### California Department of Education Review

After ETS reviews of items were completed, the items were reviewed by the CDE content teams. CDE content experts reviewed the items using the same criteria used in the ETS reviews. After CDE reviews occurred, ETS made edits to the items based on the CDE feedback, and the items were then finalized for IRMs with California educators.

### California Educator Review

Each newly developed item is reviewed during the IRMs, which are held annually in the spring or early summer. Educators participate in the meetings to review the items for alignment to the standards and appropriateness for the designated grade level or grade span.

Educators can make one of three recommendations regarding each item: accept the item as is, accept the item with revisions, or reject the item. Whenever an item is recommended to be accepted with revisions, educators specify the revisions needed to improve the text or images and the reasons for the proposed revisions.

Refer to chapter 3 of the *2021–22 Alternate English Language Proficiency Assessments for California Operational Field Test Technical Report* for the scope of work usually completed that incorporates the full range of assessment development activities (CDE, 2023).

#### California Educators as Content Experts

During IRMs, California educators serve in an advisory role to the CDE and ETS and provide guidance on matters related to item development for the Alternate ELPAC. The IRMs take place annually before newly developed items are selected for field testing and eventual operational use and are facilitated by ETS’ content experts. Typically, 15 to 20 educators are recruited to attend each meeting.

In the IRMs, the item content, test examiner *DFAs,* and alternative text were presented. ETS facilitated a discussion with the educators for each item using the *CAASPP and ELPAC Item Review Acceptance Criteria* (ETS, 2019). The educators were responsible for reviewing all newly developed items for alignment with the ELD Connectors, which are based on the 2012 ELD Standards, but reduced in depth, breadth, and complexity for students with significant cognitive disabilities. IRM participants also reviewed the items for content accuracy, language clarity, and item quality. In their examination of test items, participants could raise concerns about the appropriateness of the items as related to the grade level, age, and cognitive ability of the test takers. Additionally, passages, items, and supporting graphics were evaluated for any potential bias or sensitivity concerns associated with disability, gender, race, ethnicity, religion, or socioeconomic status. ETS recorded educator feedback for each item and adjusted item content based on approval from the CDE.

For detailed information regarding item review panel recruitment, refer to the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023).

### Data Review Meeting

After the items were administered to students as part of the Alternate ELPAC operational field test, ETS identified the statistically flagged items and the associated statistics for review by the CDE and California educators. Educators reviewed these flagged items with their statistical data in the DRM. Based on these reviews and discussions during the meeting, the participants made a recommendation to accept or reject each flagged item. Educator recommendations were then reconciled with the CDE to make a final decision on whether to include particular flagged items in the operational pool for the Summative Alternate ELPAC and Initial Alternate ELPAC. For more details about the DRM process, refer to the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023).

### References

California Department of Education. (2014). *California English language development standards: Kindergarten through grade 12.* California Department of Education website.

California Department of Education. (2020a). *Alternate English Language Proficiency Assessments for California pilot using cognitive lab methodology study* [Unpublished report]. California Department of Education.

California Department of Education. (2020b). *Alternate English Language Proficiency Assessments for California task type specifications.* California Department of Education website.

California Department of Education. (2023). *Alternate English Language Proficiency Assessments for California 2021–22 operational field test technical report.* Sacramento, CA: California Department of Education.

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.

Educational Testing Service. (2016). *ETS guidelines for fair tests and communications*. Princeton, NJ: Educational Testing Service.

Educational Testing Service. (2019). *CAASPP and ELPAC item acceptance criteria* [Unpublished manuscript]. Princeton, NJ: Educational Testing Service.

## Test Assembly

This chapter discusses the detailed procedures of test assembly for the 2022–23 Initial Alternate English Language Proficiency Assessments for California (ELPAC) administration.

### Overview

ETS’ assessment specialists assembled the Initial Alternate ELPAC, which was reviewed and approved by the California Department of Education (CDE). This process began with the creation of the high-level test design (CDE, 2019b), which the California State Board of Education (SBE) approved in May 2019 and provides the following information about the Initial Alternate ELPAC:

* Assessment purposes
* Test-taking population
* Guiding principles
* Key assumptions
* Test design recommendations
* Online test administration
* Accessibility
* Task types
* Test and blueprint development specifications
* Scoring and reporting specifications

The test form assembly process is described in the form assembly specifications (CDE, 2021). The form assembly specifications detail the content characteristics, psychometric characteristics, and number of items to be used on the 2022–23 Initial Alternate ELPAC. ETS created the form assembly specifications that the CDE reviewed and approved.

### Assessment Development

The Initial Alternate ELPAC incorporates evidence-centered design, which is especially useful for the development of new constructs and prioritizes ongoing collection of validity evidence to show that the assessment measures what it is intended to measure. It also incorporates universal design principles to ensure that it would be accessible to the intended testing population. All items and tasks were developed to grade-level standards and the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014) via the English Language Development Connectors (ELD Connectors) (CDE, 2019a). The ELD Connectors provide an aligned expectation of English language proficiency (ELP) that has been reduced in depth, breadth, and complexity to be appropriate for students identified with the most significant cognitive disabilities. This approach is consistent with that of other alternate assessments developed for California, in which Connectors are used to define how content standards are to be interpreted for this testing population.

#### High-Level Test Design

##### Test Design Principles

Three principles guided the design of the Alternate ELPAC. The principles were based on discussions with, and feedback from, various interest holder groups and local educational agencies as well as the Alternate ELPAC Test Design Advisory Team mentioned in section [*1.3 Test Content and Design*](#_Test_Content_and). The guiding principles were as follows (CDE, 2020b):

1. The assessments must be designed to ensure that the intended test-taking population is able to demonstrate its ELP.
2. The test design must be tailored to the range of needs of the students with the most significant cognitive disabilities, including providing maximum accessibility as well as ensuring linguistic and cultural fairness and sensitivity.
3. The test design must take into consideration the testing burden for students and test examiners.

##### Task Types

The Initial Alternate ELPAC assesses the four domains of Listening, Reading, Speaking, and Writing. However, it does so in an integrated manner; that is, a single task type assesses multiple domains. Receptive items assess the Listening and Reading domains, while expressive items assess the Speaking and Writing domains.

For the Initial Alternate ELPAC, the term “task type” is used to categorize test items based on their content and the evidence of student language proficiency they are designed to gather (e.g., *Recognize and Use Common Words*). In contrast, the term “item type” is used to describe items based on the form they take in the test delivery system (TDS) (e.g., selected response or constructed response [CR]). Each Initial Alternate ELPAC task type contains multiple item types. The test questions within a task type are aligned to one or more primary and secondary ELD Connectors. Additionally, to ensure that English learner students with the most significant cognitive disabilities can fully access and participate in the Initial Alternate ELPAC, these receptive and expressive task types are assessed via students’ individually preferred receptive and expressive communication modes. Such a design—one that helps ensure maximum participation of all eligible test takers—helps to eliminate the need to provide domain exemptions.

The Initial Alternate ELPAC is administered one-on-one: one test examiner assessing one student at a time. The test examiner should be an educator familiar with the student and the student’s preferred communication mode(s). The one-on-one administration model facilitates a primary test-design feature of allowing a student to have the assessment administered in a way that provides access via an individually preferred communication mode(s). Communication modes include, but are not limited to, the following:

* Verbal communication
* Communication via sign language, eye gaze, facial expressions, gestures, picture exchange system
* Use of an assistive technology device or Augmentative and Alternative Communication device

Access to each domain is provided via the communication mode(s) that are used by, and are appropriate for, an individual student.

#### Test Blueprint

The Alternate ELPAC test blueprint provides guidance for the development of all Alternate ELPAC test forms, ensuring that they appropriately sample the knowledge, skills, and abilities defined by the 2012 ELD Standards via the ELD Connectors; provide enough score points to support reliable score reporting; and support a test form that is appropriate in length for the Alternate ELPAC testing population (CDE, 2020b). The blueprint specifies the order of task types, which are sequenced from simple (lowest) to more complex (highest) linguistic complexity. The Alternate ELPAC test blueprint is similar in format to the ELPAC test blueprint and includes the

* task types assessed;
* task type linguistic complexity level;
* aligned ELD Connectors;
* number of receptive items (by task type);
* number of receptive score points (by task type);
* number of expressive items (by task type);
* number of expressive score points (by task type);
* total number of items; and
* total number of points.

In January 2020, a pilot study using cognitive lab methodology (CDE, 2020a) was administered to inform the proposed *Test Blueprint for the Alternate ELPAC* (CDE, 2020b), which the SBE approved in May 2020.

Analysis of the pilot study results led to modifications of the Alternate ELPAC test blueprint. For example, all but one of the piloted task types was retained. In addition, test design features were added to increase the accessibility of expressive (CR) items. Optional individualization was added to the assessment to allow students who use picture cards and realia as a form of expressive language to communicate, as they do within the classroom setting. Additional pictures were added to text-only answer choices to increase access. Text was culled from the story or passage and added to item stems, lessening the need for student recall.

The preliminary decision to implement a single test blueprint for both the Summative Alternate ELPAC and the Initial Alternate ELPAC was confirmed as an appropriate means of assessing the ELP of the Alternate ELPAC student population. Seven task types were retained in the final Alternate ELPAC test blueprint, with 24 operational items and, for the Summative Alternate ELPAC only, six embedded field test items. The SBE approved the Alternate ELPAC test blueprints in May 2020, which was before the first operational field test administration of the Initial Alternate ELPAC on November 1, 2021.

The test blueprint provided information about the number of receptive and expressive items and points administered per task type within each grade level and grade span. The test blueprint also identified two types of alignment between task types and the ELD Connectors: “primary” and “secondary.” Primary alignment indicated there was 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 indicated a moderate or partial match between the standard and the item in language knowledge, skills, and abilities.

#### Form Assembly Specifications

The Initial Alternate ELPAC consisted of one test form for each grade level or grade span. Each test form contained 24 operational items, described in table 4.1.

Table 4.1 displays the

* task types,
* communication modes assessed by each task type,
* number of items on the test blueprint,
* available points per item, and
* total number of items for each initial form.

Table 4.1 Kindergarten Through Grade Twelve Initial Alternate ELPAC Forms Plan

|  |  |  |  |
| --- | --- | --- | --- |
| Task Type | Communication Mode | Number of Items | Available Points |
| *Recognize and Use Common Words* | Receptive | 1 | 1 |
| *Recognize and Use Common Words* | Expressive | 1 | 1–2 |
| *Communicate About Familiar Topics* | Receptive | 1 | 1 |
| *Communicate About Familiar Topics* | Expressive | 1 | 1–2 |
| *Understand a School Exchange* | Receptive | 3 | 3 |
| *Understand a School Exchange* | Expressive | 1 | 1–2 |
| *Describe a Routine* | Receptive | 3 | 3 |
| *Describe a Routine* | Expressive | 1 | 1–2 |
| *Understand and Express an Opinion* | Receptive | 2 | 2 |
| *Understand and Express an Opinion* | Expressive | 2 | 4 |
| *Interact with a Literary Text* | Receptive | 2 | 2 |
| *Interact with a Literary Text* | Expressive | 2 | 3 |
| *Interact with an Informational Text* | Receptive | 2 | 2 |
| *Interact with an Informational Text* | Expressive | 2 | 3 |
| **Totals:** | **N/A** | **24** | **28–32** |

After the form assembly specifications were approved, ETS’ assessment specialists assembled the assessments according to the specifications into form planners. ETS’ assessment specialists and psychometricians reviewed the form planners 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.

### Test Production Process

The high-level test design for the Initial Alternate ELPAC describes the Initial Alternate ELPAC as a computer-based, linear assessment (i.e., not adaptive). The test forms are assembled so that task types are presented in order of linguistic complexity, from simple (lowest) to more complex (highest). There are three linguistic complexity levels on the Initial Alternate ELPAC: low, medium, and high. The ELD Connectors, as well as the high linguistic complexity descriptors, provide expectations for students at the highest level. The high linguistic complexity descriptors are the precise measurable skills described in each ELD Connector. The test design allows for potential exit points once the test examiner determines the student’s highest limit of linguistic skills has been reached, which may occur during or after a task type (CDE, 2020b). The system allows for students to exit when they reach content in the test that is beyond their level of ELP.

#### Selection of Task Types and Items

From the eligible item pool, assessment developers selected items that, as a whole

* met the coverage specifications of the test blueprint,
* met the form-building guidelines developed by the ETS psychometrics team,
* represented a variety of accessible item types, and
* provided a wide variety of task type context.

#### Test Forms

There was one form for each grade level and grade span of the 2022–23 Initial Alternate ELPAC. Each form met the operational test blueprint.

Table 4.1 in subsection [*4.2.3 Form Assembly Specifications*](#_Form_Assembly_Specifications_1) provides an overview of the number of operational items and points by task type and communication mode.

#### Psychometric Criteria and Review

ETS’ psychometricians reviewed and confirmed that each test form was consistent with the form assembly specifications.

The following criteria were used to review all forms:

* Forms align with the Alternate ELPAC test blueprint.
* Items selected for use meet the following criteria:
* The range for *p*-values is between 0.20 and 0.95.
* Item-total correlations are greater than 0.20.
* Items flagged for C-DIF—differential item functioning—are used only when necessary to meet the test blueprint and with CDE approval.
* Item response theory *b*-parameter estimates are within the range of −4.0 to +4.0.
* Forms should have average item difficulty that can provide sufficient information around the threshold scores for the performance levels.
* The test characteristic curve and the test information function should look reasonable.
* Multiple-choice, single-select (MCSS) items’ correct response options are approximately equally distributed among the possible answer choices A, B, and C (i.e., key balance for MCSS).
* Key runs—four or more consecutive keys of a single response option (A, B, or C)—should not occur.

#### Content Review of Forms

After psychometric approval, the proposed assessment underwent two additional content reviews and one editorial review. The content reviewers were assessment developers who had not previously worked on the development of the test forms they were reviewing. These reviewers brought a fresh perspective to the review. They were given the appropriate materials and documentation to complete the following tasks:

* Verification of item keys
* Identification of possible clueing across the items
* Verification that individual items aligned with the 2012 ELD Standards, as interpreted for English learners (ELs) with significant cognitive disabilities through the ELD Connectors
* Verification of coverage of the 2012 ELD Standards, as interpreted for ELs with significant cognitive disabilities through the ELD Connectors
* Identification of any possible grammatical or production errors

#### California Department of Education Forms Review

The CDE used a gatekeeper process to review all test materials. Test materials for review and approval by the CDE included form planners, *Directions for Administration (DFAs),* and student-facing items in the TDS. All test materials were approved before they were made available for use.

For the reviews of form planners and the *DFAs,* ETS initiated the review by submitting materials to the CDE via the gatekeeper system, along with the criteria for the review. CDE consultants performed the initial review and returned comments and requests for revisions to ETS. ETS’ staff then revised the materials as requested and returned them to the CDE consultants, who reviewed the updated materials. If the test materials needed additional revisions, they were returned to ETS for further modifications.

Once CDE consultants found that the test materials met the review criteria, the CDE consultants submitted the test materials to the CDE administrator for approval. 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 again before they could be resubmitted to the CDE administrator for approval.

#### Configuration of the Test Delivery System

Once all the test reviews were completed and concerns, if any, had been resolved, the official ordered item sequence of the proposed forms was sent to Cambium Assessment, Inc. (CAI) for configuration of the TDS. Unlike other stages of the test production process, this stage must occur prior to every administration of the Initial Alternate ELPAC, even in the case of a form reuse.

Each item underwent an extensive platform review on different operating systems, such as Windows, Linux, and iOS, to ensure that the item’s appearance was consistent across all platforms.

The platform review was conducted by a team at CAI consisting of a team leader and several team members. The team leader presented the item as it was approved in ETS and CAI item banks. Each team member was assigned a different platform—hardware device and operating system—and reviewed the item to see that it rendered as expected. This platform review meeting ensured that all items were presented consistently to all students regardless of testing device or operating system for standardization of the test administration.

Prior to operational deployment, the testing system and content were deployed to a staging server where they were subject to user acceptance testing (UAT) by both ETS and CAI staff. The TDS UAT served as both a software evaluation and a content approval.

Following the UAT by ETS and CAI staff, separate UAT cycles were conducted by the CDE. The UAT review provided the CDE with an opportunity to interact with the exact assessment that would be administered to the students. The CDE had to approve the Initial Alternate ELPAC UAT before the assessment could be released for administration to students.

#### Test Form Delivery

The TDS is the means by which the statewide computer-based assessments are delivered to students. Components of the TDS include

* the Test Administrator Interface, the web browser–based application that allows test examiners to activate student assessments;
* the Student Testing Interface, on which students take the Initial Alternate ELPAC using the secure browser and with assistance from the test examiner as needed; and
* the secure browser, the web-based application through which the Student Testing Interface may be accessed. (The secure browser prevents a student from accessing unapproved applications and resources during testing.)

### References

California Department of Education. (2014). *California English language development standards: Kindergarten through grade 12.* California Department of Education website.

California Department of Education. (2019a). *California English language development connectors for the Alternate English Language Proficiency Assessments for California.* California Department of Education website.

California Department of Education. (2019b). *Proposed high-level test design for the Alternate English Language Proficiency Assessments for California.* California Department of Education website.

California Department of Education. (2020a). *Alternate English Language Proficiency Assessments for California pilot using cognitive lab methodology study* [Unpublished report]. California Department of Education.

California Department of Education. (2020b). *Alternate English Language Proficiency Assessments for California test blueprint.* Approved by the California State Board of Education in May 2020. California Department of Education website.

California Department of Education. (2021). *Form assembly specifications for the Initial Alternate English Language Proficiency Assessments for California* [Unpublished manuscript]. Sacramento, CA: California Department of Education.

## Test Administration

This chapter details the processes involved in the administration of the 2022–23 Initial Alternate English Language Proficiency Assessments for California (ELPAC). It also describes the procedures followed by ETS to maintain test security throughout the test administration process.

### Overview

The Initial Alternate ELPAC was administered to students in kindergarten through grade twelve in 2022–23 in conjunction with the other assessments that compose the ELPAC System.

In accordance with the procedures for the computer-based ELPAC, local educational agencies (LEAs) identified test examiners and entered the test examiners as users into the Test Operations Management System (TOMS). ETS provided LEA staff with the appropriate training materials, such as test administration manuals, videos, and webinars, to ensure that the LEA staff and test examiners understood how to administer the computer-based Initial Alternate ELPAC.

The testing window for the 2022–23 administration of the Initial Alternate ELPAC was planned for July 1, 2022, through June 30, 2023. Specific test administration schedules within that window were determined locally pursuant to *California Code of Regulations,* Title 5 (5*CCR*), Section 11518(q).

The Initial Alternate ELPAC did not allow remote testing, as it uses the one-on-one administration model that all other California alternate assessments use, which requires the assessment only be given when the test examiner and student are together in the same room. The one-on-one administration model allows for the test examiner to interact with the computer on behalf of the student as appropriate to the student’s individual needs and abilities. This ensures that the test examiner provides individualized support as needed for each student.

### Administration and Scoring Training

#### Overview

The training team for the Initial Alternate ELPAC Administration and Scoring Training (AST) created a completely virtual training model for the 2022–23 administration of the Initial Alternate ELPAC.

Every LEA that had an eligible Initial Alternate ELPAC student in California was required to complete the online LEA Certification course on the Moodle Training Site (Moodle). The new Initial Alternate ELPAC Certification course was created as a supplement to the Summative Alternate ELPAC Certification course that launched earlier in the school year. It contained only content unique to the initial assessment, such as determining student eligibility for the Initial Alternate ELPAC and what to do once testing is completed. Because test administration is identical for both the Summative Alternate ELPAC and the Initial Alternate ELPAC, annual completion of the summative training course is sufficient for the administration of both assessments. All aspects of the training courses and materials were developed for the first time.

The alternative to completion of the online LEA Certification course was to coordinate with another certified LEA via a Memorandum of Understanding stating that the certified LEA would either provide test examiner training or provide a trained test examiner to administer the assessment. The LEA ELPAC coordinator, or a designee, was responsible for overseeing all training for the LEA. This includes nonpublic schools and independent charter schools so that every eligible student is provided with an opportunity to test.

An online Moodle training site was developed as a restricted site that could be accessed only by LEA ELPAC coordinators, LEA lead trainers, Alternate ELPAC test examiners, and others requiring general training specific to the Initial Alternate ELPAC. Access to Moodle was restricted in two ways. First, it was restricted by the use of Moodle keys, which are passkeys used by LEA staff to access the various secure training courses; without the Moodle keys, the contents of Moodle cannot be accessed. Second, it was restricted because the Moodle keys are stored on a secure website that can be accessed only by logging on to the site using access codes provided to LEA ELPAC coordinators.

Moodle contained all resources needed to conduct test examiner training. Because test administration is identical for both the Summative Alternate ELPAC and the Initial Alternate ELPAC, annual completion of the summative training course is sufficient for the administration of both assessments. Annually and prior to administration of either the Initial Alternate ELPAC or the Summative Alternate ELPAC, test examiners were required to complete the Summative Alternate ELPAC—Test Examiner course. This course contains all the information necessary to administer either assessment including preparation, administration, individualization, and scoring or open-ended questions.

The following items list high-level assumptions for the 2022–23 Initial Alternate ELPAC training:

* All LEA ELPAC coordinators from LEAs that had an eligible student to take the Initial Alternate ELPAC were expected to complete, or designate staff to complete, the LEA certification training requirement.
* The Test Examiner’s course in Moodle was developed to provide test examiners with the necessary knowledge to properly administer both the Initial Alternate ELPAC and Summative Alternate ELPAC. LEAs were permitted to choose to either have their test examiners complete the online training course individually or provide their own local group training.

#### Goals

The four goals of the 2022–23 Initial Alternate ELPAC AST were to

1. introduce the Initial Alternate ELPAC to new testing staff in the field;
2. standardize the administration of the Initial Alternate ELPAC;
3. train LEA trainers to score the rubric-scored items accurately and reliably so that they could effectively train test examiners and other qualified persons to locally administer and score the Initial Alternate ELPAC rubric-scored items (which includes secondary test examiners who back-scored these rubric-scored items); and
4. provide resources so test examiners could prepare for, and individualize, test administration.

#### Local Educational Agency Training Requirement

All LEA ELPAC coordinators from LEAs with an eligible student to take the Initial Alternate ELPAC were expected to complete, or designate staff to complete, the LEA certification requirement.

##### Certification of Training

ETS provided the Sacramento County Office of Education (SCOE) with a list of LEAs that registered eligible students for the Initial Alternate ELPAC. Of the 201 LEAs with eligible students, 179 LEAs completed the LEA certification, which resulted in a completion rate of 89 percent. LEAs that had not completed training were tracked and reported regularly to ETS for follow-up. ETS reached out by phone and email to those LEAs that had not completed LEA certification training. After several attempts, ETS contacted the superintendents for the LEAs that had not completed their training. When that was unsuccessful, ETS provided the California Department of Education (CDE) with a list of LEAs that had not completed LEA certification during the testing window. The CDE followed up with those specific LEAs during that same period.

Between June 1, 2022, and May 10, 2023, 424 individuals completed the Initial Alternate ELPAC—LEA Certification course, representing 605 LEAs. These numbers differ from the numbers mentioned previously because they are not limited to only those LEAs that had registered eligible students. Instead, these numbers reflect the entirety of the individuals who completed the certifications and the LEAs with which they were connected.

##### Monitoring Test Examiner Calibration

Each LEA has a unique user group in Moodle, with each LEA being issued a unique enrollment key for each of the training courses. Each LEA ELPAC coordinator can designate Alternate ELPAC trainers within the site and request that trainers have access to view reports and monitor test examiner completion status.

The LEA ELPAC coordinator, or a designee, was responsible for overseeing test examiners’ certification progress and completion. Test examiners were emailed a certificate of completion upon successfully completing and passing calibration, and test examiners were expected to email their certificate to their LEA ELPAC coordinator. LEA ELPAC coordinators could also monitor test examiners’ progress in the Activity Completion Report and Grade Book in Moodle.

#### Materials on the Moodle Training Site

The Alternate ELPAC Moodle Training Site provided California LEAs with necessary training resources to train test examiners who would administer the Initial Alternate ELPAC. The Moodle site provided a password-protected, online platform where developed course materials were made available to ELPAC coordinators, trainers, and test examiners. LEA trainers who wished to conduct a group training session could download materials to prepare. ELPAC coordinators shared access to Moodle with the test examiners within the LEA. Test examiners used the site to review training materials and complete their individual certifications.

Training materials included, but were not limited to, the following:

* *Preparing for Administration (PFA)—*This nonsecure document contained much of the information found previously in the front portion of the *Directions for Administration (DFA)*. It was a stand-alone document designed for test planning and preparation. The *PFA* was posted in Moodle and on the ELPAC website.
* *Draft Initial Alternate ELPAC DFA*—SCOE used a draft version of the *DFA,* watermarked as “DRAFT,” to allow test examiners to plan and prepare for testing. This allowed LEAs to have draft versions of the secure *DFA*s to use while watching training videos and for reference during local training, since the final version of the *DFA* was not approved until after the AST was released. Once finalized in June, the final *DFAs* then were posted in TOMS, and the draft *DFAs* were removed from Moodle.
* Initial Alternate ELPAC Test Item Preview (TIP)—This secure booklet corresponded to the *Initial Alternate ELPAC DFA* for each grade level and grade span. Using the booklet, test examiners could view each item with the *DFA* to plan and prepare materials for each student’s test administration, with considerations to the student’s individual communication mode(s). It provided a rendering of what test examiners and students would see on the screen in the test delivery system (TDS). This TIP document was similar to a print-on-demand book.
* *Alternate ELPAC Practice Test DFA*—Links to the nonsecure practice test *DFA*s for each grade level and grade span were made available as appropriate in the resources section after each video listing and at the end of the course. They were used in the administration video and the videos about rubric-scored items. Some of the training modules used items from the practice test showing students being administered the assessment. LEAs were also encouraged to have their test examiners use the practice tests and *DFA*s
* in local training as an activity to familiarize test examiners with the assessment,
* with their students to familiarize the students with the structure and format of the assessment, and
* with their students to test individualization strategies for each student.
* Check for Understanding Quizzes––Quizzes were used after each video. The content of each quiz represented the key concepts of each video. Users were required to pass each quiz before accessing the next module and were provided with the correct answer as a reinforcement of the key concepts of the video.
* Alternate ELPAC Picture Cards—The Alternate ELPAC picture cards were created and posted for test examiners to use as appropriate for individualization. SCOE printed and shipped precut and sorted, reusable picture cards at no cost to LEAs that ordered them, thus reducing the financial burden on participating LEAs.

##### Training Module Content

The Initial Alternate ELPAC—LEA Certification course contained two modules with content specific to the Initial Alternate ELPAC:

* Module 1—Identifying Potential English Learners
* Module 2—After Testing

What follows are descriptions of the modules:

* **Module 1—Identifying Potential English Learners Module**
* This module discussed the process of identifying potential English learner (EL) students. The video described the first procedures that LEAs must follow in identifying potential EL students with the most significant cognitive disabilities. Topics included the home language survey, how to access the California Longitudinal Pupil Achievement Data System (CALPADS) to determine eligibility for the assessment, the choice between administering the general Initial ELPAC or the Initial Alternate ELPAC, and the timeline for completing the identification process.
* **Module 2—After Testing Module**
* This module discussed what to do after an LEA has begun administering the Initial Alternate ELPAC. Topics included tracking completion status and parent/‌guardian notification of results.

### User Roles and Standardization

The test administration procedures were designed so that the assessments are administered in a standardized manner. ETS took all necessary measures to ensure the standardization of test administration, as described in this section.

#### Local Educational Agency ELPAC Coordinator

An LEA ELPAC coordinator was designated by the district superintendent or charter school administrator at the beginning of the 2022–23 school year. LEAs include public school districts, California State Board of Education–authorized charter schools, county office of education 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 5*CCR* Section 11518.40, their responsibilities included

* adding site ELPAC coordinators and test examiners into TOMS;
* training site ELPAC coordinators and test examiners regarding the state requirements and ELPAC administration as well as security policies and procedures;
* providing checklists for site ELPAC coordinators and test examiners to review in preparation for administering the summative assessments;
* overseeing test administration activities;
* reporting test security incidents (including testing irregularities) to the CDE using the online Security and Test Administration Incident Reporting System (STAIRS)/Appeals process;
* requesting an Appeal (if indicated by TOMS prompts while reporting an incident using the STAIRS/Appeals process);
* 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 and supplemental test materials in TOMS; and
* ensuring adequate test materials were on hand and redistributed throughout the LEA during the testing window as needed.

#### Site ELPAC Coordinator

A site ELPAC coordinator is trained by the LEA ELPAC coordinator for each test site (5*CCR* Section 11518.40[b][7]). A site ELPAC coordinator must be an employee of the LEA and must sign a security agreement (5 *CCR* Section 11518.45[b][3]).

A test site coordinator was responsible for identifying test examiners and ensuring that they have signed *ELPAC* *Test Security Affidavits* (5 *CCR* Section 11518.45[b][3]). A site ELPAC coordinator’s duties may have included

* adding test examiners into TOMS;
* entering test settings for students;
* creating testing schedules and procedures for a school consistent with state and LEA policies;
* working with technology staff to ensure secure browsers are installed and any technical issues are resolved;
* monitoring testing progress during the testing window and ensuring all students take the Initial Alternate ELPAC, as appropriate;
* coordinating and verifying the correction of student data errors in CALPADS;
* ensuring a student’s test session is rescheduled, if necessary;
* addressing testing problems;
* reporting test security incidents (including testing irregularities) to the CDE using the online STAIRS/Appeals process;
* overseeing administration activities at a school site; and
* requesting an Appeal (if indicated by TOMS prompts while reporting an incident using the STAIRS/Appeals process).

#### Test Examiner

Test examiners were identified by site ELPAC coordinators as individuals who would administer the Initial Alternate ELPAC.

A test examiner must have signed a security affidavit (5 *CCR* Section 11518.50[d]).

A test examiner’s duties may have included

* ensuring the physical conditions of the testing room meet the criteria for a secure test environment;
* administering the ELPAC, including the Initial Alternate ELPAC;
* reporting all test security incidents to the site ELPAC coordinator and LEA ELPAC coordinator in a manner consistent with state and LEA policies;
* viewing student information prior to testing to ensure that the correct student receives the proper assessment with appropriate resources and reporting potential data errors to site ELPAC coordinators and LEA ELPAC coordinators;
* monitoring student progress throughout the test session using the Test Administrator Interface; and
* fully complying with all directions provided in the *DFAs* for the Initial Alternate ELPAC (CDE, 2023f).

#### Instructions for Test Administration

##### *Preparing for Administration*

The nonsecure *PFA* documents (CDE, 2023e) contained the planning and preparation content from the *DFAs* and were posted on the Manuals and Instructions web page on the ELPAC website. The *PFA* was used by test examiners to prepare for the test administration and to become familiar with testing guidelines.

The *PFA* included the following:

* Administration notes
* Linked resources
* Necessary testing materials
* Information about student engagement
* Use of the “Mark as No Response” option

##### *Directions for Administration*

Test examiners were required to use the *DFAs*, housed securely in TOMS, to administer assessments to all eligible students. There was one *DFA* for each grade-level and grade-span assessment. Test examiners could only access the *DFAs* for the form their LEA was assigned to. LEA ELPAC coordinators could access all *DFAs*.

##### *Initial Alternate ELPAC Test Administration Manual*

The *Initial Alternate ELPAC Test Administration Manual* (CDE, 2023f) contained information and instructions on overall procedures and guidelines for all LEA and test site staff involved in the administration of computer-based assessments. Sections included the following topics:

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

Appendices included definitions of common terms and descriptions of different aspects of the assessment and systems associated with the assessment.

##### *CAASPP and ELPAC Test Operations Management System User Guide*

TOMS is a web-based application that allows LEA ELPAC coordinators to set up test administrations, add and manage users, and submit computer-based student test settings.

TOMS modules described in the *TOMS User Guide* included the following (CDE, 2023d):

* **Adding and Managing Users—**This module allowed LEA ELPAC coordinators to add site ELPAC coordinators and test examiners to TOMS so that the designated user could administer, monitor, and manage the ELPAC computer-based assessments.
* **Reports—**This module allowed LEA ELPAC coordinators and site ELPAC coordinators access to the various reports in TOMS.
* **STAIRS/Appeals—**This module allowed LEA ELPAC coordinators and site ELPAC coordinators access to create new STAIRS cases or search for STAIRS/Appeals cases.
* **Student Profile—**This module allowed LEA ELPAC coordinators, site ELPAC coordinators, and test examiners to view and manage student’s test assignments and test settings.

##### Other System Manuals

Other manuals were created to assist LEA ELPAC coordinators and others with the technological components of the ELPAC System and are listed next.

* ***California Assessment of Student Performance and Progress (CAASPP) and ELPAC Technical Specifications and Configuration Guide for Online Testing*—**This manual provided information, tools, and recommended configuration details to help technology staff prepare computers and install the secure browser to be used for the computer-based ELPAC (CDE, 2023c).
* ***CAASPP and ELPAC Security Incidents and Appeals Procedure Guide*—**This manual provided information on how to report a testing incident and submit an Appeal to reset, reopen, invalidate, or restore individual computer-based student assessments (CDE, 2023b).
* ***CAASPP and ELPAC Accessibility Guide*—**This manual provided descriptions of the accessibility features for computer-based assessments as well as information about supported hardware and software requirements for administering assessments to students using accessibility resources, including those with a braille accommodation using Job Access With Speech® (software) or a braille embosser (hardware) (CDE, 2023a).

### Local Educational Agency Training

Each year, ETS, in collaboration with the CDE and its Assessment Validity and Outreach contractor, SCOE, establishes and implements a comprehensive training plan for LEA assessment staff and educators on all aspects of the assessment program. The ETS and SCOE annual training plans specify the audience, topics, frequency, and mode (synchronous or asynchronous) of the training, including such elements as format, participants, and organization.

Knowing that educators were confronted with challenges daily that put additional demands on their time, ETS and SCOE made every effort to make the information available in a variety of ways that allowed educators access to training at a time that was responsive to their varying circumstances. This included offering training events on multiple days and times, livestreaming events, recording and archiving training, and converting training to self-paced modules that could be taken any time, at the learner’s convenience.

All training opportunities were posted in one centralized location on the ELPAC website. LEA staff were able to register for training opportunities in one place, on the Upcoming Training Opportunities web page. Archived training was made available on the Past Training Opportunities web page, making it easier for educators to find a training they missed, and providing easier access to recorded training. ETS also employed a new strategy for providing access to training materials. Participants could register to receive a copy of the training materials without registering to attend a live training. Training materials were developed in such a way that educators could consume the information independently by reading through materials.

#### Synchronous and Asynchronous Training

All synchronous training was offered on Zoom, recorded, and made available for on-demand viewing. Zoom provides an opportunity for educators to ask questions and get answers in real time. Coffee Sessions were livestreamed on YouTube.

In response to an environment where educators had competing priorities to juggle, ETS and SCOE used various strategies to increase engagement during synchronous trainings. Live polls were presented to get real-time feedback about attendees’ knowledge of a particular topic, allowing presenters to tailor presentations to the audience’s level of understanding. The chat functionality was enabled to give participants an opportunity to interact with each other or to provide open-ended feedback, or it was disabled to minimize distraction and drive attendees’ focus to the information being presented. Breakout groups were used in smaller group trainings, as appropriate. Breaks and processing time were incorporated into presentations to give attendees opportunities to attend to other responsibilities that might result as part of their work environment.

Working closely with the CDE, ETS and SCOE continued to provide informal support to educators by offering monthly Coffee Sessions. Coffee Sessions included CDE and ETS’ staff who could answer questions about all aspects of testing. ETS also offered several Office Hours for coordinators where support staff were generally available from 9 a.m. to 3 p.m., allowing coordinators to join as needed and get customized support. SCOE continued to offer Assessment and Accountability Information Meetings intended to provide LEA coordinators with regular updates about California’s assessment and accountability systems. All trainings and meetings were recorded and archived for on-demand viewing on the Past Training Opportunities web page on the ELPAC website.

#### Videos and Guides

ETS produced videos on various aspects of administering the ELPAC, including how to perform functions within TOMS, such as setting up a test administration window, adding users, assigning assessments to students, and uploading test settings. SCOE produced the accompanying quick reference guides, providing multiple avenues of support for educators administering the assessments.

In addition to the standard administration videos, ETS produced additional videos to support administration. Some videos were geared toward parents/guardians to help them understand the assessment’s purpose. Other videos were intended to help coordinators or other users complete a process, such as administering a practice or training test, starting and stopping a test session, how to monitor student completion, and how to complete second scoring that is required for some of the assessments. This list is a sampling of the available videos intended to capture the major areas of support for various interest holders. The comprehensive suite of training videos can be found on the ELPAC Videos and Quick Reference Guides web page.

#### Training for Proper Identification and Assignment of Designated Supports and Accommodations

ETS developed a video with LEA staff to help California educators learn more about the importance of implementing ELPAC accessibility resources and best practices used by educators in the field. The “Importance of Implementing CAASPP and ELPAC Accessibility Resources: Voices from Educators” video was available on the Videos and Quick Reference Guides web page on the ELPAC website.

ETS also produced short demonstration videos for every embedded accessibility resource, demonstrating how to use the resource for educators, students, and parents/guardians. The videos were available in both English and Spanish on the Accessibility Resources Demonstration Videos web page on the ELPAC website. Demonstration videos were also created for the most frequently used non-embedded accessibility resources. These videos were linked within the Individual Student Assessment Accessibility Profile (ISAAP) Tool, increasing access to the demonstration videos. Educators using the ISAAP Tool to determine the student’s needs could view the corresponding demonstration video without having to navigate away from the tool.

A video on how to use the ISAAP Tool was also available to support educators in the process of creating an individual student profile and matching accessibility resources to student needs to ensure a fair and valid testing experience for all students.

For the 2022–23 ELPAC administration, ETS produced a two-part asynchronous training module. Module A, Matching Accessibility Resources to Students’ Needs, focused on providing participants with an understanding of the importance of accessibility resources, the categories of accessibility resources, and the process for matching students with appropriate accessibility resources for daily instruction and on assessments. Module B, Using Accessibility Resources in Daily Instruction, focused on the importance of removing barriers to student learning and using accessibility resources in daily instruction. Educators could complete the training independently or had the option to attend one of two live sessions held by ETS to extend and deepen the learning experience.

At the California Assessment Conference, SCOE offered two sessions on accessibility. “Leveraging UDL and Accessibility Resources to Improve Teaching and Learning” explored Universal Design for Learning (UDL) principles to help remove barriers to student learning and provided data collection tools to participants. The session on “Introduction to Accessibility and the ISAAP Tool” provided participants with the most up-to-date information regarding accessibility resources and offered a live practical approach to identifying and matching accessibility resources to students using the ISAAP Tool. The conference also included some shared practices sessions focused on accessibility.

#### Feedback for Continuous Improvement Survey

The ELPAC program solicits feedback annually from various interest holder groups, including LEA ELPAC coordinators, site ELPAC coordinators, and test examiners, through the CAASPP and ELPAC Feedback for Continuous Improvement Survey. Feedback was collected via a post-test survey sent to more than 275,000 California educators and through focus groups. Educators provided valuable feedback for potential improvements to the future administration of CAASPP and the ELPAC—one or both—by reporting some lessons they learned in 2022–23.

Improvements made in response to survey results are detailed in [chapter 11](#_Continuous_and_Systematic). The CDE and ETS used key recommendations from educators to implement positive changes in the next test administration year.

##### Overview

LEA and site ELPAC coordinators, as well as test examiners, were invited to participate in the survey. The California educators who responded provided specific, actionable insights about their test administration experience. This survey gathered information and data from educators who were part of the administration of CAASPP, the ELPAC, or both programs. Its goal was to highlight successes and identify areas for improvement, both immediate and long term.

Overall, California educators continue to express positive experiences in their preparations for administering CAASPP and the ELPAC.

##### Communication

During the 2022–23 test administration year, the CDE and ETS continued to streamline communications and provide LEAs with relevant information throughout the year. CAASPP and ELPAC monthly communications were sent throughout the administration with timely reminders and training announcements. In addition, proactive communications were sent to help remind LEA ELPAC coordinators of important actions needed for a successful administration, such as reminders to set up a test administration window, order materials, or enter scores into the Data Entry Interface, if needed.

### Accessibility Resources

The Every Student Succeeds Act reaffirms the importance of ensuring that assessments are accessible to special populations, and the Individuals with Disabilities Education Act lays out monitoring requirements for students with disabilities. This section describes the accessibility resources used to support students in the Initial Alternate ELPAC, as well as the procedures to identify and assign students with accommodations and designated supports. Finally, the number of students who were assigned accessibility resources was reported on the basis of available data.

The 2022–23 Initial Alternate ELPAC offered commonly used accessibility resources available through the ELPAC computer-based testing platform, where applicable for the tested construct.

#### Accessibility Resource Categories

The purpose of universal tools, designated supports, and accommodations in testing is to provide *all* students with the opportunity to demonstrate what they know and what they are able to do. Universal tools, designated supports, and accommodations minimize or remove barriers that could otherwise prevent students from demonstrating their knowledge, skills, and achievement in a specific item type (expressive or receptive).

The CDE’s *California Assessment Accessibility Resources Matrix* (Accessibility Matrix) (CDE, 2022) is intended for school-level personnel and individualized education program (IEP) and Section 504 plan teams to select and administer the appropriate universal tools, designated supports, and accommodations as deemed necessary for individual students.

##### Universal Tools

Universal tools were available to all students by default, although they could be disabled if a student found them distracting. Each universal tool fell into one of two categories: embedded and non-embedded. Embedded universal tools were provided through the TDS (through the ELPAC secure browser), although they could be turned off by a test examiner.

The universal tools in the following subsections were available in the 2022–23 Initial Alternate ELPAC administration.

###### Embedded

The following embedded universal tools were available to students testing in the secure browser:

* Breaks
* Digital notepad
* Expandable items
* Expandable passages
* Highlighter
* Keyboard navigation
* Line reader (grades three through twelve)
* Mark for review (grades two through twelve)
* Strikethrough (grades three through twelve)
* Zoom (in or out)

###### Non-Embedded

The following non-embedded universal tools were available to students testing in the secure browser:

* Breaks
* Oral clarification of test directions by the test examiner in English
* Scratch paper
* Test navigation assistant

##### Designated Supports

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. These are assigned through the test settings in TOMS. The designated supports each fell into one of two categories: embedded and non-embedded. Embedded designated supports were provided through the Student Testing Interface (through the ELPAC secure browser).

The designated supports in the following subsections were available in the 2022–23 Initial Alternate ELPAC administration.

###### Embedded

The following embedded designated supports were available to students testing in the secure browser:

* Color contrast
* Masking
* Mouse pointer (size and color)
* Permissive mode
* Print size
* Streamline
* Turn off any universal tool(s)

###### Non-Embedded

The following non-embedded designated supports were available to students testing in the secure browser:

* American Sign Language or Manually Coded English
* Designated interface assistant
* Magnification
* Masking
* Medical supports
* Noise buffers
* Print-on-demand
* Read aloud
* Separate setting
* Simplified test directions
* Translated test directions (including American Sign Language or Manually Coded English)

##### Accommodations

Accommodations are changes in procedures or materials that increased equitable access during ELPAC administration and are permitted to all eligible students if specified in the student’s IEP or Section 504 plan. Assessment accommodations for students who needed them generated valid assessment results; they allowed these students to show what they know and can do. Accommodations did not compromise the learning expectations, construct, grade-level standard, or intended outcome of the assessments.

The accommodations in the following subsections were available in the 2022–23 Initial Alternate ELPAC administration.

###### Embedded

Because the Initial Alternate ELPAC was designed specifically for administration to students with the most significant cognitive disabilities and is administered one-on-one, there were no embedded accommodations available to students.

###### Non-Embedded

The following non-embedded accommodations were available to students testing in the secure browser:

* Additional instructional supports and resources for alternate assessments
* Alternate response options
* American Sign Language or Manually Coded English
* Breaks

##### Unlisted Resources

An unlisted resource is an instructional support a student regularly uses in daily instruction, assessment, or both, and has not been previously identified as a universal tool, designated support, or accommodation. The Accessibility Matrix included an inventory of unlisted resources that were already identified and were preapproved (CDE, 2022). During the 2022–23 ELPAC administration, an LEA ELPAC coordinator or a site ELPAC coordinator would use TOMS to submit a request for use of an unlisted resource. A preidentified, preapproved unlisted resource request was automatically approved. A request for an unlisted resource that was not preidentified was sent to the CDE for review and adjudication.

Unlisted resources are non-embedded resources that are made available if specified in the eligible student’s IEP or Section 504 plan and only upon approval by the CDE. Unlisted resources that changed the construct of an assessment and were approved were flagged as causing a change in construct. The lowest obtainable scale score (LOSS) would be assigned to the Initial Alternate ELPAC with the unlisted resource that changes the construct, the student’s score status would remain valid, and the student’s scale score would be reported but appear on the Student Score Report (SSR) with an asterisk and a footnote that the assessment was administered under conditions that resulted in a score that may not be an accurate representation of the student’s achievement.

Note that there were no preapproved unlisted resources associated with the Initial Alternate ELPAC.

The LEA ELPAC coordinator or site ELPAC coordinator was required to submit a request for the use of an unlisted resource to the CDE a minimum of 10 business days before the student’s first day of testing. The LOSS was reported for the affected domain when administrations included unlisted resources that changed the construct of that assessment.

#### Identification and Selection

All eligible students enrolled in a California public school participate in the ELPAC System. The CDE Accessibility Matrix (CDE, 2022) is intended for school-level personnel and IEP and Section 504 plan teams to select and administer the appropriate universal tools, designated supports, and accommodations as deemed necessary for individual students.[[3]](#footnote-4)

The full list of the universal tools, designated supports, and accommodations used in ELPAC computer-based assessments, including the Initial Alternate ELPAC, is documented in the Accessibility Matrix. Most embedded and non-embedded universal tools, designated supports, and accommodations listed in parts 1, 2, and 3 of the Accessibility Matrix are available for the Initial Alternate ELPAC through the computer-based testing interface or, in the case of non-embedded resources, from the school or LEA. Part 5 of the Accessibility Matrix includes approved unlisted resources. School-level personnel, IEP teams, and Section 504 teams used the Accessibility Matrix when deciding how best to support the student’s test-taking experience.

Test examiners are given the opportunity to administer the Alternate ELPAC practice and training tests so that students have the opportunity to familiarize themselves with a designated support or accommodation prior to testing.

Additional guidance for maximizing accessibility for students taking the Initial Alternate ELPAC was provided in the *Alternate ELPAC Accessibility and Accommodations Guidelines* (CDE, 2021)*.* It was developed by using the California Alternate Assessment accessibility guidelines in conjunction with the considerations for Initial Alternate ELPAC student population and test design.

#### Assignment

Designated supports and accommodations are assigned to individual students on the basis of identified student need. Such assignments are implemented in TOMS by the LEA ELPAC coordinator or site ELPAC coordinator, either through individual assignment through the student’s profile in TOMS or in a batch upload for multiple students. When the batch upload process was used, settings were uploaded into TOMS using a spreadsheet with data that had either been entered into a template downloaded from TOMS; or created by selecting and entering information into the web-based ISAAP Tool. The ISAAP Tool could be used by LEAs in conjunction with the *2022–23* CAASPP and ELPAC Accessibility Guide (CDE, 2023a), as well as with state regulations and policies (such as the Accessibility Matrix) related to assessment accessibility*.*

The embedded designated supports and accommodations were delivered to the student through the TDS at the time of testing; the non-embedded designated supports and accommodations were provided at the time of testing to the student by the LEA. Refer to section [*1.8 Systems Overview and Functionality*](#_Systems_Overview_and_3) in [*Chapter 1: Introduction*](#_Chapter_1:_Introduction) for more details regarding the TDS.

Once a student’s IEP or Section 504 plan team decided which accessibility resource(s) the student should use, LEA ELPAC coordinators and site ELPAC coordinators used TOMS to assign designated supports and accommodations to students prior to the start of a test session.

There were three ways a student’s accessibility resource(s) could be assigned:

1. Using the ISAAP Tool to identify the accessibility resource(s) and then uploading the spreadsheet it creates into TOMS (This process is discussed in more detail in subsection [*5.5.2 Identification and Selection*](#_Identification_and_Selection_1).)
2. Using the Online Student Test Settings template to enter students’ assignments and then uploading the spreadsheet into TOMS
3. Entering assignments for each student individually in TOMS

If a student’s IEP or Section 504 plan team identified and designated a resource not identified in the CDE Accessibility Matrix, the LEA ELPAC coordinator or site ELPAC coordinator needed to submit a request for an unlisted resource to be approved by the CDE. The CDE then determined whether the requested unlisted resource changed the construct being measured before the student started testing.

Table 5.1 and table 5.2 provide information on the number of students who were assigned accommodations and designated supports.

Table 5.1 Assignment of Accommodations and Designated Supports—Kindergarten Through Grade Two

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accessibility Resource | Kindergarten: N | Kindergarten: % of Total Tested | Grade 1: N | Grade 1: % of Total Tested | Grade 2: N | Grade 2: % of Total Tested |
| Non-Embedded Accommodation—Additional Instructional Supports and Resources for Alternate Assessments | 6 | 1% | 1 | 2% | 0 | N/A |
| Non-Embedded Accommodation—Alternate Response Options | 15 | 2% | 1 | 2% | 1 | 5% |
| Non-Embedded Accommodation—ASL or Manually Coded English | 2 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Breaks | 11 | 2% | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Scribe Items | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Unlisted Resources | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Color Contrast | 1 | 0% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Masking | 2 | 0% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Mouse Pointer (Size and Color) | 3 | 0% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Permissive Mode | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Print Size | 2 | 0% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Streamline | 2 | 0% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Turn Off Any Universal Tool(s) | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Color Contrast | 1 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Color Overlay | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Designated Interface Assistant | 1 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Magnification | 1 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Masking | 1 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Medical Supports | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Noise Buffers | 4 | 1% | 1 | 2% | 0 | N/A |
| Non-Embedded Designated Support—Print on demand | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Read-Aloud Items | 4 | 1% | 1 | 2% | 0 | N/A |
| Non-Embedded Designated Support—Separate Setting | 24 | 4% | 2 | 3% | 1 | 5% |
| Non-Embedded Designated Support—Simplified Test Directions | 17 | 2% | 1 | 2% | 0 | N/A |
| Non-Embedded Designated Support—Translated Test Directions (including ASL) | 6 | 1% | 0 | N/A | 0 | N/A |
| **Total Students Tested:** | **103** | **N/A** | **7** | **N/A** | **2** | **N/A** |

Table 5.2 Assignment of Accommodations and Designated Supports—Grade Three Through Grade Twelve

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accessibility Resource | Grade Span 3–5: N | Grade Span 3–5: % of Total Tested | Grade Span 6–8: N | Grade Span 6–8: % of Total Tested | Grade Span 9–12: N | Grade Span 9–12: % of Total Tested |
| Non-Embedded Accommodation—Additional Instructional Supports and Resources for Alternate Assessments | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Alternate Response Options | 0 | N/A | 0 | N/A | 1 | 4% |
| Non-Embedded Accommodation—ASL or Manually Coded English | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Breaks | 0 | N/A | 0 | N/A | 1 | 4% |
| Non-Embedded Accommodation—Scribe Items | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Accommodation—Unlisted Resources | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Color Contrast | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Masking | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Mouse Pointer (Size and Color) | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Permissive Mode | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Print Size | 1 | 2% | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Streamline | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Support—Turn Off Any Universal Tool(s) | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Color Contrast | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Color Overlay | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Designated Interface Assistant | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Magnification | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Masking | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Medical Supports | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Noise Buffers | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Print on demand | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Read-Aloud Items | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Support—Separate Setting | 0 | N/A | 0 | N/A | 1 | 4% |
| Non-Embedded Designated Support—Simplified Test Directions | 0 | N/A | 0 | N/A | 1 | 4% |
| Non-Embedded Designated Support—Translated Test Directions (including ASL) | 0 | N/A | 0 | N/A | 1 | 4% |
| **Total Students Tested:** | **1** | **N/A** | **0** | **N/A** | **5** | **N/A** |

#### Delivery of Embedded and Non-Embedded Resources to Students

Universal tools, designated supports, and accommodations can be delivered as either embedded or non-embedded resources. Embedded resources are digitally delivered features or settings available as part of the technology platform for Initial Alternate ELPAC testing. Examples of embedded resources include the expandable items, color contrast, and masking.

Non-embedded resources are available, when provided by the LEA, for both computer-based assessments and paper–pencil tests. These resources are not part of the technology platform for the computer-administered Initial Alternate ELPAC. Examples of non-embedded resources include magnification, noise buffers, and the use of a scribe.

Refer to subsection [*5.5.1 Accessibility Resource Categories*](#_Accessibility_Resource_Categories_2) for a detailed description of the accessibility resources available to students taking the Initial Alternate ELPAC.

#### Usage of Designated Supports and Accommodations

LEA ELPAC coordinators and site ELPAC coordinators were responsible for assigning their students’ test settings in TOMS before testing occurred and providing the necessary resources during testing. If a test setting was not applied before testing a student, a STAIRS incident could be submitted to reset the assessment so the student could be retested with the correct accommodation or designated support. If a test setting was accidentally assigned to a student, then a STAIRS incident could also be submitted to reset the assessment so the student could be retested without the accommodation or designated support.

After LEAs and schools assigned eligible students to the appropriate accommodations or designated supports, Cambium Assessment, Inc.’s (CAI’s) TDS provided and captured whether a certain designated support (or multiple designated supports) was used by a student as the student progressed through the assessment. However, because there are no embedded accommodations for the Initial Alternate ELPAC, there was no usage of accommodations to capture or report.

Two types of designated supports are captured by the TDS for the Initial Alternate ELPAC:

1. **Masking:** This resource involves blocking off content that is not of immediate need or that may be distracting to the student.
2. **Print-on-Demand:** Paper copies of passages and stimuli, items, or all of these are printed for students.

No students taking the Initial Alternate ELPAC in the 2022–23 test administration used any of the designated supports mentioned in the previous numbered list.

### Practice and Training Tests

Practice and training tests are available publicly to LEA staff, students, parent/guardians, and any other individual for the Initial Alternate ELPAC. These tests simulate the experience of the computer-based Initial Alternate ELPAC to allow anyone to experience the assessment.

Students can access practice and training tests using a web browser. They allow students and administrators to familiarize themselves with the user interface and components of the TDS and help maintain the standardization of test administration. Practice and training tests are available through the Practice and Training Test website linked on the Online Practice and Training Tests Portal web page on the ELPAC website.

The practice tests, offered at each grade level or grade span, were released to prepare students for the Initial Alternate ELPAC. These tests more closely simulate the Initial Alternate ELPAC’s length and complexity and align with the Alternate ELPAC blueprint.

*DFAs* for the practice and training tests were available on the ELPAC website for LEA staff and parents/guardians to use to help students prepare to take the Initial Alternate ELPAC. Practice test scoring guides were also provided to help LEAs and parents/guardians understand how the items are scored.

### Test Security and Confidentiality

For the Initial Alternate ELPAC, every person who worked with the assessments, communicated test results, or received testing information was responsible for maintaining the security and confidentiality of the assessments, including CDE staff, ETS’ staff, ETS’ subcontractors, LEA assessment coordinators, school assessment coordinators, students, parents/guardians, teachers, and cooperative educational service agency staff. ETS’ Code of Ethics required that all test information, including tangible materials (e.g., test items), confidential files (e.g., those containing personally identifiable student information), and processes related to test administration (e.g., the configurations of secure servers), were kept secure. ETS had systems in place that maintained tight security for test items and test results, as well as for student data. To ensure security for all assessments that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI), which is described in the next subsection.

All assessments 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* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), “The documentation should explain the steps necessary to protect test materials and to prevent inappropriate exchange of information during the test administration session” (p. 128).

This section of the *Initial Alternate ELPAC Technical Report* describes the measures intended to prevent potential test security incidents prior to testing and the actions that were taken to handle security incidents occurring during or after the testing window using the STAIRS process.

#### 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 the *ETS Standards for Quality and Fairness* (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 help the public and auditors evaluate those products and services.

The OTI’s mission is to

* prevent test security violations;
* 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 equitable way that reflects the laws and professional standards governing the integrity of testing. The OTI also implements policies designed to detect and block technologies used to gain an unfair advantage.

In its pursuit of enforcing secure testing practices, the OTI strives to safeguard the various processes involved in an assessment development and administration cycle. For the Initial Alternate ELPAC, those processes included the following:

* Assessment development
* Item and data review
* Item banking
* Transfer of forms and items to the CDE and CAI
* Security of electronic files using a firewall
* Test administration
* Test delivery
* Processing and scoring
* Data management
* Statistical analysis
* Student confidentiality

#### Procedures to Maintain Standardization of Test Security

Test security requires the accounting of all secure materials—including computer-based summative test items and student data—before, during, and after each test administration. The LEA ELPAC coordinator is responsible for keeping all electronic test materials secure, keeping student information confidential, and making sure the site ELPAC coordinators and test examiners are properly trained regarding security policies and procedures.

The site ELPAC coordinator is responsible for mitigating test security incidents at the test site and for reporting incidents to the LEA ELPAC coordinator.

The test examiner is responsible for reporting testing incidents to the site ELPAC coordinator and securely destroying printed and digital media for items and passages generated by the print-on-demand feature of the TDS (CDE, 2023f).

The following measures ensured the security of the ELPAC:

* LEA ELPAC coordinators and site ELPAC coordinators must have electronically signed and submitted an “ELPAC Test Security Agreement for LEA ELPAC coordinators and site ELPAC coordinators” form in TOMS before ETS can grant the coordinators access to TOMS (5 *CCR* Section 11518.50[d]).
* Anyone having access to the testing materials must have electronically signed and submitted a “Test Security Affidavit for Test Examiners, Test Administrators, Proctors, Translators, Scribes, and Any Other Person Having Access to ELPAC Tests” form in TOMS before receiving access to any testing materials (5 *CCR* Section 11518.50[d]).
* Anyone having access to the testing materials but not having access to TOMS must have signed the *ELPAC* *Test Security Affidavit for Non-TOMS Users*, which was available as a web-based form, before receiving access to any testing materials.

In addition, it was the responsibility of every participant in the ELPAC System to report immediately any violation or suspected violation of test security or confidentiality. The test examiner reported to the site ELPAC coordinator or LEA ELPAC coordinator, who then submitted the incident using the STAIRS/Appeals process. Breach incidents were to be reported by the LEA ELPAC coordinator to the California Technical Assistance Center (CalTAC) and entered into STAIRS within 24 hours of the incident (5 *CCR* Section 11518.40[b][13]).

#### Test Security Monitoring

The LEA and school testing staff were responsible for maintaining the security and confidentiality of testing materials and devices during the testing window and reporting any irregularities or breaches that occurred. ETS performed site visits and testing procedure audits at randomly selected LEAs and test sites throughout California during the test administration of CAASPP and the ELPAC operational assessments. Audits were performed before, during, and after test administrations to observe adherence to published procedures regarding the handling of testing materials and test administration guidelines.

To provide this service, ETS used its OTI and subcontractor staff as auditors. All auditors had a minimum of a high school diploma, a valid driver’s license, and experience in security auditing or a related field. All had passed a background check conducted by the subcontracted vendor as part of the employment process.

ETS provided a final summary report of audit findings to the CDE at the end of the test administration. In addition, the OTI reported findings and recommendations to ETS’ program management on a weekly basis as audits were completed. ETS’ program management reported a summary of these findings to the CDE after a site visit. The OTI also provided individual audit reports directly to the LEA at the completion of the testing year.

#### 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 Sacramento, California.

All electronic applications that are included in TOMS remain protected by the ETS firewall software at all times. Because of 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.

Refer to section [*1.8 Systems Overview and Functionality*](#_Systems_Overview_and_3) in[*Chapter 1: Introduction*](#_Chapter_1:_Introduction) for more information on TOMS.

#### Transfer of Scores via Secure Data Exchange

Because of the confidential nature of test results, ETS 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 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 2022–23 Initial Alternate ELPAC, ETS entered information about the deliverable into a web form on a SharePoint website when a file was posted. A CDE staff member monitored this log throughout the day for updates to the status of deliverables and downloaded and deleted the file from the SFTP server when its status showed that it had been posted.

#### Data Management in the Secure Database

ETS 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 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. Industry-standard secure protocols are used to transfer test content and student data from the ETS internal data center to any external systems. 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 have different levels of access appropriate to the role assigned to them (CDE, 2023d).

#### Statistical Analysis on Secure Servers

During ELPAC testing, ETS’ information technology staff members retrieve data files from CAI and load those files into a database. The ETS Data Quality Services staff extract the data from the database and perform quality-control procedures (e.g., the values of all variables are as expected) before passing files to the ETS statistical analysis group. The statistical analysis staff store the files on secure servers. All staff members involved with the data adhere to the ETS Code of Ethics and the ETS Information Protection Policies to prevent any unauthorized access to data.

#### 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 during the school year. ETS takes every precaution to prevent any of this information from becoming public or being used for anything other than for testing and score-reporting purposes. These procedures are applied to all documents in which student demographic data appears, such as technical reports.

#### Student Test Results

##### Types of Results

The following deliverables are produced for reporting of the Initial Alternate ELPAC:

* Individual SSRs (electronic)
* Internet reports—available on the CDE Test Results for California’s Assessments website—aggregated by state, county, LEA, or test site

##### Security of Results Files

ETS takes measures to protect files and reports that show students’ scores and reporting 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. Staff access to production databases is limited to personnel with a business need to access the data. User IDs for production systems must be 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 both the ISO 27001 standard for information security and the ISO 22301 standard for business continuity, and conducts disaster recovery exercises annually.

Access to the ETS Computer Processing Center is controlled by employee and visitor identification badges. The Center is secured by doors that can be unlocked only 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.

##### Security of Individual Results

ETS protects individual students’ results during the following conditions:

* Scoring
* Transfer of scores by means of secure data exchange
* Reporting
* Posting of aggregated 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 take an assessment developed by ETS. 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.

#### Security and Test Administration Incident Reporting System Process

Test security incidents, such as improprieties, irregularities, and breaches, are prohibited behaviors that give a student an unfair advantage or compromise the secure administration of the assessments, which, in turn, compromise the reliability and validity of test results (CDE, 2023b). Whether intentional or unintentional, failure by staff or students to comply with security rules constitutes a test security incident. Test security incidents impact scoring and affect students’ performance on the assessment.

LEA ELPAC coordinators and site ELPAC coordinators ensured that all test security and test administration incidents were documented by following the prompts in TOMS that guided coordinators in their submittal. An Appeal is a request to reset, restore, reopen, invalidate, or grant a grace period extension to a student’s assessment. If an Appeal to a student’s assessment was warranted, TOMS provided additional prompts to file the Appeal.

After a case was submitted, an email containing a case number and next steps was sent to the submitter (and to the LEA ELPAC coordinator, if the case was submitted by the site ELPAC coordinator). The STAIRScase in TOMS provided the LEA ELPAC coordinator, the CDE, and the LEA Outreach Administrator with the opportunity to interact and communicate regarding the STAIRS process (CDE, 2023b).

Prior to the assessment administration, ETS and the CDE agreed that the following types of STAIRS cases would also be forwarded to the CDE:

* Student cheating or accessing unauthorized devices
* Security breach (where a student exposed secure materials)
* Student unable to review previous answers (i.e., 20-minute pause rule)
* Student disruption (student left the test room without authorization or disrupted the test session)

Appeals requests were reviewed by the CDE or an ETS LEA Outreach Administrator. When a request to submit an Appeal was approved, the coordinator received a system-generated email with the Appeal type that was approved (CDE, 2023b).

Types of Appeals available during the 2022–23 ELPAC administration are described in table 5.3.

Table 5.3 Types of Appeals

|  |  |
| --- | --- |
| Type of Appeal | Description |
| Reset | Resetting a student’s assessment removed that assessment from the system and enabled the student to start a new assessment from the beginning. |
| Re-open | Reopening an assessment allowed a student to access an assessment that had already been submitted or had expired. |
| Restore | Restoring an assessment returned an assessment from the Reset status to its prior status. This action could be performed only on assessments that were reset previously. |
| Grace Period Extension | Permitting a grace period extension allowed the student to review previously answered items upon logging back on to the assessment after expiration of the pause rule.  A grace period extension was granted only in cases where there was a disruption to a test session, such as a technical difficulty, fire drill, schoolwide power outage, earthquake, or other act beyond the control of the test examiner. |

##### Impropriety

A testing impropriety is an unusual circumstance that has a low impact on the individual or group of students who are testing and has a low risk of potentially affecting student performance on the assessment, test security, or test validity. An example of an impropriety could be if students were making distracting gestures or sounds or talking during the test session that creates a disruption in the test session for other students, or a student left the test room without authorization.

An impropriety can be corrected and contained at a local level. An impropriety should be reported to the LEA ELPAC coordinator and site ELPAC coordinator immediately. The coordinator must report the incident within 24 hours, using the STAIRS/Appeals process in TOMS.

##### Irregularity

A testing irregularity is an unusual circumstance that impacts an individual or a group of students who are testing and may potentially affect student performance on the assessment or impact test security or test validity. An example of an irregularity could be that students were assigned an incorrect designated support or accommodation, or students cheated or provided answers to each other.

These circumstances can be corrected and contained at the local level and submitted using the STAIRS/Appeals process in TOMS. An irregularity must be reported to the LEA ELPAC coordinator and site ELPAC coordinator immediately. The coordinator must report the irregularity within 24 hours, using the online STAIRS/Appeals process in TOMS.

##### Breach

A testing breach is an event that poses a threat to the validity of the assessment. Examples may include such situations as a release of secure materials or a security or system risk. These circumstances have external implications for the CDE and may result in a decision to remove the test item(s) from the available secure item bank.

Breaches require immediate attention; a breach that was due to social media exposure on the part of a student or adult or due to media coverage of an administration was to be escalated to CalTAC via a telephone call from the LEA ELPAC coordinator. Following the call, the site ELPAC coordinator or LEA ELPAC coordinator must report the incident using the online STAIRS/Appeals process in TOMS within 24 hours. All other breaches were to be entered into STAIRS directly.

#### Appeals

For test security incidents reported in STAIRS that resulted in a need to reset, reopen, or restore individual computer-based student assessments, the request had to be approved by the CDE. Requests to reset and reopen assessments were processed by an LEA Outreach Administrator.

In most instances, an Appeal was submitted to address a test security breach or irregularity. The LEA ELPAC coordinator or site ELPAC coordinator submitted Appeals in TOMS. All submitted Appeals were available for retrieval and reviewed by LEA and site coordinators within a given organization. However, the view of Appeals was restricted according to the user role as established in TOMS. An Appeal could be requested only by the LEA ELPAC coordinator or site ELPAC coordinator if prompted while filing a STAIRS case in TOMS (CDE, 2023d). Types of Appeals available during the 2022–23 ELPAC administration are described in table 5.3.

The total number of incidents reported in STAIRS for the Initial Alternate ELPAC was 66. The number of STAIRS incidents that required an Appeal was 41. The most common Appeal type was Re-open, and the second most common was Reset. These counts exclude incidents that were in draft form, pending, or partially approved. As noted in table 5.4, some Statewide Student Identifiers (SSIDs) were submitted with multiple Appeal types, and some Appeal types were submitted with multiple SSIDs; therefore, the numbers within table 5.4 will not add up to the numbers reported within this paragraph.

Table 5.4 provides the list of incident or issue types, the Appeal type associated with it, the number of incidents reported for that issue, and number of SSIDs affected. The incidents involving exposing secure materials or security breaches ranged from students and parents/guardians taking pictures of the testing device or test materials; to test examiners accidentally sharing the *DFAs* with parents/guardians; to test materials becoming lost at the school site because they were not kept in a secure, locked room. Counts exclude incidents that were in draft form, pending, or partially approved.

Table 5.4 Number and Types of Incidents Submitted in STAIRS

|  |  |  |  |
| --- | --- | --- | --- |
| Description | Appeal Type | Number of Incidents | Total Number of SSIDs Submitted |
| Accessibility Issue | Reset | 1 | 1 |
| Accidentally Submitted Test | Re-open | 23 | 23 |
| Administered Incorrect Assessment | Reset or No Appeal | 11 | 11 |
| Administration Error | No Appeal | 1 | 1 |
| Data Entry Issue | Reset or Re-open | 0 | 0 |
| Exposing Secure Material | No Appeal | 0 | 0 |
| Incorrect SSID Used | Reset or No Appeal | 0 | 0 |
| Irregularity Flag Submitted in Error | No Appeal | 0 | 0 |
| Other Issues | No Appeal | 0 | 0 |
| Restore from Reset | Restore | 3 | 3 |
| Student Cheating or Accessing Unauthorized Devices | No Appeal | 0 | 0 |
| Student Disruption | No Appeal | 0 | 0 |
| Validity Issue | No Appeal | 0 | 0 |

Table 5.5 provides the counts of approved Appeals.

Table 5.5 Number of Appeals Approved in STAIRS in the 2022–23 Administration—All Grade Levels and Grade Spans

|  |  |
| --- | --- |
| Appeal Type | Total Number of Appeals |
| Reset | 11 |
| Re-open | 23 |
| Restore | 3 |
| No Appeal | 1 |

### 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 Code of Regulations,* Title 5, Education, Division 1, Chapter 11, Article 2.

California Department of Education. (2021). *Alternate ELPAC accessibility and accommodations guidelines.* Sacramento, CA: California Department of Education.

California Department of Education. (2022). *California assessment accessibility resources matrix*. Sacramento, CA: California Department of Education.

California Department of Education. (2023a). *CAASPP and ELPAC accessibility guide*. Sacramento, CA: California Department of Education.

California Department of Education. (2023b). *CAASPP and ELPAC security incidents and Appeals procedure guide.* Sacramento, CA: California Department of Education.

California Department of Education. (2023c). *CAASPP and ELPAC technical specifications and configuration guide for online testing*. Sacramento, CA: California Department of Education.

California Department of Education. (2023d). *CAASPP and ELPAC Test Operations Management System user guide*. Sacramento, CA: California Department of Education.

California Department of Education. (2023e). *Initial Alternate ELPAC preparing for administration*. Sacramento, CA: California Department of Education.

California Department of Education. (2023f). *Initial Alternate ELPAC test administration manual.* Sacramento, CA: California Department of Education.

Educational Testing Service. (2014). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.

## Standard Setting

### Description

Standard setting, which also is referred to as performance level setting, refers to a class of methodologies by which one or more thresholds are used to determine performance levels. The California Department of Education (CDE) set three performance levels—*Level 1—Novice, Level 2—Intermediate,* and *Level 3—Initial Fluent English Proficient*—with two threshold cuts for each grade level.

The CDE and ETS implemented an extensive performance level–setting process involving software development, item mapping, review panels, committees, workshops, and extensive validity research to set the final thresholds and performance level descriptors. For detailed information regarding this process, refer to the *Standard Setting Technical Report for the Alternate English Language Proficiency Assessments for California* (CDE, 2022).

### Reference

California Department of Education. (2022). *Standard setting technical report for the Alternate English Language Proficiency Assessments for California.* Sacramento, CA: California Department of Education.

## Scoring and Reporting

To determine individual students’ scores for the Initial Alternate English Language Proficiency Assessments for California (ELPAC), student item responses were scored, and individual student scores were calculated on the basis of the item responses. In addition, student test scores were aggregated to produce information for schools and local educational agencies (LEAs).

This chapter describes how various types of student responses were scored, as well as the various types of scores and score reports that were produced at the end of administration of the Initial Alternate ELPAC.

### Student Test Scores

Overall scale scores were reported for the Initial Alternate ELPAC at the individual student level. To obtain scale scores, the ability (theta) scores first need to be estimated.

Prior to the test administration, ETS’ Assessment & Learning Technology Development staff reviewed each item and verified the answer keys. The keys were provided to Cambium Assessment, Inc. (CAI) for implementation in the test delivery system (TDS). After CAI finished machine-scoring item responses, scores and responses were delivered to ETS. The ETS Enterprise Score Key Management (eSKM) system calculated and collected individual students’ overall raw scores (i.e., total raw scores). ETS’ Psychometric Analysis & Research (PAR) team linked the individual item scores of the test samples to the base scale and produced the raw-to-scale-score conversion tables. When the conversion tables were implemented, eSKM produced the scale scores and performance levels using the score reporting ranges for students who completed the assessment.

ETS used two parallel scoring systems to produce and verify students’ scores. The eSKM scoring system received individual students’ item scores and item responses from CAI and computed individual student scores for the ETS reporting system. The ETS PAR team also computed individual student scores based on the same data files. The two sets of scores were then compared for the purpose of internal quality control. Any inconsistencies found in the total raw scores were investigated and resolved. The parallel scoring process ensured the quality and accuracy of scoring and supported the transfer of scores into the database of the student records scoring system, the Test Operations Management System (TOMS).

#### Raw Scores

Raw scores were obtained by summing the number of multiple-choice, single-select items answered correctly and the number of a rubric-based item score points obtained. The number and percentage of students at each raw score point are reported in table 7.A.1 through table 7.A.6 in [appendix 7.A](#_Appendix_6.A:_Raw).

#### Theta Scores

All items presented to students were calibrated onto the theta scale so that students’ raw scores could be transformed into ability (theta) estimates by using the item response theory inverse test characteristic curve (TCC) method (Stocking, 1996). With this method, each student’s estimated ability is the ability value at which the student’s expected raw score is equal to the student’s obtained raw score. Refer to subsection [*8.3.3 Scaling the Scores*](#_Toc120784038) for more information about the inverse TCC method and transforming theta scores to scale scores. Each grade level and grade span has its own theta scale.

Once a conversion table from the raw score to theta score is created for each grade level or grade span, the estimated ability (theta) score of each individual student can be obtained from the conversion table. The theta score can later be transformed into a scale score through a linear transformation. Refer to subsection [*8.3.3.2 Transformation from Theta Scores to Scale Scores*](#_Transformation_from_Theta_3) for more information.

#### Scale Scores for the Total Assessment

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. Instead, student performance on the Initial Alternate ELPAC is reported in terms of scale scores that express student proficiency in terms of a constant metric. Thus, a scale score of 150 for grade one in one administration represents the same level of proficiency as 150 for grade one in another administration even though each scale score may represent a different raw score.

The following requirements were used to develop and define the Initial Alternate ELPAC reporting scale ranges, as described in subsection [*8.3.3 Scaling the Scores*](#_Toc120784038)*:*

1. Each scale score has three digits, where the first digit is always 1 (i.e., possible scores range from 101 to 199) and the second and third digits represent the scale score as derived from the transformation from the raw scores to the scale scores as described in the previous subsection.
2. Each threshold score on the scale is the same from year to year. Also, across the grade levels and grade spans, the last two digits corresponding to the Level 2—Intermediate and Level 3—Initial Fluent English Proficient threshold scores are the same (refer to subsection [*7.1.4 Performance Levels*](#_Performance_Levels)for a brief description of performance levels).

Scale score frequency distributions by grade levels and grade spans are presented in [appendix 7.A](#_Appendix_6.A:_Raw), table 7.A.1 through table 7.A.6.

#### Performance Levels

Initial Alternate ELPAC reporting scales classify each student’s performance into one of the three performance levels, with Level 1—Novice indicating the lowest level of performance and Level 3—Initial Fluent English Proficient indicating the highest level of performance. Student test results are reported in the following overall performance levels:

* **Level 1—Novice:** The student is beginning to develop the English skills needed to communicate and learn in school.
* **Level 2—Intermediate:** The student can sometimes use English to communicate and learn in school.
* **Level 3—Initial Fluent English Proficient:** The student has sufficient English skills to communicate and learn in school.

Detailed information regarding the determination of the performance levels can be found in the *Standard Setting Technical Report for the Alternate ELPAC* (California Department of Education [CDE], 2022).

Scale score ranges for each reporting level are the following:

* **Level 1—**101–143
* **Level 2—**144–159
* **Level 3—**160–199

### Overview of Score Aggregation Procedures

To provide meaningful results to the interest holders, test scores for a given grade level are aggregated at the school, LEA or direct funded charter school, county, and state levels. The aggregated scores are generated both for selected groups and for the population. The next subsection contains a description of the types of aggregation performed on ELPAC computer-based assessment scores. Score aggregation includes only students with valid scores; refer to subsection [*7.3.2 Special Cases*](#_Special_Cases_1) for more information.

#### Student Score Distributions and Summary Statistics

Table 7.1 presents the means and standard deviations (SDs) of overall scale scores by grade level.

Table 7.1 Mean and SD of Scale Scores by Grade Level

|  |  |  |  |
| --- | --- | --- | --- |
| Grade Level | Number of Students Tested | Overall Scale Score Mean | Overall Scale Score SD |
| Kindergarten | 683 | 138 | 20 |
| 1 | 59 | 135 | 22 |
| 2 | 21 | 136 | 18 |
| 3 | 13 | 130 | 17 |
| 4 | 14 | 124 | 20 |
| 5 | 15 | 129 | 25 |
| 6 | 11 | 127 | 24 |
| 7 | 4 | 149 | 35 |
| 8 | 8 | 141 | 9 |
| 9 | 7 | 138 | 17 |
| 10 | 3 | 126 | 24 |
| 11 | 3 | 126 | 21 |
| 12 | 15 | 139 | 17 |

The percentage of students in each performance level for the overall assessment is presented in table 7.2. Note that numbers might not add to 100 because of rounding.

Table 7.2 Percentage of Students in Each Performance Level by Grade Level

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Level 1 N | Level 1 % | Level 2 N | Level 2 % | Level 3 N | Level 3 % |
| Kindergarten | 365 | 53 | 249 | 36 | 69 | 10 |
| 1 | 33 | 56 | 18 | 31 | 8 | 14 |
| 2 | 13 | 62 | 7 | 33 | 1 | 5 |
| 3 | 12 | 92 | 1 | 8 | 0 | 0 |
| 4 | 11 | 79 | 3 | 21 | 0 | 0 |
| 5 | 10 | 67 | 4 | 27 | 1 | 7 |
| 6 | 6 | 55 | 5 | 45 | 0 | 0 |
| 7 | 1 | 25 | 1 | 25 | 2 | 50 |
| 8 | 4 | 50 | 4 | 50 | 0 | 0 |
| 9 | 5 | 71 | 2 | 29 | 0 | 0 |
| 10 | 2 | 67 | 1 | 33 | 0 | 0 |
| 11 | 3 | 100 | 0 | 0 | 0 | 0 |
| 12 | 7 | 47 | 7 | 47 | 1 | 7 |

Figure 7.1, which is derived from the data in table 7.2, presents the percentage of students at each performance level by grade level. Because of the very small sample sizes at most grade levels, the percentage of students at Level 3 has a very large range, from 0 percent to 50 percent. In the two grade levels with the most test takers, kindergarten and grade one, the percentage of students in Level 3 is between 10 percent and 14 percent.

Figure 7.1 Percentage of students at each performance level

#### Demographic Student Group Summaries

Table 7.B.1 through table 7.B.6 in [appendix 7.B](#_Appendix_7.B:_Means_1) provide, for all grade levels and grade spans, the number and the percentage of students for selected student groups with completed test scores for the 2022–23 administration of the Initial Alternate ELPAC.

The tables in [appendix 7.B](#_Appendix_7.B:_Demographic) group students by demographic characteristics, including economic status, length of enrollment in US schools, ethnicity, gender, migrant status, primary disability type, military status, homeless status, and foster youth status. The list of student groups is presented in table 7.3.

Table 7.3 Demographic Student Groups to Be Reported

|  |  |
| --- | --- |
| Category | Student Groups |
| **Economic Status** | * Not economically disadvantaged * Economically disadvantaged |
| **Enrollment in US Schools** | * Less than 12 months * 12 months or more * Duration unknown |
| **Ethnicity** | * American Indian or Alaska Native * Asian * Native Hawaiian or Other Pacific Islander * Filipino * Hispanic or Latino * Black or African American * White * Two or more races |
| **Gender** | * Male * Female * Nonbinary |
| **Migrant Status** | * Migrant education * Not migrant education |
| **Primary Disability Type** | * Intellectual disability * Hearing impairment * Speech or language impairment * Visual impairment * Emotional impairment * Orthopedic impairment * Other health impairment * Specific learning disability * Deaf-blindness * Multiple disabilities * Autism * Traumatic brain injury * Not classified |
| **Military Status** | * Armed forces family member * Not armed forces family member |
| **Homeless Status** | * Homeless * Not homeless |
| **Foster Youth Status** | * Foster youth * Not foster youth |

#### Student Group Distributions

Table 7.4 shows means and SDs of scale scores for the overall assessment by student group in kindergarten. Means and SDs are not reported for the other grade levels or grade spans because of insufficient sample sizes. Even for student groups with 11 or more students in other grade levels or grade spans, the small sample sizes would make it difficult to meaningfully compare the student groups.

Table 7.4 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Kindergarten

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 683 | 138 | 20 | 53 | 36 | 10 |
| Male | 488 | 138 | 21 | 53 | 36 | 11 |
| Female | 195 | 139 | 19 | 54 | 38 | 8 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 0 | N/A | N/A | N/A | N/A | N/A |
| Asian | 101 | 138 | 20 | 55 | 36 | 9 |
| Native Hawaiian or Other Pacific Islander | 2 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 8 | N/A | N/A | N/A | N/A | N/A |
| Hispanic or Latino | 521 | 138 | 20 | 54 | 36 | 9 |
| Black or African American | 8 | N/A | N/A | N/A | N/A | N/A |
| White | 27 | 147 | 23 | 41 | 30 | 30 |
| Two or more races | 16 | 151 | 20 | 31 | 50 | 19 |
| Intellectual disability | 138 | 136 | 19 | 57 | 38 | 6 |
| Hearing impairment | 2 | N/A | N/A | N/A | N/A | N/A |
| Speech or language impairment | 28 | 146 | 18 | 43 | 36 | 21 |
| Visual impairment | 0 | N/A | N/A | N/A | N/A | N/A |
| Emotional impairment | 0 | N/A | N/A | N/A | N/A | N/A |
| Orthopedic impairment | 11 | 136 | 36 | 73 | 9 | 18 |
| Other health impairment | 47 | 142 | 20 | 43 | 45 | 13 |
| Specific learning disability | 0 | N/A | N/A | N/A | N/A | N/A |
| Deaf-blindness | 0 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 33 | 130 | 18 | 76 | 24 | 0 |
| Autism | 423 | 139 | 20 | 52 | 37 | 11 |
| Traumatic brain injury | 1 | N/A | N/A | N/A | N/A | N/A |
| Not classified | 0 | N/A | N/A | N/A | N/A | N/A |
| Not economically disadvantaged | 233 | 139 | 20 | 54 | 35 | 11 |
| Economically disadvantaged | 450 | 138 | 20 | 53 | 37 | 10 |
| In US schools less than 12 months | 609 | 138 | 20 | 54 | 37 | 8 |
| In US schools 12 months or more | 23 | 141 | 18 | 43 | 48 | 9 |
| Duration unknown | 51 | 143 | 23 | 47 | 22 | 31 |
| Migrant education | 9 | N/A | N/A | N/A | N/A | N/A |
| Not migrant education | 674 | 138 | 20 | 53 | 36 | 10 |
| Armed forces family member | 8 | N/A | N/A | N/A | N/A | N/A |
| Not armed forces family member | 675 | 138 | 20 | 53 | 36 | 10 |
| Homeless | 24 | 138 | 19 | 63 | 29 | 8 |
| Not homeless | 659 | 139 | 20 | 53 | 37 | 10 |
| Foster youth | 6 | N/A | N/A | N/A | N/A | N/A |
| Not foster youth | 677 | 138 | 20 | 53 | 37 | 10 |

**Note:** To protect privacy, when the number of students in a student group is 10 or fewer, the summary statistics of scale scores and performance levels are not reported, but are replaced by “N/A.” Because of very low samples sizes, student group results are presented for kindergarten only.

For kindergarten students, female students scored, on average, 1 scale score point higher than male students. Hispanic or Latino and Asian students scored, on average, the same as the overall group of test takers; White students scored 9 scale score points higher; and students reporting two or more races scored 13 scale score points higher. Students with intellectual disabilities or orthopedic impairments scored, on average, 2 scale score points lower than the overall group of test takers; students with multiple abilities scored 8 scale score points lower; students with autism scored 1 scale score point higher; students with other health impairments scored 4 scale score points higher; and students with speech or language impairments scored 8 scale score points higher.

### Reports Produced and Scores for Each Report

The assessments that make up the ELPAC computer-based assessments provide results or score summaries that are reported for different purposes. The four major purposes are to

1. help facilitate conversations between parents/guardians and teachers about student performance,
2. serve as a tool to help parents/guardians and teachers work together to improve student learning,
3. help schools and LEAs identify strengths and areas that need improvement in their educational programs, and
4. provide the public and policymakers with information about student performance.

This section provides detailed descriptions of the uses and applications of ELPAC reporting for students.

#### Online Reporting

TOMS is a secure website hosted by ETS that permits LEA users to manage the ELPAC computer-based assessments and to inform the TDS. This system uses a role-specific design to restrict access to certain tools and applications based on the user’s designated role. Specific functions of TOMS include the following:

* Manage user access privileges
* Manage test administration calendars and testing windows
* Manage student test assignments
* Manage and confirm the accuracy of students’ test settings (i.e., designated supports and accommodations) prior to testing
* Generate and download various reports

#### Special Cases

Student scores were not reported for the following cases:

* The student did not log on to test systems.
* The student partially tested and did not meet participation requirements.

These students were excluded from aggregated results as well.

#### Types of Score Reports

There are two categories of ELPAC reports. The specific reports within each category are presented in this subsection.

1. **Student Score Report (SSR)—**The SSR was the official score report for parents/‌guardians. An SSR described the student’s results and was made available only to students who met the program’s participation requirement.
2. **LEA student data files and aggregations—**LEA student data files were available for download on demand by the LEA in TOMS to coincide with availability of the SSRs. Aggregated data was available to view in the Test Results for California’s Assessments website.

##### Student Score Report

The Initial Alternate ELPAC SSR is the official score report for parents/guardians and includes the following information:

* Overall scale score
* Overall performance level

As mentioned previously, overall scale scores placed students into one of the three Initial Alternate ELPAC performance levels: Initial Fluent English Proficient, Intermediate, and Novice.

LEAs had four options for accessing and distributing SSRs to parents/guardians:

1. Accessing electronic SSR PDFs using a locally provided parent/guardian or student portal
2. Downloading SSR PDFs from TOMS and making them available electronically using a secure local method
3. Downloading SSR PDFs from TOMS, printing them, and making them available locally
4. Purchasing paper SSRs from ETS

The LEA ELPAC coordinator could forward the appropriate reports to test sites. In the case of a locally printed Initial Alternate ELPAC SSR, the LEA sent the printed report(s) to the child’s parent/guardian. Initial Alternate ELPAC SSRs that included individual student results were not distributed beyond the student’s school.

Scores for students who were assigned accommodations or designated supports are reported in the same way as for students who were not assigned accommodations or designated supports. Detailed information about accessibility resources is described in subsection [*5.5.1 Accessibility Resource Categories*](#_Accessibility_Resource_Categories_2).

For the 2022–23 test administration, SSRs were made available to the LEAs in English, Spanish, Filipino, Chinese (Traditional), Vietnamese, and Korean. An SSR in a supported language was created if the student’s primary language as reported in the California Longitudinal Pupil 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 the supported language. SSRs were made available only to students who met the participation requirement by responding to at least one expressive and one receptive item. These reports were available as PDFs for the LEA to download from TOMS.

Further information about the SSR and its interpretation is provided on the ELPAC Starting Smarter website for California assessments.

###### Access via Student or Parent Portal

LEAs had the option to provide SSRs electronically using a locally provided parent or student portal.

Amazon Web Services—with the Amazon Simple Storage Service and the Amazon Key Management Service—ensured encrypted access for parents/guardians to view a child’s electronic SSR, which was available as a PDF.

###### Access via the Test Operations Management System

The LEA ELPAC coordinator downloaded the electronic PDFs directly from TOMS and could forward the appropriate reports to test sites. Optionally, the LEA could download and then print the SSR PDF and then send the printed report(s) to the child’s parent/‌guardian.

##### Local Educational Agency Student Data Files and Aggregations

The ELPAC student data files for the LEA were available for the LEA ELPAC coordinator and site ELPAC coordinator to download from TOMS.

Current and historical aggregated results are accessible to the public on the CDE Test Results for California’s Assessments website.

#### Score Report Applications

Initial Alternate ELPAC results, presented in SSRs, provided parents/guardians and LEAs with information about a child’s English proficiency as the child entered a California public school for the first time. Identifying students with the most significant cognitive disabilities who need help in English is important because it helps students obtain the extra support and resources they need to do well in school.

#### Criteria for Interpreting Test Scores

An LEA may use ELPAC computer-based summative assessment results to help make decisions about student placement, promotion, retention, or other considerations related to student achievement. However, it is important to remember that a single assessment can provide only limited information. Other relevant information should be considered as well. It is advisable for parents/guardians to evaluate their child’s strengths and weaknesses in the relevant topics by reviewing classroom work and progress reports in addition to the child’s ELPAC computer-based summative assessment results. It is also important to note that a student’s score in an item type (expressive or receptive) could vary somewhat if the student were retested.

#### Criteria for Interpreting Score Reports

The information presented in various reports must be interpreted with caution when making performance comparisons. When comparing scale score and performance-level results, the user is limited to comparisons within a grade level or grade span. The user may compare scale scores for the same grade level or grade span, within a school, between schools, or between a school and its LEA, its county, or the state. The ELPAC user can also make comparisons within the same grade level or grade span across years.

However, comparing scale scores from different grade levels for the ELPAC is not appropriate, because the curricula are different across grade levels and the scale scores are not vertically linked between grade levels or grade spans.

For more details on the criteria for interpreting information provided on the score reports, refer to the ELPAC Starting Smarter website for California assessments.

### References

California Department of Education. (2022). *Standard setting technical report for the Alternate English Language Proficiency Assessments for California.* Sacramento, CA: California Department of Education.

Stocking, M.L. (1996). An alternative method for scoring adaptive tests. *Journal of Educational and Behavioral Statistics, 21,* 365–89.

### Appendix 7.A: Raw-Score-to-Scale-Score Distributions

Table 7.A.1 Raw-Score-to-Scale-Score Distribution for Kindergarten

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 21 | 88 | 12.88 |
| 1 | 116 | 1 | 11 | 47 | 6.88 |
| 2 | 123 | 1 | 8 | 27 | 3.95 |
| 3 | 128 | 1 | 6 | 32 | 4.69 |
| 4 | 131 | 1 | 6 | 23 | 3.37 |
| 5 | 134 | 1 | 5 | 24 | 3.51 |
| 6 | 136 | 1 | 5 | 21 | 3.07 |
| 7 | 138 | 1 | 4 | 26 | 3.81 |
| 8 | 140 | 1 | 4 | 21 | 3.07 |
| 9 | 142 | 1 | 4 | 27 | 3.95 |
| 10 | 143 | 1 | 4 | 29 | 4.25 |
| 11 | 145 | 2 | 4 | 24 | 3.51 |
| 12 | 146 | 2 | 4 | 26 | 3.81 |
| 13 | 147 | 2 | 4 | 22 | 3.22 |
| 14 | 149 | 2 | 4 | 22 | 3.22 |
| 15 | 150 | 2 | 4 | 26 | 3.81 |
| 16 | 151 | 2 | 4 | 26 | 3.81 |
| 17 | 153 | 2 | 4 | 25 | 3.66 |
| 18 | 154 | 2 | 4 | 19 | 2.78 |
| 19 | 155 | 2 | 4 | 15 | 2.20 |
| 20 | 157 | 2 | 4 | 27 | 3.95 |
| 21 | 158 | 2 | 4 | 17 | 2.49 |
| 22 | 160 | 3 | 4 | 10 | 1.46 |
| 23 | 162 | 3 | 4 | 14 | 2.05 |
| 24 | 163 | 3 | 5 | 9 | 1.32 |
| 25 | 166 | 3 | 5 | 11 | 1.61 |
| 26 | 168 | 3 | 5 | 8 | 1.17 |
| 27 | 171 | 3 | 6 | 3 | 0.44 |
| 28 | 174 | 3 | 6 | 3 | 0.44 |
| 29 | 179 | 3 | 8 | 3 | 0.44 |
| 30 | 187 | 3 | 11 | 4 | 0.59 |
| 31 | 199 | 3 | 18 | 4 | 0.59 |

Table 7.A.2 Raw-Score-to-Scale-Score Distribution for Grade One

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 18 | 10 | 16.95 |
| 1 | 112 | 1 | 11 | 4 | 6.78 |
| 2 | 120 | 1 | 8 | 3 | 5.08 |
| 3 | 125 | 1 | 7 | 2 | 3.39 |
| 4 | 128 | 1 | 6 | 4 | 6.78 |
| 5 | 131 | 1 | 5 | 2 | 3.39 |
| 6 | 134 | 1 | 5 | 2 | 3.39 |
| 7 | 136 | 1 | 5 | 2 | 3.39 |
| 8 | 138 | 1 | 5 | 1 | 1.69 |
| 9 | 140 | 1 | 4 | 0 | 0.00 |
| 10 | 141 | 1 | 4 | 1 | 1.69 |
| 11 | 143 | 1 | 4 | 2 | 3.39 |
| 12 | 144 | 2 | 4 | 4 | 6.78 |
| 13 | 146 | 2 | 4 | 1 | 1.69 |
| 14 | 147 | 2 | 4 | 0 | 0.00 |
| 15 | 149 | 2 | 4 | 0 | 0.00 |
| 16 | 150 | 2 | 4 | 3 | 5.08 |
| 17 | 152 | 2 | 4 | 2 | 3.39 |
| 18 | 153 | 2 | 4 | 2 | 3.39 |
| 19 | 154 | 2 | 4 | 1 | 1.69 |
| 20 | 156 | 2 | 4 | 2 | 3.39 |
| 21 | 158 | 2 | 4 | 1 | 1.69 |
| 22 | 159 | 2 | 4 | 2 | 3.39 |
| 23 | 161 | 3 | 5 | 2 | 3.39 |
| 24 | 163 | 3 | 5 | 3 | 5.08 |
| 25 | 165 | 3 | 5 | 1 | 1.69 |
| 26 | 168 | 3 | 6 | 2 | 3.39 |
| 27 | 171 | 3 | 6 | 0 | 0.00 |
| 28 | 176 | 3 | 8 | 0 | 0.00 |
| 29 | 184 | 3 | 11 | 0 | 0.00 |
| 30 | 199 | 3 | 22 | 0 | 0.00 |

Table 7.A.3 Raw-Score-to-Scale-Score Distribution for Grade Two

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 16 | 1 | 4.76 |
| 1 | 109 | 1 | 11 | 1 | 4.76 |
| 2 | 117 | 1 | 8 | 2 | 9.52 |
| 3 | 122 | 1 | 7 | 0 | 0.00 |
| 4 | 125 | 1 | 6 | 3 | 14.29 |
| 5 | 128 | 1 | 5 | 2 | 9.52 |
| 6 | 130 | 1 | 5 | 1 | 4.76 |
| 7 | 132 | 1 | 5 | 0 | 0.00 |
| 8 | 134 | 1 | 5 | 0 | 0.00 |
| 9 | 136 | 1 | 4 | 0 | 0.00 |
| 10 | 138 | 1 | 4 | 2 | 9.52 |
| 11 | 139 | 1 | 4 | 1 | 4.76 |
| 12 | 141 | 1 | 4 | 0 | 0.00 |
| 13 | 142 | 1 | 4 | 0 | 0.00 |
| 14 | 144 | 2 | 4 | 0 | 0.00 |
| 15 | 145 | 2 | 4 | 2 | 9.52 |
| 16 | 147 | 2 | 4 | 1 | 4.76 |
| 17 | 148 | 2 | 4 | 0 | 0.00 |
| 18 | 150 | 2 | 4 | 0 | 0.00 |
| 19 | 151 | 2 | 4 | 0 | 0.00 |
| 20 | 153 | 2 | 4 | 0 | 0.00 |
| 21 | 155 | 2 | 4 | 0 | 0.00 |
| 22 | 157 | 2 | 5 | 2 | 9.52 |
| 23 | 159 | 2 | 5 | 2 | 9.52 |
| 24 | 161 | 3 | 5 | 0 | 0.00 |
| 25 | 163 | 3 | 5 | 1 | 4.76 |
| 26 | 166 | 3 | 6 | 0 | 0.00 |
| 27 | 170 | 3 | 7 | 0 | 0.00 |
| 28 | 175 | 3 | 8 | 0 | 0.00 |
| 29 | 184 | 3 | 12 | 0 | 0.00 |
| 30 | 199 | 3 | 22 | 0 | 0.00 |

Table 7.A.4 Raw-Score-to-Scale-Score Distribution for Grade Span Three Through Five

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 21 | 13 | 30.95 |
| 1 | 116 | 1 | 9 | 3 | 7.14 |
| 2 | 123 | 1 | 7 | 1 | 2.38 |
| 3 | 127 | 1 | 6 | 1 | 2.38 |
| 4 | 130 | 1 | 5 | 0 | 0.00 |
| 5 | 132 | 1 | 5 | 1 | 2.38 |
| 6 | 134 | 1 | 4 | 0 | 0.00 |
| 7 | 136 | 1 | 4 | 3 | 7.14 |
| 8 | 138 | 1 | 4 | 3 | 7.14 |
| 9 | 139 | 1 | 4 | 2 | 4.76 |
| 10 | 141 | 1 | 3 | 2 | 4.76 |
| 11 | 142 | 1 | 3 | 3 | 7.14 |
| 12 | 143 | 1 | 3 | 1 | 2.38 |
| 13 | 144 | 2 | 3 | 1 | 2.38 |
| 14 | 145 | 2 | 3 | 0 | 0.00 |
| 15 | 146 | 2 | 3 | 2 | 4.76 |
| 16 | 148 | 2 | 3 | 1 | 2.38 |
| 17 | 149 | 2 | 3 | 0 | 0.00 |
| 18 | 150 | 2 | 3 | 1 | 2.38 |
| 19 | 151 | 2 | 3 | 0 | 0.00 |
| 20 | 152 | 2 | 3 | 1 | 2.38 |
| 21 | 154 | 2 | 4 | 1 | 2.38 |
| 22 | 155 | 2 | 4 | 1 | 2.38 |
| 23 | 156 | 2 | 4 | 0 | 0.00 |
| 24 | 158 | 2 | 4 | 0 | 0.00 |
| 25 | 160 | 3 | 4 | 0 | 0.00 |
| 26 | 162 | 3 | 5 | 0 | 0.00 |
| 27 | 165 | 3 | 5 | 0 | 0.00 |
| 28 | 169 | 3 | 7 | 1 | 2.38 |
| 29 | 176 | 3 | 9 | 0 | 0.00 |
| 30 | 199 | 3 | 32 | 0 | 0.00 |

Table 7.A.5 Raw-Score-to-Scale-Score Distribution for Grade Span Six Through Eight

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 18 | 5 | 21.74 |
| 1 | 114 | 1 | 9 | 1 | 4.35 |
| 2 | 120 | 1 | 7 | 0 | 0.00 |
| 3 | 125 | 1 | 6 | 0 | 0.00 |
| 4 | 128 | 1 | 5 | 3 | 13.04 |
| 5 | 130 | 1 | 5 | 0 | 0.00 |
| 6 | 132 | 1 | 4 | 0 | 0.00 |
| 7 | 134 | 1 | 4 | 0 | 0.00 |
| 8 | 136 | 1 | 4 | 0 | 0.00 |
| 9 | 137 | 1 | 4 | 1 | 4.35 |
| 10 | 139 | 1 | 4 | 0 | 0.00 |
| 11 | 140 | 1 | 4 | 1 | 4.35 |
| 12 | 141 | 1 | 3 | 0 | 0.00 |
| 13 | 143 | 1 | 3 | 0 | 0.00 |
| 14 | 144 | 2 | 3 | 1 | 4.35 |
| 15 | 145 | 2 | 3 | 2 | 8.70 |
| 16 | 147 | 2 | 3 | 2 | 8.70 |
| 17 | 148 | 2 | 3 | 1 | 4.35 |
| 18 | 149 | 2 | 4 | 0 | 0.00 |
| 19 | 151 | 2 | 4 | 1 | 4.35 |
| 20 | 152 | 2 | 4 | 1 | 4.35 |
| 21 | 154 | 2 | 4 | 1 | 4.35 |
| 22 | 155 | 2 | 4 | 1 | 4.35 |
| 23 | 157 | 2 | 4 | 0 | 0.00 |
| 24 | 159 | 2 | 4 | 0 | 0.00 |
| 25 | 161 | 3 | 5 | 0 | 0.00 |
| 26 | 163 | 3 | 5 | 0 | 0.00 |
| 27 | 166 | 3 | 5 | 0 | 0.00 |
| 28 | 171 | 3 | 7 | 1 | 4.35 |
| 29 | 177 | 3 | 9 | 1 | 4.35 |
| 30 | 199 | 3 | 30 | 0 | 0.00 |

Table 7.A.6 Raw-Score-to-Scale-Score Distribution for Grade Span Nine Through Twelve

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Raw Score | Scale Score | Level | CSEM | N | Percent |
| 0 | 101 | 1 | 19 | 5 | 17.86 |
| 1 | 115 | 1 | 8 | 0 | 0.00 |
| 2 | 121 | 1 | 6 | 0 | 0.00 |
| 3 | 125 | 1 | 5 | 0 | 0.00 |
| 4 | 128 | 1 | 4 | 1 | 3.57 |
| 5 | 130 | 1 | 4 | 0 | 0.00 |
| 6 | 131 | 1 | 4 | 0 | 0.00 |
| 7 | 133 | 1 | 3 | 1 | 3.57 |
| 8 | 135 | 1 | 3 | 0 | 0.00 |
| 9 | 136 | 1 | 3 | 0 | 0.00 |
| 10 | 137 | 1 | 3 | 2 | 7.14 |
| 11 | 138 | 1 | 3 | 4 | 14.29 |
| 12 | 140 | 1 | 3 | 3 | 10.71 |
| 13 | 141 | 1 | 3 | 0 | 0.00 |
| 14 | 142 | 1 | 3 | 1 | 3.57 |
| 15 | 143 | 1 | 3 | 0 | 0.00 |
| 16 | 144 | 2 | 3 | 1 | 3.57 |
| 17 | 146 | 2 | 3 | 0 | 0.00 |
| 18 | 147 | 2 | 3 | 3 | 10.71 |
| 19 | 148 | 2 | 3 | 1 | 3.57 |
| 20 | 149 | 2 | 3 | 2 | 7.14 |
| 21 | 151 | 2 | 3 | 1 | 3.57 |
| 22 | 152 | 2 | 3 | 1 | 3.57 |
| 23 | 154 | 2 | 4 | 0 | 0.00 |
| 24 | 155 | 2 | 4 | 0 | 0.00 |
| 25 | 157 | 2 | 4 | 1 | 3.57 |
| 26 | 159 | 2 | 4 | 0 | 0.00 |
| 27 | 162 | 3 | 5 | 1 | 3.57 |
| 28 | 165 | 3 | 6 | 0 | 0.00 |
| 29 | 171 | 3 | 8 | 0 | 0.00 |
| 30 | 199 | 3 | 44 | 0 | 0.00 |

### Appendix 7.B: Demographic Student Group Summaries

What follows are details about the data tables in appendix 7.B:

* The student group “All” represents all students who took an assessment.
* The *Number Tested* columns contain the number of students in each demographic student group who took the assessment.
* The *Number Analyzed* columns contain the number of students included in item analyses after data cleaning rules were applied:
* *Number Analyzed*—Students who responded to at least two items

Table 7.B.1 **Demographic Summary for Students: Kindergarten**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 691 | 689 | 99.71 | 683 | 98.84 |
| Male | 492 | 491 | 99.80 | 488 | 99.19 |
| Female | 199 | 198 | 99.50 | 195 | 97.99 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 0 | 0 | N/A | 0 | N/A |
| Asian | 102 | 102 | 100.00 | 101 | 99.02 |
| Native Hawaiian or Other Pacific Islander | 2 | 2 | 100.00 | 2 | 100.00 |
| Filipino | 8 | 8 | 100.00 | 8 | 100.00 |
| Hispanic or Latino | 527 | 525 | 99.62 | 521 | 98.86 |
| Black or African American | 8 | 8 | 100.00 | 8 | 100.00 |
| White | 27 | 27 | 100.00 | 27 | 100.00 |
| Two or more races | 17 | 17 | 100.00 | 16 | 94.12 |
| Intellectual disability | 138 | 138 | 100.00 | 138 | 100.00 |
| Hearing impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Speech or language impairment | 28 | 28 | 100.00 | 28 | 100.00 |
| Visual impairment | 0 | 0 | N/A | 0 | N/A |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 11 | 11 | 100.00 | 11 | 100.00 |
| Other health impairment | 47 | 47 | 100.00 | 47 | 100.00 |
| Specific learning disability | 0 | 0 | N/A | 0 | N/A |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 38 | 37 | 97.37 | 33 | 86.84 |
| Autism | 426 | 425 | 99.77 | 423 | 99.30 |
| Traumatic brain injury | 1 | 1 | 100.00 | 1 | 100.00 |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 235 | 233 | 99.15 | 233 | 99.15 |
| Economically disadvantaged | 456 | 456 | 100.00 | 450 | 98.68 |
| In US schools less than 12 months | 616 | 615 | 99.84 | 609 | 98.86 |
| In US schools 12 months or more | 23 | 23 | 100.00 | 23 | 100.00 |
| Duration unknown | 52 | 51 | 98.08 | 51 | 98.08 |
| Migrant education | 9 | 9 | 100.00 | 9 | 100.00 |
| Not migrant education | 682 | 680 | 99.71 | 674 | 98.83 |
| Armed forces family member | 8 | 8 | 100.00 | 8 | 100.00 |
| Not armed forces family member | 683 | 681 | 99.71 | 675 | 98.83 |
| Homeless | 25 | 25 | 100.00 | 24 | 96.00 |
| Not homeless | 666 | 664 | 99.70 | 659 | 98.95 |
| Foster youth | 6 | 6 | 100.00 | 6 | 100.00 |
| Not foster youth | 685 | 683 | 99.71 | 677 | 98.83 |

Table 7.B.2 **Demographic Summary for Students: Grade One**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 61 | 60 | 98.36 | 59 | 96.72 |
| Male | 43 | 43 | 100.00 | 42 | 97.67 |
| Female | 18 | 17 | 94.44 | 17 | 94.44 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 0 | 0 | N/A | 0 | N/A |
| Asian | 11 | 11 | 100.00 | 11 | 100.00 |
| Native Hawaiian or Other Pacific Islander | 0 | 0 | N/A | 0 | N/A |
| Filipino | 0 | 0 | N/A | 0 | N/A |
| Hispanic or Latino | 43 | 42 | 97.67 | 41 | 95.35 |
| Black or African American | 3 | 3 | 100.00 | 3 | 100.00 |
| White | 4 | 4 | 100.00 | 4 | 100.00 |
| Two or more races | 0 | 0 | N/A | 0 | N/A |
| Intellectual disability | 18 | 18 | 100.00 | 18 | 100.00 |
| Hearing impairment | 0 | 0 | N/A | 0 | N/A |
| Speech or language impairment | 1 | 1 | 100.00 | 1 | 100.00 |
| Visual impairment | 1 | 1 | 100.00 | 1 | 100.00 |
| Emotional impairment | 1 | 1 | 100.00 | 1 | 100.00 |
| Orthopedic impairment | 0 | 0 | N/A | 0 | N/A |
| Other health impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Specific learning disability | 0 | 0 | N/A | 0 | N/A |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 14 | 13 | 92.86 | 12 | 85.71 |
| Autism | 24 | 24 | 100.00 | 24 | 100.00 |
| Traumatic brain injury | 0 | 0 | N/A | 0 | N/A |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 21 | 21 | 100.00 | 21 | 100.00 |
| Economically disadvantaged | 40 | 39 | 97.50 | 38 | 95.00 |
| In US schools less than 12 months | 27 | 27 | 100.00 | 27 | 100.00 |
| In US schools 12 months or more | 24 | 24 | 100.00 | 23 | 95.83 |
| Duration unknown | 10 | 9 | 90.00 | 9 | 90.00 |
| Migrant education | 1 | 1 | 100.00 | 1 | 100.00 |
| Not migrant education | 60 | 59 | 98.33 | 58 | 96.67 |
| Armed forces family member | 1 | 1 | 100.00 | 1 | 100.00 |
| Not armed forces family member | 60 | 59 | 98.33 | 58 | 96.67 |
| Homeless | 4 | 3 | 75.00 | 3 | 75.00 |
| Not homeless | 57 | 57 | 100.00 | 56 | 98.25 |
| Foster youth | 0 | 0 | N/A | 0 | N/A |
| Not foster youth | 61 | 60 | 98.36 | 59 | 96.72 |

Table 7.B.3 **Demographic Summary for Students: Grade Two**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 21 | 21 | 100.00 | 21 | 100.00 |
| Male | 18 | 18 | 100.00 | 18 | 100.00 |
| Female | 3 | 3 | 100.00 | 3 | 100.00 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 0 | 0 | N/A | 0 | N/A |
| Asian | 1 | 1 | 100.00 | 1 | 100.00 |
| Native Hawaiian or Other Pacific Islander | 0 | 0 | N/A | 0 | N/A |
| Filipino | 0 | 0 | N/A | 0 | N/A |
| Hispanic or Latino | 14 | 14 | 100.00 | 14 | 100.00 |
| Black or African American | 0 | 0 | N/A | 0 | N/A |
| White | 4 | 4 | 100.00 | 4 | 100.00 |
| Two or more races | 2 | 2 | 100.00 | 2 | 100.00 |
| Intellectual disability | 1 | 1 | 100.00 | 1 | 100.00 |
| Hearing impairment | 0 | 0 | N/A | 0 | N/A |
| Speech or language impairment | 0 | 0 | N/A | 0 | N/A |
| Visual impairment | 0 | 0 | N/A | 0 | N/A |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 0 | 0 | N/A | 0 | N/A |
| Other health impairment | 1 | 1 | 100.00 | 1 | 100.00 |
| Specific learning disability | 0 | 0 | N/A | 0 | N/A |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 6 | 6 | 100.00 | 6 | 100.00 |
| Autism | 13 | 13 | 100.00 | 13 | 100.00 |
| Traumatic brain injury | 0 | 0 | N/A | 0 | N/A |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 5 | 5 | 100.00 | 5 | 100.00 |
| Economically disadvantaged | 16 | 16 | 100.00 | 16 | 100.00 |
| In US schools less than 12 months | 11 | 11 | 100.00 | 11 | 100.00 |
| In US schools 12 months or more | 9 | 9 | 100.00 | 9 | 100.00 |
| Duration unknown | 1 | 1 | 100.00 | 1 | 100.00 |
| Migrant education | 0 | 0 | N/A | 0 | N/A |
| Not migrant education | 21 | 21 | 100.00 | 21 | 100.00 |
| Armed forces family member | 0 | 0 | N/A | 0 | N/A |
| Not armed forces family member | 21 | 21 | 100.00 | 21 | 100.00 |
| Homeless | 0 | 0 | N/A | 0 | N/A |
| Not homeless | 21 | 21 | 100.00 | 21 | 100.00 |
| Foster youth | 0 | 0 | N/A | 0 | N/A |
| Not foster youth | 21 | 21 | 100.00 | 21 | 100.00 |

Table 7.B.4 **Demographic Summary for Students: Grade Span Three Through Five**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 45 | 43 | 95.56 | 42 | 93.33 |
| Male | 26 | 25 | 96.15 | 24 | 92.31 |
| Female | 19 | 18 | 94.74 | 18 | 94.74 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 1 | 1 | 100.00 | 1 | 100.00 |
| Asian | 7 | 7 | 100.00 | 6 | 85.71 |
| Native Hawaiian or Other Pacific Islander | 1 | 1 | 100.00 | 1 | 100.00 |
| Filipino | 3 | 3 | 100.00 | 3 | 100.00 |
| Hispanic or Latino | 22 | 22 | 100.00 | 22 | 100.00 |
| Black or African American | 0 | 0 | N/A | 0 | N/A |
| White | 7 | 6 | 85.71 | 6 | 85.71 |
| Two or more races | 4 | 3 | 75.00 | 3 | 75.00 |
| Intellectual disability | 10 | 9 | 90.00 | 8 | 80.00 |
| Hearing impairment | 0 | 0 | N/A | 0 | N/A |
| Speech or language impairment | 0 | 0 | N/A | 0 | N/A |
| Visual impairment | 0 | 0 | N/A | 0 | N/A |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 4 | 4 | 100.00 | 4 | 100.00 |
| Other health impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Specific learning disability | 0 | 0 | N/A | 0 | N/A |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 14 | 13 | 92.86 | 13 | 92.86 |
| Autism | 12 | 12 | 100.00 | 12 | 100.00 |
| Traumatic brain injury | 2 | 2 | 100.00 | 2 | 100.00 |
| Not classified | 1 | 1 | 100.00 | 1 | 100.00 |
| Not economically disadvantaged | 22 | 20 | 90.91 | 19 | 86.36 |
| Economically disadvantaged | 23 | 23 | 100.00 | 23 | 100.00 |
| In US schools less than 12 months | 23 | 23 | 100.00 | 23 | 100.00 |
| In US schools 12 months or more | 16 | 15 | 93.75 | 14 | 87.50 |
| Duration unknown | 6 | 5 | 83.33 | 5 | 83.33 |
| Migrant education | 0 | 0 | N/A | 0 | N/A |
| Not migrant education | 45 | 43 | 95.56 | 42 | 93.33 |
| Armed forces family member | 1 | 1 | 100.00 | 1 | 100.00 |
| Not armed forces family member | 44 | 42 | 95.45 | 41 | 93.18 |
| Homeless | 5 | 5 | 100.00 | 5 | 100.00 |
| Not homeless | 40 | 38 | 95.00 | 37 | 92.50 |
| Foster youth | 0 | 0 | N/A | 0 | N/A |
| Not foster youth | 45 | 43 | 95.56 | 42 | 93.33 |

Table 7.B.5 **Demographic Summary for Students: Grade Span Six Through Eight**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 24 | 23 | 95.83 | 23 | 95.83 |
| Male | 16 | 16 | 100.00 | 16 | 100.00 |
| Female | 8 | 7 | 87.50 | 7 | 87.50 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 0 | 0 | N/A | 0 | N/A |
| Asian | 6 | 5 | 83.33 | 5 | 83.33 |
| Native Hawaiian or Other Pacific Islander | 1 | 1 | 100.00 | 1 | 100.00 |
| Filipino | 1 | 1 | 100.00 | 1 | 100.00 |
| Hispanic or Latino | 12 | 12 | 100.00 | 12 | 100.00 |
| Black or African American | 0 | 0 | N/A | 0 | N/A |
| White | 3 | 3 | 100.00 | 3 | 100.00 |
| Two or more races | 1 | 1 | 100.00 | 1 | 100.00 |
| Intellectual disability | 7 | 6 | 85.71 | 6 | 85.71 |
| Hearing impairment | 0 | 0 | N/A | 0 | N/A |
| Speech or language impairment | 0 | 0 | N/A | 0 | N/A |
| Visual impairment | 0 | 0 | N/A | 0 | N/A |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 0 | 0 | N/A | 0 | N/A |
| Other health impairment | 4 | 4 | 100.00 | 4 | 100.00 |
| Specific learning disability | 0 | 0 | N/A | 0 | N/A |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 7 | 7 | 100.00 | 7 | 100.00 |
| Autism | 6 | 6 | 100.00 | 6 | 100.00 |
| Traumatic brain injury | 0 | 0 | N/A | 0 | N/A |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 11 | 10 | 90.91 | 10 | 90.91 |
| Economically disadvantaged | 13 | 13 | 100.00 | 13 | 100.00 |
| In US schools less than 12 months | 10 | 10 | 100.00 | 10 | 100.00 |
| In US schools 12 months or more | 12 | 12 | 100.00 | 12 | 100.00 |
| Duration unknown | 2 | 1 | 50.00 | 1 | 50.00 |
| Migrant education | 0 | 0 | N/A | 0 | N/A |
| Not migrant education | 24 | 23 | 95.83 | 23 | 95.83 |
| Armed forces family member | 1 | 1 | 100.00 | 1 | 100.00 |
| Not armed forces family member | 23 | 22 | 95.65 | 22 | 95.65 |
| Homeless | 3 | 3 | 100.00 | 3 | 100.00 |
| Not homeless | 21 | 20 | 95.24 | 20 | 95.24 |
| Foster youth | 1 | 1 | 100.00 | 1 | 100.00 |
| Not foster youth | 23 | 22 | 95.65 | 22 | 95.65 |

Table 7.B.6 **Demographic Summary for Students: Grade Span Nine Through Twelve**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 31 | 28 | 90.32 | 28 | 90.32 |
| Male | 21 | 18 | 85.71 | 18 | 85.71 |
| Female | 10 | 10 | 100.00 | 10 | 100.00 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 0 | 0 | N/A | 0 | N/A |
| Asian | 8 | 7 | 87.50 | 7 | 87.50 |
| Native Hawaiian or Other Pacific Islander | 0 | 0 | N/A | 0 | N/A |
| Filipino | 6 | 5 | 83.33 | 5 | 83.33 |
| Hispanic or Latino | 16 | 15 | 93.75 | 15 | 93.75 |
| Black or African American | 0 | 0 | N/A | 0 | N/A |
| White | 0 | 0 | N/A | 0 | N/A |
| Two or more races | 1 | 1 | 100.00 | 1 | 100.00 |
| Intellectual disability | 15 | 14 | 93.33 | 14 | 93.33 |
| Hearing impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Speech or language impairment | 0 | 0 | N/A | 0 | N/A |
| Visual impairment | 0 | 0 | N/A | 0 | N/A |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 0 | 0 | N/A | 0 | N/A |
| Other health impairment | 1 | 1 | 100.00 | 1 | 100.00 |
| Specific learning disability | 1 | 1 | 100.00 | 1 | 100.00 |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 2 | 2 | 100.00 | 2 | 100.00 |
| Autism | 9 | 7 | 77.78 | 7 | 77.78 |
| Traumatic brain injury | 0 | 0 | N/A | 0 | N/A |
| Not classified | 1 | 1 | 100.00 | 1 | 100.00 |
| Not economically disadvantaged | 16 | 14 | 87.50 | 14 | 87.50 |
| Economically disadvantaged | 15 | 14 | 93.33 | 14 | 93.33 |
| In US schools less than 12 months | 14 | 14 | 100.00 | 14 | 100.00 |
| In US schools 12 months or more | 16 | 13 | 81.25 | 13 | 81.25 |
| Duration unknown | 1 | 1 | 100.00 | 1 | 100.00 |
| Migrant education | 0 | 0 | N/A | 0 | N/A |
| Not migrant education | 31 | 28 | 90.32 | 28 | 90.32 |
| Armed forces family member | 0 | 0 | N/A | 0 | N/A |
| Not armed forces family member | 31 | 28 | 90.32 | 28 | 90.32 |
| Homeless | 0 | 0 | N/A | 0 | N/A |
| Not homeless | 31 | 28 | 90.32 | 28 | 90.32 |
| Foster youth | 0 | 0 | N/A | 0 | N/A |
| Not foster youth | 31 | 28 | 90.32 | 28 | 90.32 |

## Psychometric Analyses

This chapter summarizes the item- and test-level statistics from the analyses conducted for the 2022–23 administration of the Initial Alternate English Language Proficiency Assessments for California (ELPAC).

### Overview

This chapter describes the psychometric analyses conducted by ETS for the Initial Alternate ELPAC, including classical item analyses and analyses to support reliability and validity evidence.

#### Summary of the Analyses

Each of these analyses of the Initial Alternate ELPAC data is presented in the body of the text and in the listed appendices.

1. **Classical Item Analyses—**Classical item analysis for the Initial Alternate ELPAC is described in section [*8.2 Classical Item Analyses*](#_Demographic_Student_Group). The results of the item-level classical item analyses, by grade level and grade span, including item difficulty indices (*p*-values), and item-total correlation coefficients for dichotomous and polytomous items are provided in table 8.A.1 through table 8.A.6 in [appendix 8.A](#_Alternative_Text_for_59). Summary statistics, including the mean, minimum, and maximum values are presented in table 8.2.
2. **Item Response Theory (IRT) Analyses—**IRT models and analyses are presented in section [*8.3 Item Response Theory Analyses*](#_Item_Response_Theory). Table 8.3 presents the slopes and intercepts that convert theta scores to reported scale scores for the operational items. Table 8.B.1 through table 8.B.6 in [appendix 8.B](#_Appendix_7.C:_Item) provide IRT results for individual items by grade level or grade span.
3. **Reliability Analyses—**Reliability estimation for the Initial Alternate ELPAC is illustrated in section [*8.4 Reliability Analyses*](#_Reliability__Analyses_1).
4. **Validity Evidence—**Validity evidence related to the Initial Alternate ELPAC is discussed in section [*8.5 Validity Evidence*](#_Validity_Evidence).

#### Samples Used for the Analyses

All analyses were conducted using the final data sample from June 2023. Students who did not respond to at least one receptive item and one expressive item were removed from the analysis sample.

There are two types of missing data: (1) “marked as no response” and omitted responses and (2) not-reached responses. These two types are treated differently in the analyses. The first occurs when a question has been received in the test delivery system but was not answered (i.e., the question was left blank or skipped) even though the student received and responded to subsequent items. The second is generated when a student ends the assessment early. A student may not reach the end of an assessment, or the item or items were not presented to the student. “Marked as no response” and omitted responses were treated as incorrect, and not-reached responses were treated as not presented in all statistical analyses.

#### Test-Taking Rates

The decision to assign a student to take the Initial Alternate ELPAC is determined by the student’s individualized education program (IEP) team using the information in the California Department of Education (CDE) Alternate Assessment Decision-Making Tool for California web document. This web document describes the criteria for taking alternate assessments and the students who should be identified to take alternate assessments (CDE, 2023a).

All students who are identified by an IEP team to take the Initial Alternate ELPAC are required to take alternate assessments for all state standardized assessments. All students who are logged on and respond to at least one receptive item and one expressive item are counted as having taken the assessment and, therefore, have a valid score (CDE, 2020).

Table 8.1 shows the test-taking rates by grade level.

Table 8.1 The Test-Taking Rates by Grade Level

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level | Number of Registered Students | Number of Test Takers | Number of Students with Valid Score | Test Takers as a Percent of Registered Students |
| Kindergarten | 691 | 689 | 683 | 99.7 |
| 1 | 61 | 60 | 59 | 98.4 |
| 2 | 21 | 21 | 21 | 100.0 |
| 3 | 14 | 13 | 13 | 92.9 |
| 4 | 15 | 15 | 14 | 100.0 |
| 5 | 16 | 15 | 15 | 93.8 |
| 6 | 12 | 11 | 11 | 91.7 |
| 7 | 4 | 4 | 4 | 100.0 |
| 8 | 8 | 8 | 8 | 100.0 |
| 9 | 8 | 7 | 7 | 87.5 |
| 10 | 3 | 3 | 3 | 100.0 |
| 11 | 4 | 3 | 3 | 75.0 |
| 12 | 16 | 15 | 15 | 93.8 |
| **Total:** | **873** | **864** | **856** | **99.0** |

### Classical Item Analyses

The classical item analyses include the item difficulty indices and the item-total correlation indices. Items that are not performing as expected are identified on the basis of flagging rules associated with the item statistics. 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 polytomous items are also included in the results of the classical item analyses.

#### Classical Item Difficulty Indices (*p*-value and Average Item Score)

Items scored as one (correct) or zero (incorrect) are referred to as dichotomous items. Items scored from zero to some number of points greater than one are called polytomous items.

For dichotomous items, item difficulty is indicated by its *p*-value, which is the proportion of students who answer the item correctly. The range of *p*-values is from 0.00 to 1.00. Items with high *p*-values are easier items; those with low *p*-values are more difficult. Dichotomous items are flagged for review if their *p*-values are above 0.95 (i.e., too easy). Two-choice dichotomous single-select items, three-choice dichotomous single-select items, and all other dichotomous items are flagged as too difficult if their *p*-values are below 0.50, 0.30, and 0.20, respectively.

The formula for the *p*-value for a dichotomous item is presented in equation 8.1. *Refer to the* [*Alternative Text for Equation 8.1*](#_Alternative_Text_for_2) *for a description of this equation.*

 (8.1)

where,

*Xij* is the score (0 or 1) received for a given dichotomous item *i* for student *j*, and

*Ji* is the total number of students who were presented with item *i*.

For polytomous items, the difficulty is indicated by either the average item score (AIS) or *p*-‍value. The AIS can range from 0.00 to the maximum total possible points for an item. Desired AIS values for polytomous items generally fall within the range of 20 percent to 80 percent of the maximum obtainable item score; items with values outside this range are flagged for review. To facilitate the interpretation, the AIS values for polytomous items are often expressed as the proportion of the maximum possible score, which are equivalent to the *p-*values for dichotomous items.

For polytomous items, the *p-*value is defined as presented in equation 8.2. *Refer to the* [*Alternative Text for Equation 8.2*](#_Alternative_Text_for_38) *for a description of this equation.*

 (8.2)

where,

*Xij* is the score assigned for a given polytomous item *i* and student *j*,

*Ji* is the total number of students who were presented with item *i*, and

*Mi* is the maximum possible score for item *i*.

#### Item-Total Correlation

An important indicator of item discrimination is the item-total correlation, defined as the correlation between student scores on an individual item and student “total” scores on the assessment.

The item-total correlation statistic describes the relationship between students’ performance on a specific item and students’ performance on the total assessment. It is calculated as the correlation coefficient between the item score and total score—specifically, the polyserial correlation is used as the index of item-total correlation for both polytomous and dichotomous items. Statistically, it is calculated as the correlation between an observed continuous variable and an unobserved continuous variable hypothesized to underlie the variable with ordered categories (Olsson, Drasgow, & Dorans, 1982). The total scale score or the raw score is used as the criterion score for this analysis.

Theoretically, the polyserial correlation ranges from −1.0 (for a perfect negative relationship) to 1.0 (for a perfect positive relationship) and is estimated as presented in equation 8.3. *Refer to the* [*Alternative Text for Equation 8.3*](#_Alternative_Text_for_39) *for a description of this equation.*

 (8.3)

where,

*β* is the item parameter to be estimated from the data, with the estimate denoted as , using maximum likelihood estimation; it is a regression coefficient (slope) for predicting the continuous version of an item score onto the continuous version of the total score;

*s2tot* is the variance of the criterion (for example, the students’ total score); and

*stot* is the standard deviation (SD) of the criterion.

For a polytomous item, there is a regression for each boundary between item scores, with all regressions for the same item sharing a common slope, *β*. For a polytomous item with *m* possible score values, there are *m*−1 regressions.

Acceptable values for this correlation coefficient are positive and greater than 0.20. A relatively high item-total correlation coefficient value is preferred, as it indicates that higher-performing students tend to perform better on the item than lower-performing students. An item with a negative item-total correlation typically signifies a problem with the item, as that indicates that

* the higher-performing students on the overall assessment tend to respond incorrectly to the item if dichotomous, or are assigned a low score for the item if polytomous; or
* the lower-performing students on the overall assessment are responding correctly to the item if dichotomous, or are assigned a high score for that item if polytomous.

#### Distribution of Item Scores

For polytomous items, examination of the distribution of scores assists in showing how well items performed. If no students were given the highest possible score, the item may not be functioning as expected because the item may be confusing, poorly worded, or just unexpectedly difficult; the scoring rubric may be flawed; or students may not have had an opportunity to learn the content. If the rubric for an item allowed for partial credit but nearly all students received either full credit or partial credit, the rubric should be reviewed for whether the rubric for the partial credit score category should be revised.

#### Omit Rates

If a student views an item, leaves it unanswered, and then goes on to view and answer another item, the missing response is classified as an “omit.” If the student omits an item—that is, leaves the item unanswered—and does not view additional items, the responses for the successive items are classified as “not seen.”

##### Rates for Dichotomous and Polytomous Items

For both dichotomous and polytomous items, examining the omit rate is useful for identifying potential problems with test features such as testing time and item or test layout. Items with high omit rates are flagged for further investigation by content specialists to ensure that no issues are found with these items. Omit rates for polytomous items tend to be higher than for dichotomous items.

##### No-Response Rate

The *Mark as No Response* contextual menu option is a specific case of an omitted item. The *Mark as No Response* option should be used when the item was presented to the student and the student did not provide a response despite the test examiner’s best efforts to elicit a response. Similar to the omit rate, the Mark as No Response information is useful for identifying potential problems with an item.

#### Classical Item Analyses Results

Table 8.2 presents the summary statistics of item difficulty and item discrimination for all operational items. The *p*-value columns contain *p*-values for dichotomous items and AIS for polytomous items. For both item types, these values can be interpreted as the proportion of students obtaining full credit on the item.

Table 8.2 Summary Statistics for Classical Item Analysis

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Number of Items | Number of Students | Mean *p*-value | Minimum *p*-value | Maximum *p*-value | Mean Item-Total Correlation | Minimum Item-Total Correlation | Maximum Item-Total Correlation |
| Kindergarten | 24 | 656 | 0.35 | 0.13 | 0.64 | 0.67 | 0.49 | 0.82 |
| 1 | 24 | 57 | 0.36 | 0.23 | 0.56 | 0.70 | 0.44 | 0.89 |
| 2 | 24 | 20 | 0.38 | 0.10 | 0.75 | 0.71 | −0.11 | 1.00 |
| 3–5 | 24 | 39 | 0.27 | 0.13 | 0.49 | 0.70 | 0.34 | 0.98 |
| 6–8 | 24 | 22 | 0.39 | 0.18 | 0.73 | 0.81 | 0.53 | 0.98 |
| 9–12 | 24 | 25 | 0.47 | 0.16 | 0.78 | 0.63 | 0.36 | 0.86 |

The item difficulty and item discrimination statistics for each item by grade level or grade span are presented in table 8.A.1 through table 8.A.6 in [appendix 8.A](#_Alternative_Text_for_59).

The minimum *p*-value at each grade level or grade span is less than the target minimum *p*-value, but the majority of items have much higher *p*-values, as can be seen in the aforementioned appendix tables. One item in grade two had a negative polyserial correlation, suggesting an inverse relationship between answering the item correctly and doing well on the assessment as a whole. However, in the 2021–22 operational field test administration, this item, VR151643, had a polyserial correlation of 0.44, based on 1,134 test-takers. The very small sample sizes at every grade level or grade span, except for kindergarten, are likely to lead to large fluctuations in item statistics between test administrations.

### Item Response Theory Analyses

IRT is a family of mathematical models that characterizes the probability of a given response as a function of a test taker’s true ability and one or more features of the items, such as its difficulty or discrimination. IRT can be used to calibrate items, link item parameter estimates, scale or equate test scores across different forms or test administrations, evaluate item performance, build an item bank, and assemble test forms.

This section describes how IRT models were used during the 2021–22 Alternate ELPAC operational field test administration to calibrate all the operational field test items to establish the base IRT scale. This administration was the source of all items on the Initial Alternate ELPAC, so all calibration activities refer to the 2021–22 Alternate ELPAC operational field test administration.

#### Item Response Theory Model

The one-parameter logistic item response theory (1PL-IRT) model was used for the item calibration and was selected after consultation with the CDE. In particular, the generalized partial credit model (GPCM) (Muraki, 1992) restricted for 1PL-IRT, which is essentially the partial credit model (Masters, 1982), was applied to both dichotomous and polytomous items.

The mathematical form of the GPCM is presented in equation 8.4. *Refer to the* [*Alternative Text for Equation 8.4*](#_Alternative_Text_for_40) *for a description of this equation.*

 (8.4)

where,

 is the probability of student with proficiency  obtaining score *h* on item *i*,

*Mi* is the maximum number of score points for item *i*,

*ai* is the discrimination parameter, which is fixed to 0.588 for every item,

*bi* is the location parameter for item *i*,

*div* is the category parameter for item *i* on item score *v*,

*D* is a scaling constant of 1.7,

*c* indexes the item score, and

*v* indexes the non-zero item score.

When *Mi* = 1, equation 8.4 becomes an expression of the one-parameter logistic model for dichotomous items.

#### Equating

Equating is a procedure where test scores, from different test forms assembled on the basis of the same specifications, are placed onto a reference scale so that scores from different test administrations are comparable. There are two approaches to equate the test forms: preequating and postequating.

A preequating design allows for conversion tables that describe the relationship between raw scores and scale scores, or theta scores and scale scores, to be established prior to the current test administration using data from prior administrations. Preequating relies on having a well-calibrated item bank, robust embedded field-testing processes, and stability in item performance over time.

A postequating design uses the data from the current administration to establish the raw-to-scale-score relationship for the current administration’s form.

Both preequating and postequating involve a common‑item nonequivalent groups design (Kolen & Brennan, 2004).

The Initial Alternate ELPAC was preequated using the IRT parameters from the base scale established in the 2021–22 Alternate ELPAC operational field test administration. For detailed information on the method to establish the raw score to scale relationship, refer to subsection [*8.3.3* *Scaling the Scores*](#_Toc120784038)*.*

#### Scaling the Scores

The raw scores on each new form were transformed to scale scores on the reference scale using a two-step procedure. First, the number-correct scores (raw scores) were transformed to ability (theta) scores on the reference scale by the inverse test characteristic curve (TCC) procedure described in the next subsection. Then, these ability (theta) scores were transformed to scale scores through the linear transformation described in subsection [*8.3.3.2 Transformation from Theta Scores to Scale Scores*](#_Transformation_from_Theta_3).

##### Inverse Test Characteristic Curve Procedure

After all the item difficulty estimates are transformed to the reference scale, students’ overall ability estimates can be derived from the input data filethrough the IRT inverse TCC method (Stocking, 1996). This method transforms the sum of the student’s item scores into an ability estimate. That estimate is the ability value that makes the sum of the expected scores on the items administered to the student equal to the sum of the scores that the student actually received on those items.

The TCC expresses the expected total score on a set of items as a function of the student’s ability, which is shown in equation 8.5. *Refer to the* [*Alternative Text for Equation 8.5*](#_Alternative_Text_for_44) *for a description of this equation.*

 (8.5)

where,

*i* indexes dichotomous items,

*j* indexes polytomous items,

*ndich* is the number of dichotomous items in the assessment,

*pi(θ)* is the probability of a correct response to item *i* at ability *θ* on the dichotomous item in equation 8.4,

*npoly* is the number of polytomous items in the assessment,

*m* is the number of score categories for each polytomous item,

*sxj* is the value for score category x for the polytomous item *j*,

*pxj(θ)* is the probability that an examinee with ability *θ* obtains score sx on the polytomous item *j* in equation 8.4, and

*ξ(θ)* is the corresponding expected total score.

##### Transformation from Theta Scores to Scale Scores

Students’ ability estimates (theta scores) were transformed to the scale score metric by applying a linear transformation based on threshold theta values. Those threshold values were determined after standard setting and approved by the California State Board of Education (SBE). There were two threshold theta values (for Level 2 and Level 3) for each grade level or grade span. The scaling transformation was the linear transformation that transformed the Level 2 threshold to scale score 44 and the Level 3 threshold to scale score 60 (refer to equations 8.6, 8.7, and 8.8). *Refer to the* [*Alternative Text for Equation 8.6*](#_Alternative_Text_for_46) *for a description of this equation.*

 (8.6)

where,

 denotes the scale score for student *j*,

 represents student ability estimate for student *j*,

*A* is the slope parameter (scaling factor) needed to transform theta to the scale score metric, and

*B* is the intercept parameter needed to transform theta to the scale score metric.

The slope and intercept parameters are derived by mapping the equated Level 2 and Level 3 threshold scores from the standard setting to the prespecified scale score threshold scores. Specifically, the slope and intercept in equation 8.6 are derived using the threshold scores from standard setting approved by the SBE ( and  in equations 8.7 and 8.8) and the desired threshold scale scores (two-digit scale score) ( and  in equations 8.7 and 8.8) using the following two formulas:

*Refer to the* [*Alternative Text for Equation 8.7*](#_Alternative_Text_for_47) *for a description of this equation.*

 (8.7)

*Refer to the* [*Alternative Text for Equation 8.8*](#_Alternative_Text_for_48) *for a description of this equation.*

 (8.8)

where,

 represents the threshold score for Level 3—Alternate on the reporting scale, which is set to be 60;

 represents the threshold score for Level 2—Alternate on the reporting scale, which is set to be 44;

 represents the threshold score for Level 3—Alternate on the theta scale; and

 represents the threshold score for Level 2—Alternate on the theta scale. (For more information on  and , refer to [*Chapter 6: Standard Setting*](#_Appendix_4.A:_Demographic).)

The slopes and intercepts for each grade level are shown in table 8.3. Also, refer to subsection [*7.1.3 Scale Scores for the Total Assessment*](#_Scale_Scores_for)for the special requirements for the Initial Alternate ELPAC reporting scale.

Table 8.3 Slopes and Intercepts That Convert Theta Scores to Reporting Scale Scores

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Threshold Theta Score for Level 2—Alternate | Threshold Theta Score for Level 3—Alternate | Reporting Scale Score for Level 2—Alternate | Reporting Scale Score for Level 3—Alternate | Slope | Intercept |
| Kindergarten | −0.0341 | 1.4732 | 144 | 160 | 10.62 | 44.35 |
| 1 | −0.2819 | 1.1807 | 144 | 160 | 10.94 | 47.08 |
| 2 | −0.1351 | 1.2967 | 144 | 160 | 11.17 | 45.52 |
| 3ؘ–5 | −0.7329 | 1.0099 | 144 | 160 | 9.18 | 50.73 |
| 6–8 | −1.2416 | 0.5178 | 144 | 160 | 9.09 | 55.29 |
| 9–12 | −1.0927 | 0.9151 | 144 | 160 | 7.97 | 52.71 |

#### Parameter Estimates

Ranges of IRT *b*-parameter estimates for each grade level or grade span are shown in table 8.4. These parameter estimates are derived from the 2021–22 Alternate ELPAC operational field test administration. Summary statistics (i.e., minimum, maximum, mean, and SD values) are also presented.

The means of *b*-parameter estimates range from −1.26 to 0.65. The mean item difficulty level decreased for higher grade levels or grade spans. Overall, there are not many items with large negative *b-*parameters (i.e., very easy items), and even fewer with large positive *b*-values (i.e., very difficult items). All *b*-parameters are within acceptable ranges.

Table 8.4 Distribution of *b*-values

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| IRT *b*-value | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–5 | Grade Span 6–8 | Grade Span 9–12 |
| −4.0 ≤ b < −3.8 | 0 | 0 | 0 | 0 | 0 | 0 |
| −3.8 ≤ b < −3.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| −3.6 ≤ b < −3.4 | 0 | 0 | 0 | 0 | 1 | 0 |
| −3.4 ≤ b < −3.2 | 0 | 0 | 0 | 0 | 0 | 0 |
| −3.2 ≤ b < −3.0 | 0 | 0 | 0 | 0 | 0 | 2 |
| −3.0 ≤ b < −2.8 | 0 | 0 | 0 | 0 | 0 | 0 |
| −2.8 ≤ b < −2.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| −2.6 ≤ b < −2.4 | 0 | 0 | 0 | 0 | 0 | 0 |
| −2.4 ≤ b < −2.2 | 0 | 0 | 0 | 0 | 1 | 0 |
| −2.2 ≤ b < −2.0 | 0 | 0 | 0 | 0 | 0 | 1 |
| −2.0 ≤ b < −1.8 | 0 | 0 | 0 | 1 | 3 | 2 |
| −1.8 ≤ b < −1.6 | 0 | 0 | 0 | 0 | 2 | 3 |
| −1.6 ≤ b < −1.4 | 0 | 1 | 0 | 2 | 2 | 2 |
| −1.4 ≤ b < −1.2 | 0 | 0 | 0 | 0 | 2 | 3 |
| −1.2 ≤ b < −1.0 | 0 | 0 | 2 | 1 | 3 | 2 |
| −1.0 ≤ b < −0.8 | 2 | 1 | 2 | 4 | 2 | 3 |
| −0.8 ≤ b < −0.6 | 0 | 4 | 1 | 3 | 1 | 1 |
| −0.6 ≤ b < −0.4 | 0 | 0 | 1 | 2 | 2 | 2 |
| −0.4 ≤ b < −0.2 | 1 | 2 | 2 | 6 | 1 | 0 |
| −0.2 ≤ b < 0.0 | 1 | 1 | 2 | 1 | 1 | 0 |
| 0.0 ≤ b < 0.2 | 3 | 3 | 5 | 0 | 1 | 2 |
| 0.2 ≤ b < 0.4 | 3 | 3 | 4 | 2 | 2 | 0 |
| 0.4 ≤ b < 0.6 | 0 | 4 | 0 | 1 | 0 | 0 |
| 0.6 ≤ b < 0.8 | 5 | 2 | 3 | 1 | 0 | 0 |
| 0.8 ≤ b < 1.0 | 1 | 1 | 1 | 0 | 0 | 1 |
| 1.0 ≤ b < 1.2 | 1 | 2 | 0 | 0 | 0 | 0 |
| 1.2 ≤ b < 1.4 | 3 | 0 | 0 | 0 | 0 | 0 |
| 1.4 ≤ b < 1.6 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1.6 ≤ b < 1.8 | 1 | 0 | 1 | 0 | 0 | 0 |
| 1.8 ≤ b < 2.0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2.0 ≤ b < 2.2 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2.2 ≤ b < 2.4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2.4 ≤ b < 2.6 | 1 | 0 | 0 | 0 | 0 | 0 |
| 2.6 ≤ b < 2.8 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2.8 ≤ b < 3.0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.0 ≤ b < 3.2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.2 ≤ b < 3.4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.4 ≤ b < 3.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.6 ≤ b < 3.8 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3.8 ≤ b ≤ 4.0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Minimum | −0.94 | −1.47 | −1.15 | −1.86 | −3.45 | −3.20 |
| Maximum | 2.51 | 1.18 | 1.76 | 0.71 | 0.35 | 0.90 |
| Mean | 0.65 | 0.08 | 0.02 | −0.53 | −1.12 | −1.26 |
| SD | 0.85 | 0.66 | 0.68 | 0.62 | 0.88 | 0.92 |
| **Number of Items:** | **24** | **24** | **24** | **24** | **24** | **24** |

### Reliability Analyses

The reliability for a particular group of students’ test scores is the extent to which the scores would remain consistent if those same students were retested with a parallel version of the same assessment. There are many definitions of reliability (Haertel, 2006) that have their genesis in classical test theory and a variety of methods that can be used to estimate reliability.

The general concept of reliability concerns the extent to which the test scores measure *a particular construct* consistently. The variance in the distribution of test scores—essentially, the observed differences among individuals—is partly due to differences that are consistent and partly due to differences that are not consistent. The measure of variation associated with the first kind of differences—consistent differences—is called “true variance”; this would include actual differences in students’ knowledge. The measure of variation associated with the remaining differences—those that operate essentially at random—is called “error variance.” Error variance includes a variety of underlying differences such as selections of test content, which may cause a student’s test score to be slightly higher in one evaluation and slightly lower in another. Reliability is the proportion of total variance that is due to true variance. The standard error of measurement (SEM) is a statistic that characterizes the error variance.

Reliability coefficients range from zero to one. The higher the reliability coefficient for a set of scores, the more likely individuals are to obtain very similar scores upon repeated testing occasions, if the students do not change in their level of the knowledge or skills measured by the assessment.

#### Reliability Measures

In a specified population of students, the reliability of test scores, *X*, is defined as the proportion of the test score variance that is attributable to true differences in student abilities and is sometimes operationalized as the correlation between scores on two replications of the same testing procedure, .

Reliability coefficients may range from 0 to 1. The higher the reliability coefficient for a set of scores, the more likely students would be to obtain very similar scores if they were retested. In applied settings, the requirement of repeated administrations is impractical, and methodologies estimating reliability from relationships among student performances on items within a single test form are often used.

An IRT-based approach called marginal reliability (Green et al., 1984) can be used to estimate the reliability of the scores. The estimates of reliability coefficients reported here are for IRT-based ability estimates.

This reliability coefficient for theta estimates, , is defined on the basis of a single test administration, as shown in equation 8.9. *Refer to the* [*Alternative Text for Equation 8.9*](#_Alternative_Text_for_49) *for a description of this equation.*

 (8.9)

where,

*J* is the number of students who took the assessment,

 is the measure of variance in ability estimates, and

 is the squared conditional standard error of measurement (CSEM) (i.e., error variances) for student *j* with ability estimate .

#### Standard Error of Measurement

The SEM is a measure of how much students’ scores would vary from the scores they would earn on a perfectly reliable assessment. 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 SEMs are an indication of how much the errors of measurement affect 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.

The SEM is the square root of the error variance in the scores (i.e., the SD of the distribution of the differences between students’ observed scores and their true scores). The SEM is calculated using equation 8.10. *Refer to the* [*Alternative Text for Equation 8.10*](#_Alternative_Text_for_64) *for a description of this equation.*

 (8.10)

where,

 is the reliability estimated in equation 8.9,

 is the SD of the total test theta score, and

*A* is the slope of the scaling transformation of theta scoresto the reporting scale.

The SEM is useful in determining the confidence interval (CI) that likely captures a student’s true score. A student’s true score can be thought of as the mean of observed scores a student would earn over an infinite number of independent administrations of the assessment. Across those administrations, approximately 95 percent of the time the interval ranging from the student’s observed score minus 1.96 SEMs to the student’s observed score plus 1.96 SEMs would contain that student’s true score (Crocker & Algina, 1986). Therefore, this interval is called a 95 percent CI for the student’s true score. For example, if a student’s observed score on a given assessment equals 150 points and the SEM equals 5, one can be 95 percent confident that the student’s true score lies between 140 and 160 points (150 ± 10).

Table 8.5 gives the mean and SD of the scale scores, the reliability, and the SEM for the test form for each grade level or grade span. These results indicate that the reliability estimates for all assessments are moderately high. Results based on samples that contain 50 or fewer students should be interpreted with caution because these estimates may meaningfully deviate from the true reliability.

Table 8.5 Reliability of Reporting Scale Scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Students | Mean | SD | Reliability | SEM—Scale Score |
| Kindergarten | 591 | 144 | 14.2 | 0.85 | 5.47 |
| 1 | 49 | 142 | 16.4 | 0.87 | 5.79 |
| 2 | 20 | 138 | 16.1 | 0.87 | 5.78 |
| 3–5 | 29 | 139 | 11.9 | 0.84 | 4.83 |
| 6–8 | 18 | 145 | 15.1 | 0.89 | 4.95 |
| 9–12 | 23 | 144 | 7.9 | 0.82 | 3.37 |

#### Student Group Reliabilities

Table 8.6 provides student group reliabilities and SEM statistics for kindergarten. The student groups included in these analyses were defined by their gender, ethnicity, primary disability type, economic status, enrollment in US schools, migrant status, military status, homeless status, and foster youth status. Student group reliabilities and SEM statistics are not provided for the other grade levels or grade spans because of insufficient sample sizes.

Reliability values are estimates that approach the true reliability as the number of students whose scores contribute to the estimates increases. Reliabilities are not reported for samples that comprise 10 or fewer students. Results based on samples that contain 50 or fewer students should be interpreted with caution because these estimates may meaningfully deviate from the true reliability. In some cases, score reliabilities could not be estimated and are presented in the tables as “N/A.”

Table 8.6 Student Group Reliabilities for Kindergarten

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | Kindergarten N | Kindergarten Reliability | Kindergarten SEM |
| All | 591 | 0.85 | 5.47 |
| Male | 417 | 0.85 | 5.47 |
| Female | 174 | 0.85 | 5.47 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 0 | N/A | N/A |
| Asian | 86 | 0.85 | 5.07 |
| Native Hawaiian or Other Pacific Islander | 1 | N/A | N/A |
| Filipino | 7 | N/A | N/A |
| Hispanic or Latino | 449 | 0.85 | 5.49 |
| Black or African American | 8 | N/A | N/A |
| White | 25 | 0.90 | 6.29 |
| Two or more races | 15 | 0.88 | 5.44 |
| Intellectual disability | 115 | 0.82 | 5.26 |
| Hearing impairment | 1 | N/A | N/A |
| Speech or language impairment | 27 | 0.88 | 5.44 |
| Visual impairment | 0 | N/A | N/A |
| Emotional impairment | 0 | N/A | N/A |
| Orthopedic impairment | 6 | N/A | N/A |
| Other health impairment | 41 | 0.85 | 5.09 |
| Specific learning disability | 0 | N/A | N/A |
| Deaf-blindness | 0 | N/A | N/A |
| Multiple disabilities | 26 | 0.75 | 5.90 |
| Autism | 374 | 0.86 | 5.52 |
| Traumatic brain injury | 1 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 204 | 0.85 | 5.59 |
| Economically disadvantaged | 387 | 0.85 | 5.40 |
| In US schools less than 12 months | 526 | 0.85 | 5.45 |
| In US schools 12 months or more | 21 | 0.86 | 5.20 |
| Duration unknown | 44 | 0.88 | 5.77 |
| Migrant education | 9 | N/A | N/A |
| Not migrant education | 582 | 0.85 | 5.45 |
| Armed forces family member | 6 | N/A | N/A |
| Not armed forces family member | 585 | 0.85 | 5.47 |
| Homeless | 21 | 0.85 | 5.14 |
| Not homeless | 570 | 0.85 | 5.48 |
| Foster youth | 6 | N/A | N/A |
| Not foster youth | 585 | 0.85 | 5.47 |

Compared to reliability for the overall group of kindergarten test takers, reliability was lower for students with multiple disabilities and was slightly lower for students with intellectual disabilities. Every other student group with sufficient sample sizes to report results had scores that were at least as reliable, if not more reliable, than the overall sample. For instance, reliability was higher for White students and was slightly higher for students reporting two or more races.

#### Conditional Standard Errors of Measurement

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 mid-score 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.

##### Methodology

CSEMs are estimated as part of the IRT-based scoring procedure. CSEMs for scale scores are based on IRT and are estimated as a function of measured ability. The CSEMs of theta scores (or of linearly transformed theta scores) are smaller at points of the scale in the test metric where more items are located. A student’s CSEM under the IRT framework is equal to the reciprocal of the square root of the test information function based on the items taken by each student. The CSEM for a student with proficiency  is calculated using equation 8.11. *Refer to the* [*Alternative Text for Equation 8.11*](#_Alternative_Text_for_65) *for a description of this equation.*

 (8.11)

where,

 is the test information for student *j* and is calculated using equation 8.12. *Refer to the* [*Alternative Text for Equation 8.12*](#_Alternative_Text_for_54) *for a description of this equation.*

 (8.12)

where,

*I* is the number of items on the test form, and

 is the item information of item *i* for student *j*.

Item information is calculated as presented in equation 8.13. *Refer to the* [*Alternative Text for Equation 8.13*](#_Alternative_Text_for_55) *for a description of this equation.*

 (8.13)

where,

 and  are the first and second order moments of the item score for item *i* for a student with theta score .

The expected score of item *i* for student *j* is calculated as presented in equation 8.14. *Refer to the* [*Alternative Text for Equation 8.14*](#_Alternative_Text_for_56) *for a description of this equation.*

 (8.14)

The expected squared score of item *i* for student *j* is calculated as presented in equation 8.15. *Refer to the* [*Alternative Text for Equation 8.15*](#_Alternative_Text_for_57) *for a description of this equation.*

 (8.15)

where is the probability of an examinee with proficiency  obtaining score *h* on item *I*, the computation of which is shown in equation 8.16. *Refer to the* [*Alternative Text for Equation 8.16*](#_Alternative_Text_for_37) *for a description of this equation.*

P sub I h of theta sub j equals the numerator exp open parenthesis the sum from v equals 1 to h of D times a sub i of the quantity open parenthesis theta sub j minus b sub I plus d sub iv close parenthesis close parenthesis and denominator 1 plus the sum from c equals 1 to n sub I exp open parenthesis the sum from v equals 1 to c D times a sub i of the quantity open parenthesis theta sub j minus b sub I plus d sub iv close parenthesis close parenthesis, if score h equals 1, 2, …., n sub i.

P sub I h of theta sub j equals 1 divided by denominator 1 plus the sum from c equals 1 to n sub I exp open parenthesis the sum from v equals 1 to c D times a sub i of the quantity open parenthesis theta sub j minus b sub I plus d sub iv close parenthesis close parenthesis, if score h equals 0. (8.16)

where,

*Mi* is the maximum number of score points for item *i*;

*ai* is the discrimination parameter for item *i,* which is fixed to 0.588 for all items;

*bi* is the location parameter for item *i*;

*div* is the category parameter for item *i* on score *v*;

*D* is a scaling constant of 1.7;

*c* indexes the item score; and

*v* indexes the non-zero item score.

When *Mi* = 1, equation 8.16 becomes an expression of the two-parameter logistic model for dichotomous items.

CSEMs for scale scores are computed by transforming CSEMs of theta scores onto the reporting scale. Refer to subsection [*7.1.3* *Scale Scores for the Total Assessment*](#_Scale_Scores_for) for scaling procedures. A student’s CSEM for scale scores under the IRT framework is equal to the CSEM for the theta score multiplied by the scaling factor *A*, as presented in equation 8.17. *Refer to the* [*Alternative Text for Equation 8.17*](#_Alternative_Text_for_66) *for a description of this equation.*

 (8.17)

where,

 is the CSEM on the scale score metric for student *j*;

 is the CSEM on the theta score metric for student *j* estimated in equation 8.11;

 is the  test information for student *j* as calculated in equation 8.12; and

*A* is the scaling factor (the slope) needed to transform theta to the scale score metric calculated in equation 8.7.

Table 8.7 presents the scale score CSEMs at the lowest scale score required for a student to be classified in the score reporting range two and score reporting range three for the Initial Alternate ELPAC.

Table 8.7 Scale Score CSEMs at the Score Reporting Range Thresholds

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Level 2―‌Score Reporting Range Threshold | Level 2—‌CSEM | Level 3―‌Score Reporting Range Threshold | Level 3—‌CSEM |
| Kindergarten | 144 | 4 | 160 | 4 |
| 1 | 144 | 4 | 160 | 5 |
| 2 | 144 | 4 | 160 | 5 |
| 3–5 | 144 | 3 | 160 | 4 |
| 6–8 | 144 | 3 | 160 | 5 |
| 9–12 | 144 | 3 | 160 | 5 |

#### Decision Classification Analyses

When an assessment uses performance levels as the primary method to report test results, accuracy and consistency of decisions become key indicators of the quality of the assessment.

##### Methodology

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

Decision accuracy is the extent to which students are classified in the same way as they would be if each student’s score were the average over all possible forms of the assessment (the student’s true score). Decision accuracy answers the following question: How closely does the actual classification of test takers, based on their single-form scores, agree with the classification that would be made on the basis of their true scores, if their true scores could somehow be known?

Decision consistency is the extent to which students are classified in the same way as they would be on the basis of a single form of an assessment other than the one for which data is available. Decision consistency answers the following question: What is the agreement between the classifications based on two nonoverlapping, equally difficult forms of the assessment?

The methodology used for estimating the reliability of classification decisions is described in Livingston and Lewis (1995). The necessary input information includes only the maximum and minimum possible scores on the assessment and the observed score distribution and the reliability coefficient for the group of students referenced by the estimates. The method was implemented by the ETS proprietary computer program RELCLASS-COMP (Version 4.14).

Reliability of classification at a threshold is estimated by combining the performance levels above a particular threshold and combining the performance levels below that threshold. The result is a two-by-two table indicating whether the students are above or below the threshold. The sum of the entries in the main diagonal is the number of students accurately (or consistently) classified as above or below that threshold.

Table 8.8 and table 8.9 illustrate these two-by-two contingency tables. The proportion of students being accurately classified is determined by summing across the diagonals of the upper tables. The proportion of consistently classified students is determined by summing the diagonals of the lower tables.

Table 8.8 Decision Accuracy for Reaching a Performance Level

|  |  |  |
| --- | --- | --- |
| Performance Level Status | Does Not Reach a Performance Level Based on True Score | Reaches a Performance Level Based on True Score |
| Does not reach a performance level | Correct classification | Incorrect classification |
| Reaches a performance level | Incorrect classification | Correct classification |

Table 8.9 Decision Consistency for Reaching a Performance Level

|  |  |  |
| --- | --- | --- |
| Performance Level Status | Does Not Reach a Performance Level Based on an Alternate Form | Reaches a Performance Level Based on an Alternate Form |
| Does not reach a performance level | Consistent classification | Inconsistent classification |
| Reaches a performance level | Inconsistent classification | Consistent classification |

The results of decision accuracy and consistency at each threshold performance level for the Initial Alternate ELPAC are presented in table 8.10 for all grade levels and grade spans.

Table 8.10 Classification Accuracy and Consistency

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Cut Between Level 1 and Level 2 Accuracy | Cut Between Level 1 and Level 2 Consistency | Cut Between Level 2 and Level 3 Accuracy | Cut Between Level 2 and Level 3 Consistency | Overall Accuracy | Overall Consistency |
| Kindergarten | 0.87 | 0.83 | 0.92 | 0.90 | 0.79 | 0.73 |
| 1 | 0.89 | 0.86 | 0.92 | 0.90 | 0.81 | 0.76 |
| 2 | 0.88 | 0.84 | 0.93 | 0.91 | 0.81 | 0.75 |
| 3–5 | 0.90 | 0.87 | 0.95 | 0.94 | 0.86 | 0.82 |
| 6–8 | 0.82 | 0.82 | 0.93 | 0.91 | 0.75 | 0.73 |
| 9–12 | 0.86 | 0.82 | 0.94 | 0.92 | 0.80 | 0.74 |

At each threshold, the classification at adjacent performance levels has acceptable reliability and consistency. The classification accuracy ranged from 0.82 to 0.90 for the cut between Level 1 and Level 2, and from 0.92 to 0.95 for the cut between Level 2 and Level 3. The classification consistency ranged from 0.82 to 0.87 for the cut between Level 1 and Level 2, and 0.90 to 0.94 for the cut between Level 2 and Level 3. Overall, the accuracy ranged from 0.75 to 0.86 and classification consistency ranged from 0.73 to 0.82. These values, which are limited by small sample sizes at each grade level or grade span, are nevertheless all within acceptable bounds.

### Validity Evidence

Validity refers to the degree to which each interpretation or use of a test score is supported by the accumulated evidence (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014; ETS, 2014). Concerns about validity drive the development, administration, and scoring of an assessment. Validity evidence also determines the appropriateness of test score interpretations and uses.

Validation is the process of accumulating evidence to support each proposed score interpretation or use. This validation process does not rely on a single study or gathering only one type of evidence. Rather, validation involves multiple investigations and different kinds of supporting evidence (AERA, APA, & NCME, 2014; Cronbach, 1971; ETS, 2014; Kane, 2006). It begins with the test design and is implicit throughout the entire assessment process, which includes item development and field testing, analyses of items, standard setting, test scaling and linking, scoring, reporting, and score usage.

In this section, the evidence gathered is presented to support the intended uses and interpretations of scores for the Initial Alternate ELPAC. This section discusses some of the principles prescribed by AERA, APA, and NCME’s *Standards for Educational and Psychological Testing* (2014). These *Standards* require a clear definition of the purpose of the assessment, a description of the constructs to be assessed, and the population to be assessed, as well as how the scores are to be interpreted and used.

The *Standards* identify five kinds of evidence that can provide support for score interpretations and uses:

1. Evidence based on test content
2. Evidence based on relations to other variables
3. Evidence based on response processes
4. Evidence based on internal structure
5. Evidence based on the consequences of testing

The next subsection defines the purpose of the Initial Alternate ELPAC, followed by a description and discussion of different kinds of validity evidence that have been gathered.

#### Design of the Initial Alternate ELPAC

##### Purposes

The Initial Alternate ELPAC is designed to assess students with the most significant cognitive disabilities whose IEP teams have designated the use of an alternate assessment on statewide assessments. The goal of the Initial Alternate ELPAC is to provide information to determine a student’s initial classification as an English learner (EL) or as initial fluent English proficient (IFEP). The goal for the Summative Alternate ELPAC is to provide information on annual student progress toward English language proficiency (ELP) and support decisions on student reclassification as fluent English proficient.

Based on the results of the home language survey (HLS), those students whose primary language is not English take the Initial Alternate ELPAC one time only. Those students who are identified as ELs, as a result of the Initial Alternate ELPAC, take the Summative Alternate ELPAC each year to track their progress until they are reclassified as fluent English proficient.

##### The Constructs to Be Measured

The Initial Alternate ELPAC is designed to align with the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* via the English Language Development Connectors (ELD Connectors), which reduce the depth, breadth, and complexity of the standards, as appropriate for students with the most significant cognitive disabilities. The ELD Connectors were developed through collaboration among California educators, the CDE, and ETS’ research and assessment experts, as well as with guidance from the Alternate ELPAC Test Design Advisory Team of four nationally recognized experts on the assessment of EL students with the most significant cognitive disabilities.

The ELD Connectors represent the highest level of expected performance in ELP for EL students with the most significant cognitive disabilities at a given grade level or grade span. The ELD Connectors are not intended to represent the full range of performance in ELP that may be measured by a standardized ELP assessment.

Test blueprints are used to measure students’ mastery of the ELD Connectors. They also provide an operational definition of the construct to which each set of standards refers and define the following for each content area:

* Subject to be assessed
* Tasks to be presented
* Administration instructions to be given
* Rules used to score student responses

The test blueprints control as many aspects of the measurement procedure as possible so that the testing conditions will remain the same over test administrations (Cronbach, 1971) to minimize construct-irrelevant score variance (Messick, 1989).

ETS developed all Alternate ELPAC test items to conform to the SBE-approved ELD Connectors and test blueprint (CDE, 2019).

##### Interpretations and Uses of the Scores

Overall student performance is expressed as a scale score that is generated for the Initial Alternate ELPAC. The total score is also used to classify students in terms of their performance level, by applying threshold scores resulting from standard setting procedures.

The grade level– and grade span–performance level descriptors (PLDs) describe what students at each performance level know and can do, by grade level or grade span. The PLDs reflect the level of expectation on students’ performance on the contents aligned with ELD Connectors. California educators gathered to develop the grade level– or grade span–range PLDs using the general PLDs, which provided the number of reporting levels and the general definition of each reporting level. The importance of the grade level– or grade span–PLDs is that they define the knowledge or skill expectations at each performance level on a functional basis, define the standards as they apply to threshold scores, and give standardized meaning to scores or score ranges.

A local educational agency may use Initial Alternate ELPAC results to help make decisions about student placement in programs that support the student’s ongoing development toward ELP, student exit from EL programs, and student growth in proficiency while in EL programs. The Initial Alternate 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 students; other relevant information should be considered as well. 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 Initial Alternate ELPAC results.

##### Intended Test Population

The Initial Alternate ELPAC is the required state assessment for ELP that must be administered to students with the most significant cognitive disabilities who are determined by IEP teams to be eligible for an alternate assessment in kindergarten through grade twelve (K–12); and are enrolled in California schools for the first time who are potentially ELs based on a home language other than English, as indicated by the results of an HLS. The Initial Alternate ELPAC is used to identify students as being either an EL or IFEP. The Initial Alternate ELPAC is administered only once during a student’s time in California public schools. The decision to administer the Initial Alternate ELPAC is based on the results of the HLS. The Initial Alternate ELPAC is administered to K–12 students who enrolled in a California public school for the first time. This includes students who enroll in transitional kindergarten, which is the first year of a two-year kindergarten program.

#### Content

Evidence based on test content refers to traditional forms of content validity evidence, such as the rating (scoring) of test specifications and test items (Crocker et al., 1989; Sireci, 1998), as well as alignment methods for educational assessments that evaluate the interactions between curriculum frameworks, testing, and instruction (Rothman et al., 2002; Bhola, Impara, & Buckendahl, 2003; Martone & Sireci, 2009).

[Chapter 3](#_Item_Development_and_1) and [chapter 4](#_Toc122102494) of this technical report describe the procedures for item development and test assembly for the 2022–23 Initial Alternate ELPAC administration and include descriptions of the Alternate ELPAC blueprint, item review process, and procedures to review test forms to ensure appropriate content coverage and psychometric targets.

#### Response Processes

Validity evidence based on response processes refers to “evidence concerning the fit between the construct and the detailed nature of performance or response actually engaged in by students” (AERA et al., 2014, p. 12). This type of evidence generally includes documentation of activities such as

* systematic observations of test response behavior;
* showing the relationships of items intended to require demonstrations or applications of knowledge and skills to other measures that require similar levels of cognitive complexity in the content (i.e., teacher ratings of student performance); and
* evaluation of the reasoning processes students employ when solving test items (Embretson, 1983; Messick, 1989).

##### Observations of Test Response Behavior

An important step in developing the Initial Alternate ELPAC was conducting a cognitive lab methodology study to understand how students interact with the assessment and to ensure that the assessment is accessible to all students in the intended population (CDE, 2020). The final report details all of the changes that were made to the assessment before the operational field test was administered, including steps taken to minimize construct irrelevance, improve face validity, and increase accessibility of items.

##### Threshold Score Validation Study

[Chapter 10](#_In-Test_Survey_3) describes the threshold score validation study that was carried out to verify the appropriateness of the performance level cut scores for the Initial Alternate ELPAC. This study was intended to provide validity evidence for the performance level cut scores for all grade levels or grade spans by showing high levels of agreement between the Initial Alternate ELPAC’s classification of students as ELs and their teachers’ observations of their performance in the classroom. However, because of the very limited sample size of test takers, the study was only able to be carried out for kindergarten students. Results showed that, for 83 percent of kindergarten test takers, their teachers “Agreed” or “Strongly Agreed” with the assessment’s classification of those students as ELs (CDE, 2023c).

#### Internal Structure

Internal structure evidence evaluates the strength or salience of the major dimensions underlying an assessment using dimensionality evaluation, which includes differential item functioning analyses. These analyses were conducted using the 2021–22 Summative Alternate ELPAC operational field test data.

##### Test Dimensionality

Results of the test dimensionality study are summarized in section 8.7 and appendix 8.E of the *Alternate ELPAC 2021–22 Operational Field Test Technical Report* (CDE, 2023b).

Evidence collected from the 2021–22 Alternate ELPAC operational field test data supported the receptive and expressive communication mode performances being reported together as a single Alternate ELPAC test score. This conclusion was based on the results of confirmatory factor analyses using correlated common factor model and bifactor models.

##### Overall Reliability Estimates

The results of reliability and SEMs on the scale score for each assessment are presented in subsection [*8.4.2 Standard Error of Measurement*](#_Standard_Error_of_5). Results indicate that the scores for the Initial Alternate ELPAC are reliable.

##### Student Group Reliability Estimates

The reliabilities are also examined for various student groups. The student groups considered are based on gender, ethnicity, primary disability type, economic status, enrollment in US schools, migrant status, military status, homeless status, and foster youth status. These results are presented in subsection [*8.4.3 Student Group Reliabilities*](#_Student_Group_Reliabilities).

##### Reliability of Performance Classifications

The methodology used for estimating the reliability of classification decisions is described with the decision classification analyses in subsection [*8.4.5 Decision Classification Analyses*](#_Decision_Classification_Analyses).

### 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.

Bhola, D. S., Impara, J. C., & Buckendahl, C. W. (2003). Aligning tests with states’ content standards: Methods and issues. *Educational Measurement: Issues and Practice, 22,* 21–‍29.

California Department of Education. (2019). *ELPAC summative dimensionality report* [Unpublished report]. Sacramento, CA: California Department of Education.

California Department of Education. (2020). *Computer-based* *Summative English Language Proficiency Assessments for California fall 2019 field test technical report* [Unpublished report]. Sacramento, CA: California Department of Education.

California Department of Education. (2023a). *Alternate assessment decision-making tool for California.* Sacramento, CA: California Department of Education.

California Department of Education. (2023b). *Alternate English Language Proficiency Assessments for California 2021–22 operational field test technical report*. Sacramento, CA: California Department of Education.

California Department of Education. (2023c). *Initial Alternate English Language Proficiency Assessments for California threshold score validation study final report*[Unpublished report]. Sacramento, CA: California Department of Education.

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt.

Crocker, L. M., Miller, D., & Franks, E. A. (1989). Quantitative methods for assessing the fit between test and curriculum. *Applied Measurement in Education, 2,* 179–94.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

Educational Testing Service. (2014). *ETS standards for quality and fairness.* Princeton, NJ: Educational Testing Service.

Embretson (Whitley), S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin,* *93*, 179–97.

Feldt, L. S. & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd edition) (pp. 105–46). New York: Macmillan.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist,* 18. 519–32.

Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, *21*(4), 347–60.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Washington, DC: American Council on Education and National Council on Measurement in Education.

Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Washington, DC: American Council on Education/Praeger.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, linking, and scaling: Methods and practices* (2nd ed.). New York, NY: Springer-Verlag.

Livingston, S. A. & Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal of Educational Measurement, 32*, 179–97.

Martone, A., & Sireci, S. G. (2009)*. Evaluating alignment between curriculum, assessments, and instruction. Review of Educational Research, 4,* 1332–61*.*

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149–‍74.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed.). Washington, DC: American Council on Education.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*, 159–76.

Olsson, U., Drasgow, F., & Dorans, N. J. (1982). The polyserial correlation coefficient. *Psychometrika, 47*, 337–47.

Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002).Benchmarking and alignment of standards and testing. [Technical Report 566]. Washington, DC: Center for the Study of Evaluation.

Sireci, S. G. (1998).Gathering and analyzing content validity data*. Educational Assessment, 5,* 299–321.

Stocking, M. L. (1996). An alternative method for scoring adaptive tests. *Journal of Educational and Behavioral Statistics, 21*, 365–89.

### Accessibility Information

#### Alternative Text for Equation 8.1

p value sub dich equals the fraction with the numerator the sum from j equals 1 to J sub i of X sub ij and the denominator J sub i end fraction. *(Return to* [*equation 8.1*](#EQ8_1)*.)*

#### Alternative Text for Equation 8.2

p value sub poly equals the fraction with the numerator the sum from j equals 1 to J sub i of X sub ij and the denominator J sub i times M sub i end fraction. *(Return to* [*equation 8.2*](#EQ8_2)*.)*

#### Alternative Text for Equation 8.3

r sub polyreg equals the fraction beta-hat times s sub tot divided by the square root of beta-hat squared times s squared sub tot plus 1. *(Return to* [*equation 8.3*](#EQ8_3)*.)*

#### Alternative Text for Equation 8.4

p sub ih of theta-hat sub j equals the numerator exp open parenthesis the sum from v equals 1 to h of D times a sub i open parenthesis theta-hat sub j minus b sub i plus d sub iv close parenthesis close parenthesis divided by the denominator open parenthesis 1 plus the sum from c equals 1 to m sub i exp open parenthesis the sum from v equals 1 to c of D times a sub i open parenthesis theta-hat sub j minus b sub i plus d sub iv close parenthesis close parenthesis close parenthesis, if score h equals 1, 2, …, n sub i.

p sub ih of theta-hat sub j equals 1 divided by the denominator open parenthesis 1 plus the sum from c equals 1 to m sub i exp open parenthesis the sum from v equals 1 to c of D times a sub i open parenthesis theta-hat sub j minus b sub i plus d sub iv close parenthesis close parenthesis close parenthesis, if score h equals 0. *(Return to* [*equation 8.4*](#EQ8_4)*.)*

#### Alternative Text for Equation 8.5

epsilon of theta equals the sum from i equals 1 to n sub dich of P sub i of theta plus the sum from j equals 1 to n sub poly over each sum of x equals 1 to m of s sub xj times P sub xj of theta. *(Return to* [*equation 8.5*](#EQ8_5)*.)*

#### Alternative Text for Equation 8.6

SS sub j equals A times theta-hat plus B. *(Return to* [*equation 8.6*](#EQ8_6)*.)*

#### Alternative Text for Equation 8.7

A equals the numerator SS sub Level 3 minus SS sub Level 2 divided by the denominator theta-hat sub Level 3 minus theta-hat sub Level 2. *(Return to* [*equation 8.7*](#EQ8_7)*.)*

#### Alternative Text for Equation 8.8

B equals scale score sub Level 3 minus theta-hat sub Level 3 multiplied by the numerator open parentheses scale score sub Level 3 minus scale score sub Level 2 divided by the denominator theta-hat sub Level 3 minus theta-hat sub Level 2 close parentheses. *(Return to* [*equation 8.8*](#EQ8_8)*.)*

#### Alternative Text for Equation 8.9

rho sub theta-hat prime equals 1 minus the fraction with the numerator sum from j equals 1 to J of CSEM squared sub theta-hat sub j divided by the denominator J times s squared sub theta-hat. *(Return to* [*equation 8.9*](#EQ8_9)*.)*

#### Alternative Text for Equation 8.10

SEM sub scaled equals A times s sub theta-hat times the square root of 1 minus rho sub theta-hat prime. *(Return to* [*equation 8.10*](#EQ8_10)*.)*

#### Alternative Text for Equation 8.11

CSEM of theta-hat sub j equals 1 divided by the square root of I of theta sub j. *(Return to* [*equation 8.11*](#EQ8_11)*.)*

#### Alternative Text for Equation 8.12

I of theta-hat sub j equals the sum from i equals 1 to I of I sub i of theta-hat sub j. *(Return to* [*equation 8.12*](#EQ8_12)*.)*

#### Alternative Text for Equation 8.13

I sub i of theta-hat sub j equals open bracket s sub i2 of theta-hat sub j minus s sub i squared of theta-hat sub j. *(Return to* [*equation 8.13*](#EQ8_13)*.)*

#### Alternative Text for Equation 8.14

s sub i of theta-hat sub j equals the sum from h equals 0 to M sub i of h times p sub ih of theta-hat sub j. *(Return to* [*equation 8.14*](#EQ8_14)*.)*

#### Alternative Text for Equation 8.15

s sub i2 of theta-hat sub j equals the sum from h equals 0 to M sub I of h squared times p sub ih of theta-hat sub j. *(Return to* [*equation 8.15*](#EQ8_15)*.)*

#### Alternative Text for Equation 8.16

p sub ih of theta-hat sub j equals the numerator exp open parenthesis the sum from v equals 1 to h of D times a sub i open parenthesis theta-hat sub j minus b sub i plus d sub iv close parenthesis close parenthesis divided by the denominator open parenthesis 1 plus the sum from c equals 1 to M sub i exp open parenthesis the sum from v equals 1 to c of D times a sub i open parenthesis theta-hat sub j minus b sub i plus d sub iv close parenthesis close parenthesis close parenthesis, if score h equals 1, 2, …, M sub i.

p sub ih of theta-hat sub j equals 1 divided by the denominator open parenthesis 1 plus the sum from c equals 1 to M sub i exp open parenthesis the sum from v equals 1 to c of D times a sub i open parenthesis theta sub j minus b sub I plus d sub iv close parenthesis close parenthesis close parenthesis, if score h equals 0. *(Return to* [*equation 8.16*](#EQ8_16)*.)*

#### Alternative Text for Equation 8.17

CSEM of SS sub j equals A times CSEM of theta-hat sub j. *(Return to* [*equation 8.17*](#EQ8_17)*.)*

### Appendix 8.A: Classical Item Analyses Results

**Notes:**

* In table 8.A.1 through table 8.A.6,
* D = dichotomous item, and
* P = polytomous item.

Table 8.A.1 Classical Item Statistics for Kindergarten

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR052871 | D | 656 | 0.27 | 0.58 | 29 | 73 | 27 | N/A |
| VR131310 | D | 656 | 0.13 | 0.49 | 27 | 87 | 13 | N/A |
| VR131342 | D | 656 | 0.35 | 0.64 | 26 | 65 | 35 | N/A |
| VR131450 | D | 656 | 0.43 | 0.69 | 29 | 57 | 43 | N/A |
| VR131594 | D | 656 | 0.22 | 0.51 | 29 | 78 | 22 | N/A |
| VR137085 | D | 656 | 0.64 | 0.69 | 18 | 36 | 64 | N/A |
| VR138924 | D | 656 | 0.47 | 0.64 | 24 | 53 | 47 | N/A |
| VR138950 | D | 656 | 0.19 | 0.57 | 37 | 81 | 19 | N/A |
| VR138982 | D | 656 | 0.47 | 0.73 | 25 | 53 | 47 | N/A |
| VR139666 | D | 656 | 0.35 | 0.66 | 31 | 65 | 35 | N/A |
| VR139673 | D | 656 | 0.39 | 0.66 | 28 | 61 | 39 | N/A |
| VR139729 | D | 656 | 0.29 | 0.62 | 31 | 71 | 29 | N/A |
| VR154406 | D | 656 | 0.42 | 0.76 | 34 | 58 | 42 | N/A |
| VR154449 | D | 656 | 0.29 | 0.62 | 28 | 71 | 29 | N/A |
| VR170322 | D | 656 | 0.29 | 0.70 | 32 | 71 | 29 | N/A |
| VR193093 | D | 656 | 0.58 | 0.59 | 16 | 42 | 58 | N/A |
| VR244385 | D | 656 | 0.27 | 0.61 | 30 | 73 | 27 | N/A |
| VR139022 | P | 656 | 0.68 | 0.74 | 12 | 57 | 17 | 25 |
| VR139973 | P | 656 | 0.77 | 0.78 | 11 | 54 | 15 | 31 |
| VR154458 | P | 656 | 0.50 | 0.74 | 19 | 66 | 18 | 16 |
| VR154465 | P | 656 | 1.01 | 0.82 | 15 | 42 | 16 | 43 |
| VR193113 | P | 656 | 0.84 | 0.73 | 10 | 49 | 18 | 33 |
| VR215978 | P | 656 | 0.57 | 0.77 | 17 | 63 | 18 | 19 |
| VR223164 | P | 656 | 0.38 | 0.74 | 22 | 73 | 16 | 11 |

Table 8.A.2 Classical Item Statistics for Grade One

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR130319 | D | 57 | 0.37 | 0.79 | 33 | 63 | 37 | N/A |
| VR130326 | D | 57 | 0.28 | 0.57 | 33 | 72 | 28 | N/A |
| VR130331 | D | 57 | 0.28 | 0.75 | 37 | 72 | 28 | N/A |
| VR130345 | D | 57 | 0.32 | 0.44 | 33 | 68 | 32 | N/A |
| VR133917 | D | 57 | 0.32 | 0.80 | 30 | 68 | 32 | N/A |
| VR133975 | D | 57 | 0.23 | 0.68 | 33 | 77 | 23 | N/A |
| VR134007 | D | 57 | 0.47 | 0.64 | 30 | 53 | 47 | N/A |
| VR137615 | D | 57 | 0.49 | 0.58 | 25 | 51 | 49 | N/A |
| VR138495 | D | 57 | 0.53 | 0.73 | 25 | 47 | 53 | N/A |
| VR138505 | D | 57 | 0.49 | 0.61 | 23 | 51 | 49 | N/A |
| VR138567 | D | 57 | 0.51 | 0.85 | 33 | 49 | 51 | N/A |
| VR138628 | D | 57 | 0.28 | 0.64 | 40 | 72 | 28 | N/A |
| VR150660 | D | 57 | 0.33 | 0.71 | 33 | 67 | 33 | N/A |
| VR150685 | D | 57 | 0.28 | 0.59 | 40 | 72 | 28 | N/A |
| VR150709 | D | 57 | 0.32 | 0.71 | 40 | 68 | 32 | N/A |
| VR154742 | D | 57 | 0.28 | 0.69 | 32 | 72 | 28 | N/A |
| VR154751 | D | 57 | 0.25 | 0.58 | 39 | 75 | 25 | N/A |
| VR193603 | D | 57 | 0.56 | 0.71 | 25 | 44 | 56 | N/A |
| VR133983 | P | 57 | 0.63 | 0.83 | 25 | 65 | 7 | 28 |
| VR137618 | P | 57 | 0.68 | 0.73 | 19 | 61 | 9 | 30 |
| VR150707 | P | 57 | 0.60 | 0.89 | 28 | 67 | 7 | 26 |
| VR154753 | P | 57 | 0.51 | 0.79 | 28 | 68 | 12 | 19 |
| VR154755 | P | 57 | 0.72 | 0.78 | 26 | 54 | 19 | 26 |
| VR193651 | P | 57 | 0.91 | 0.77 | 12 | 51 | 7 | 42 |

Table 8.A.3 Classical Item Statistics for Grade Two

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR130125 | D | 20 | 0.40 | 0.65 | 25 | 60 | 40 | N/A |
| VR130141 | D | 20 | 0.30 | 0.96 | 35 | 70 | 30 | N/A |
| VR130233 | D | 20 | 0.30 | 0.68 | 40 | 70 | 30 | N/A |
| VR130277 | D | 20 | 0.30 | 0.77 | 30 | 70 | 30 | N/A |
| VR134637 | D | 20 | 0.45 | 0.72 | 40 | 55 | 45 | N/A |
| VR134649 | D | 20 | 0.40 | 0.53 | 35 | 60 | 40 | N/A |
| VR134668 | D | 20 | 0.30 | 0.96 | 35 | 70 | 30 | N/A |
| VR140204 | D | 20 | 0.75 | 0.66 | 10 | 25 | 75 | N/A |
| VR140495 | D | 20 | 0.50 | 0.79 | 20 | 50 | 50 | N/A |
| VR140498 | D | 20 | 0.40 | 0.78 | 30 | 60 | 40 | N/A |
| VR140501 | D | 20 | 0.60 | 0.97 | 30 | 40 | 60 | N/A |
| VR140520 | D | 20 | 0.45 | 0.82 | 40 | 55 | 45 | N/A |
| VR151565 | D | 20 | 0.20 | 0.64 | 45 | 80 | 20 | N/A |
| VR151573 | D | 20 | 0.15 | 0.93 | 40 | 85 | 15 | N/A |
| VR151643 | D | 20 | 0.10 | −0.11 | 45 | 90 | 10 | N/A |
| VR155513 | D | 20 | 0.30 | 0.84 | 45 | 70 | 30 | N/A |
| VR155674 | D | 20 | 0.40 | 0.67 | 40 | 60 | 40 | N/A |
| VR193873 | D | 20 | 0.65 | 0.55 | 10 | 35 | 65 | N/A |
| VR134677 | P | 20 | 0.55 | 0.69 | 20 | 70 | 5 | 25 |
| VR140209 | P | 20 | 0.90 | 1.00 | 10 | 50 | 10 | 40 |
| VR151624 | P | 20 | 0.40 | 0.47 | 30 | 70 | 20 | 10 |
| VR155670 | P | 20 | 0.65 | 0.69 | 20 | 45 | 45 | 10 |
| VR193885 | P | 20 | 1.30 | 0.72 | 5 | 25 | 20 | 55 |
| VR223063 | P | 20 | 0.50 | 0.63 | 25 | 60 | 30 | 10 |

Table 8.A.4 Classical Item Statistics for Grade Span Three Through Five

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR060298 | D | 39 | 0.33 | 0.63 | 33 | 67 | 33 | N/A |
| VR131591 | D | 39 | 0.13 | 0.60 | 31 | 87 | 13 | N/A |
| VR131622 | D | 39 | 0.38 | 0.34 | 28 | 62 | 38 | N/A |
| VR131627 | D | 39 | 0.23 | 0.64 | 38 | 77 | 23 | N/A |
| VR131628 | D | 39 | 0.31 | 0.78 | 38 | 69 | 31 | N/A |
| VR140200 | D | 39 | 0.26 | 0.64 | 51 | 74 | 26 | N/A |
| VR140214 | D | 39 | 0.28 | 0.72 | 36 | 72 | 28 | N/A |
| VR140221 | D | 39 | 0.15 | 0.98 | 46 | 85 | 15 | N/A |
| VR144676 | D | 39 | 0.38 | 0.53 | 28 | 62 | 38 | N/A |
| VR145701 | D | 39 | 0.46 | 0.80 | 36 | 54 | 46 | N/A |
| VR145817 | D | 39 | 0.38 | 0.71 | 28 | 62 | 38 | N/A |
| VR145916 | D | 39 | 0.31 | 0.44 | 33 | 69 | 31 | N/A |
| VR146024 | D | 39 | 0.28 | 0.72 | 46 | 72 | 28 | N/A |
| VR150967 | D | 39 | 0.18 | 0.54 | 33 | 82 | 18 | N/A |
| VR150970 | D | 39 | 0.15 | 0.72 | 46 | 85 | 15 | N/A |
| VR151004 | D | 39 | 0.15 | 0.84 | 59 | 85 | 15 | N/A |
| VR194284 | D | 39 | 0.49 | 0.61 | 23 | 51 | 49 | N/A |
| VR223852 | D | 39 | 0.15 | 0.57 | 41 | 85 | 15 | N/A |
| VR140236 | P | 39 | 0.36 | 0.75 | 23 | 74 | 15 | 10 |
| VR144686 | P | 39 | 0.67 | 0.81 | 15 | 59 | 15 | 26 |
| VR150995 | P | 39 | 0.54 | 0.83 | 28 | 69 | 8 | 23 |
| VR194301 | P | 39 | 0.38 | 0.90 | 21 | 77 | 8 | 15 |
| VR222572 | P | 39 | 0.56 | 0.78 | 28 | 62 | 21 | 18 |
| VR222573 | P | 39 | 0.38 | 0.84 | 26 | 74 | 13 | 13 |

Table 8.A.5 Classical Item Statistics for Grade Span Six Through Eight

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR132196 | D | 22 | 0.36 | 0.77 | 32 | 64 | 36 | N/A |
| VR132209 | D | 22 | 0.36 | 0.61 | 32 | 64 | 36 | N/A |
| VR132300 | D | 22 | 0.18 | 0.69 | 32 | 82 | 18 | N/A |
| VR132305 | D | 22 | 0.36 | 0.60 | 32 | 64 | 36 | N/A |
| VR133740 | D | 22 | 0.55 | 0.93 | 27 | 45 | 55 | N/A |
| VR133759 | D | 22 | 0.32 | 0.53 | 27 | 68 | 32 | N/A |
| VR133797 | D | 22 | 0.41 | 0.97 | 36 | 59 | 41 | N/A |
| VR147461 | D | 22 | 0.59 | 0.91 | 36 | 41 | 59 | N/A |
| VR148853 | D | 22 | 0.36 | 0.78 | 36 | 64 | 36 | N/A |
| VR148858 | D | 22 | 0.18 | 0.90 | 36 | 82 | 18 | N/A |
| VR148916 | D | 22 | 0.23 | 0.98 | 41 | 77 | 23 | N/A |
| VR150186 | D | 22 | 0.41 | 0.75 | 36 | 59 | 41 | N/A |
| VR150187 | D | 22 | 0.23 | 0.78 | 36 | 77 | 23 | N/A |
| VR150188 | D | 22 | 0.55 | 0.91 | 36 | 45 | 55 | N/A |
| VR150189 | D | 22 | 0.27 | 0.69 | 36 | 73 | 27 | N/A |
| VR166709 | D | 22 | 0.32 | 0.92 | 36 | 68 | 32 | N/A |
| VR167935 | D | 22 | 0.27 | 0.87 | 36 | 73 | 27 | N/A |
| VR196656 | D | 22 | 0.73 | 0.94 | 23 | 27 | 73 | N/A |
| VR133811 | P | 22 | 0.73 | 0.82 | 27 | 55 | 18 | 27 |
| VR147469 | P | 22 | 1.00 | 0.78 | 14 | 45 | 9 | 45 |
| VR148864 | P | 22 | 0.82 | 0.71 | 27 | 55 | 9 | 36 |
| VR167959 | P | 22 | 0.64 | 0.80 | 27 | 64 | 9 | 27 |
| VR167974 | P | 22 | 0.95 | 0.87 | 27 | 32 | 41 | 27 |
| VR196675 | P | 22 | 1.23 | 0.85 | 9 | 32 | 14 | 55 |

Table 8.A.6 Classical Item Statistics for Grade Span Nine Through Twelve

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 |
| VR132681 | D | 25 | 0.40 | 0.64 | 8 | 60 | 40 | N/A |
| VR132695 | D | 25 | 0.48 | 0.62 | 8 | 52 | 48 | N/A |
| VR132848 | D | 25 | 0.48 | 0.65 | 12 | 52 | 48 | N/A |
| VR133507 | D | 25 | 0.28 | 0.56 | 12 | 72 | 28 | N/A |
| VR133887 | D | 25 | 0.60 | 0.54 | 8 | 40 | 60 | N/A |
| VR133915 | D | 25 | 0.44 | 0.86 | 8 | 56 | 44 | N/A |
| VR134023 | D | 25 | 0.52 | 0.69 | 8 | 48 | 52 | N/A |
| VR144835 | D | 25 | 0.72 | 0.71 | 4 | 28 | 72 | N/A |
| VR147932 | D | 25 | 0.60 | 0.70 | 4 | 40 | 60 | N/A |
| VR148029 | D | 25 | 0.44 | 0.64 | 4 | 56 | 44 | N/A |
| VR148031 | D | 25 | 0.52 | 0.41 | 4 | 48 | 52 | N/A |
| VR148050 | D | 25 | 0.40 | 0.36 | 12 | 60 | 40 | N/A |
| VR150493 | D | 25 | 0.44 | 0.65 | 12 | 56 | 44 | N/A |
| VR150497 | D | 25 | 0.52 | 0.83 | 12 | 48 | 52 | N/A |
| VR150530 | D | 25 | 0.20 | 0.64 | 12 | 80 | 20 | N/A |
| VR154631 | D | 25 | 0.48 | 0.76 | 8 | 52 | 48 | N/A |
| VR154835 | D | 25 | 0.28 | 0.43 | 8 | 72 | 28 | N/A |
| VR191181 | D | 25 | 0.68 | 0.44 | 4 | 32 | 68 | N/A |
| VR132823 | P | 25 | 0.68 | 0.68 | 4 | 56 | 20 | 24 |
| VR144875 | P | 25 | 1.56 | 0.53 | 0 | 12 | 20 | 68 |
| VR150525 | P | 25 | 0.32 | 0.74 | 4 | 76 | 16 | 8 |
| VR154860 | P | 25 | 1.12 | 0.80 | 4 | 24 | 40 | 36 |
| VR154926 | P | 25 | 0.84 | 0.77 | 4 | 48 | 20 | 32 |
| VR191268 | P | 25 | 0.96 | 0.53 | 0 | 40 | 24 | 36 |

### Appendix 8.B: Item Response Theory Results

**Note:** In table 8.B.1 through table 8.B.6, “N/A” indicates that these items did not have *d*-‍parameter estimates, and SE = standard error.

Table 8.B.1 IRT Item Statistics, Kindergarten

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR052871 | 1.49 | 0.12 | N/A | N/A |
| VR131310 | 2.51 | 0.14 | N/A | N/A |
| VR131342 | 0.62 | 0.11 | N/A | N/A |
| VR131450 | 0.60 | 0.11 | N/A | N/A |
| VR131594 | 1.74 | 0.12 | N/A | N/A |
| VR137085 | −0.94 | 0.12 | N/A | N/A |
| VR138924 | 0.03 | 0.11 | N/A | N/A |
| VR138950 | 2.10 | 0.13 | N/A | N/A |
| VR138982 | −0.18 | 0.11 | N/A | N/A |
| VR139022 | 0.64 | 0.09 | −0.62 | 0.62 |
| VR139666 | 0.33 | 0.12 | N/A | N/A |
| VR139673 | 0.38 | 0.11 | N/A | N/A |
| VR139729 | 1.23 | 0.12 | N/A | N/A |
| VR139973 | 0.21 | 0.09 | −0.89 | 0.89 |
| VR154406 | 0.01 | 0.12 | N/A | N/A |
| VR154449 | 0.95 | 0.11 | N/A | N/A |
| VR154458 | 1.06 | 0.09 | −0.23 | 0.23 |
| VR154465 | −0.31 | 0.09 | −0.66 | 0.66 |
| VR170322 | 0.76 | 0.12 | N/A | N/A |
| VR193093 | −0.81 | 0.12 | N/A | N/A |
| VR193113 | 0.05 | 0.09 | −0.11 | 0.11 |
| VR215978 | 0.66 | 0.09 | −0.46 | 0.46 |
| VR223164 | 1.28 | 0.09 | −0.22 | 0.22 |
| VR244385 | 1.23 | 0.12 | N/A | N/A |

Table 8.B.2 IRT Item Statistics, Grade One

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR130319 | −0.36 | 0.10 | N/A | N/A |
| VR130326 | 0.53 | 0.10 | N/A | N/A |
| VR130331 | 0.11 | 0.10 | N/A | N/A |
| VR130345 | 0.09 | 0.10 | N/A | N/A |
| VR133917 | −0.07 | 0.10 | N/A | N/A |
| VR133975 | 1.18 | 0.11 | N/A | N/A |
| VR133983 | 0.68 | 0.08 | −0.32 | 0.32 |
| VR134007 | −0.65 | 0.11 | N/A | N/A |
| VR137615 | −0.69 | 0.11 | N/A | N/A |
| VR137618 | 0.36 | 0.08 | −0.66 | 0.66 |
| VR138495 | −0.24 | 0.10 | N/A | N/A |
| VR138505 | −0.68 | 0.11 | N/A | N/A |
| VR138567 | −0.84 | 0.11 | N/A | N/A |
| VR138628 | 0.07 | 0.10 | N/A | N/A |
| VR150660 | 0.25 | 0.10 | N/A | N/A |
| VR150685 | 1.00 | 0.10 | N/A | N/A |
| VR150707 | 0.50 | 0.08 | −0.65 | 0.65 |
| VR150709 | 0.76 | 0.10 | N/A | N/A |
| VR154742 | 0.49 | 0.10 | N/A | N/A |
| VR154751 | 0.41 | 0.10 | N/A | N/A |
| VR154753 | 0.80 | 0.08 | −0.17 | 0.17 |
| VR154755 | 0.27 | 0.08 | 0.57 | −0.57 |
| VR193603 | −1.47 | 0.12 | N/A | N/A |
| VR193651 | −0.70 | 0.08 | −0.83 | 0.83 |

Table 8.B.3 IRT Item Statistics, Grade Two

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR130125 | 0.39 | 0.11 | N/A | N/A |
| VR130141 | 0.10 | 0.11 | N/A | N/A |
| VR130233 | −0.22 | 0.11 | N/A | N/A |
| VR130277 | −0.20 | 0.11 | N/A | N/A |
| VR134637 | 0.04 | 0.11 | N/A | N/A |
| VR134649 | −0.10 | 0.11 | N/A | N/A |
| VR134668 | 0.78 | 0.11 | N/A | N/A |
| VR134677 | 0.24 | 0.08 | −0.54 | 0.54 |
| VR140204 | −0.89 | 0.11 | N/A | N/A |
| VR140209 | −0.65 | 0.08 | −1.11 | 1.11 |
| VR140495 | −0.30 | 0.11 | N/A | N/A |
| VR140498 | 0.38 | 0.11 | N/A | N/A |
| VR140501 | −1.15 | 0.12 | N/A | N/A |
| VR140520 | −0.42 | 0.11 | N/A | N/A |
| VR151565 | 0.61 | 0.11 | N/A | N/A |
| VR151573 | 0.38 | 0.11 | N/A | N/A |
| VR151624 | 0.69 | 0.08 | −0.37 | 0.37 |
| VR151643 | 1.76 | 0.12 | N/A | N/A |
| VR155513 | 0.00 | 0.11 | N/A | N/A |
| VR155670 | 0.15 | 0.09 | 0.80 | −0.80 |
| VR155674 | 0.87 | 0.11 | N/A | N/A |
| VR193873 | −1.05 | 0.11 | N/A | N/A |
| VR193885 | −0.87 | 0.09 | −0.96 | 0.96 |
| VR223063 | 0.03 | 0.09 | 0.73 | −0.73 |

Table 8.B.4 IRT Item Statistics, Grade Span Three Through Five

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR060298 | −0.34 | 0.07 | N/A | N/A |
| VR131591 | −0.29 | 0.06 | N/A | N/A |
| VR131622 | −0.29 | 0.06 | N/A | N/A |
| VR131627 | 0.56 | 0.06 | N/A | N/A |
| VR131628 | −1.03 | 0.06 | N/A | N/A |
| VR140200 | −0.88 | 0.06 | N/A | N/A |
| VR140214 | −0.27 | 0.06 | N/A | N/A |
| VR140221 | −0.44 | 0.06 | N/A | N/A |
| VR140236 | 0.24 | 0.04 | −0.31 | 0.31 |
| VR144676 | −1.48 | 0.06 | N/A | N/A |
| VR144686 | −0.87 | 0.04 | −1.16 | 1.16 |
| VR145701 | −1.86 | 0.09 | N/A | N/A |
| VR145817 | −0.71 | 0.08 | N/A | N/A |
| VR145916 | −0.81 | 0.07 | N/A | N/A |
| VR146024 | −1.00 | 0.08 | N/A | N/A |
| VR150967 | −0.02 | 0.07 | N/A | N/A |
| VR150970 | −0.30 | 0.08 | N/A | N/A |
| VR150995 | −0.37 | 0.05 | −0.88 | 0.88 |
| VR151004 | 0.71 | 0.07 | N/A | N/A |
| VR194284 | −1.54 | 0.08 | N/A | N/A |
| VR194301 | −0.61 | 0.05 | −0.58 | 0.58 |
| VR222572 | −0.62 | 0.06 | 0.42 | −0.42 |
| VR222573 | 0.21 | 0.05 | −0.58 | 0.58 |
| VR223852 | −0.59 | 0.08 | N/A | N/A |

Table 8.B.5 IRT Item Statistics, Grade Span Six Through Eight

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR132196 | −1.28 | 0.08 | N/A | N/A |
| VR132209 | −1.64 | 0.08 | N/A | N/A |
| VR132300 | 0.19 | 0.07 | N/A | N/A |
| VR132305 | −1.09 | 0.08 | N/A | N/A |
| VR133740 | −1.71 | 0.08 | N/A | N/A |
| VR133759 | −1.28 | 0.07 | N/A | N/A |
| VR133797 | −1.54 | 0.08 | N/A | N/A |
| VR133811 | −0.56 | 0.06 | 0.22 | −0.22 |
| VR147461 | −2.31 | 0.08 | N/A | N/A |
| VR147469 | −1.93 | 0.06 | −0.89 | 0.89 |
| VR148853 | −0.79 | 0.09 | N/A | N/A |
| VR148858 | −0.10 | 0.09 | N/A | N/A |
| VR148864 | −0.59 | 0.06 | −0.47 | 0.47 |
| VR148916 | 0.35 | 0.08 | N/A | N/A |
| VR150186 | −1.81 | 0.10 | N/A | N/A |
| VR150187 | −0.91 | 0.09 | N/A | N/A |
| VR150188 | −1.49 | 0.10 | N/A | N/A |
| VR150189 | −1.01 | 0.10 | N/A | N/A |
| VR166709 | −0.35 | 0.08 | N/A | N/A |
| VR167935 | −1.07 | 0.09 | N/A | N/A |
| VR167959 | 0.28 | 0.06 | 0.28 | −0.28 |
| VR167974 | −0.94 | 0.07 | 0.83 | −0.83 |
| VR196656 | −3.45 | 0.10 | N/A | N/A |
| VR196675 | −1.93 | 0.06 | −0.79 | 0.79 |

Table 8.B.6 IRT Item Statistics, Grade Span Nine Through Twelve

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | D1 | D2 |
| VR132681 | −0.98 | 0.06 | N/A | N/A |
| VR132695 | −1.47 | 0.06 | N/A | N/A |
| VR132823 | −0.60 | 0.05 | −0.44 | 0.44 |
| VR132848 | −1.69 | 0.06 | N/A | N/A |
| VR133507 | −0.79 | 0.06 | N/A | N/A |
| VR133887 | −1.60 | 0.06 | N/A | N/A |
| VR133915 | −1.35 | 0.06 | N/A | N/A |
| VR134023 | −1.89 | 0.07 | N/A | N/A |
| VR144835 | −3.08 | 0.11 | N/A | N/A |
| VR144875 | −2.16 | 0.06 | −1.27 | 1.27 |
| VR147932 | −1.37 | 0.09 | N/A | N/A |
| VR148029 | −1.26 | 0.09 | N/A | N/A |
| VR148031 | −1.95 | 0.09 | N/A | N/A |
| VR148050 | −0.89 | 0.08 | N/A | N/A |
| VR150493 | −1.72 | 0.09 | N/A | N/A |
| VR150497 | −0.99 | 0.08 | N/A | N/A |
| VR150525 | 0.18 | 0.05 | 0.10 | −0.10 |
| VR150530 | 0.90 | 0.07 | N/A | N/A |
| VR154631 | −1.04 | 0.09 | N/A | N/A |
| VR154835 | −0.45 | 0.07 | N/A | N/A |
| VR154860 | −1.11 | 0.06 | 0.60 | −0.60 |
| VR154926 | 0.01 | 0.06 | 0.36 | −0.36 |
| VR191181 | −3.20 | 0.12 | N/A | N/A |
| VR191268 | −1.69 | 0.06 | −0.48 | 0.48 |

## Quality Control

The California Department of Education (CDE) and ETS implemented rigorous quality-control procedures throughout the assessment development, administration, scoring, analyses, and reporting processes for the Initial Alternate English Language Proficiency Assessments for California (ELPAC). As part of this effort, ETS’ staff worked with its 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 evaluating those products and services. Quality-control procedures are outlined in this chapter.

### Quality Control of Item Development

ETS’ goal is to provide the best standards-based and innovative items for the Initial Alternate ELPAC. Items developed for the Initial Alternate ELPAC were subject to an extensive item review process. The item writers responsible for developing Initial Alternate ELPAC items were trained in ELPAC and ETS’ policies on quality control of item content, bias and sensitivity guidelines, as well as guidelines for accessibility, to ensure that the items allow the widest possible range of students to demonstrate their abilities.

Once a draft item was accepted for authoring—that is, once it was entered into ETS’ item bank and formatted for use in an assessment—ETS employed a series of internal reviews and an initial CDE review. These reviews used established criteria and specifications to judge the quality of an item’s content and ensured that each item measured what it was intended to measure. These reviews also examined the overall quality of the test items before presentation to the CDE and item reviewers. To finish the process, a group of California educators reviewed the items for accessibility, bias and sensitivity, and content, and made recommendations for item enhancement. The details on quality control of item development are described in section [*3.2 ETS Item Review Process*](#_ETS_Item_Review_2).

During administrations of the Initial Alternate ELPAC, when sufficient student response data on each item became available, ETS’ Psychometric Analysis & Research (PAR) staff conducted item analyses and a key check to examine whether the items performed as expected. ETS’ psychometric staff conducted a thorough evaluation of all item statistics to flag items that were potentially problematic because of poor item performance, content issues, item bias, or accessibility challenges. Flagged items were then reviewed by ETS’ Assessment & Learning Technology Development (ALTD) staff, the CDE, and California educators to determine whether issues existed.

### Quality Control of Test Assembly and Delivery

The assembly of all test forms must conform to blueprints that represent a set of constraints and specifications. ETS conducted multiple levels of quality assurance (QA) checks on each assembled Initial Alternate ELPAC form to ensure it met the form-building specifications. Both ETS’ ALTD and PAR staff reviewed and signed off on the accuracy of forms before the test forms were posted for CDE review. Detailed information related to test assembly can be found in [*Chapter 4: Test Assembly*](#_Toc122102494).

In particular, the assembly of all test forms went through a certification process that involved various checks, including verifying that

* all item answers in the key were correctly identified and documented in the scoring system;
* items were scored correctly in the item bank and incorrect responses were scored as incorrect;
* all items assessed the intended standard;
* all content in the item was correct with the exception of distractors, which are intended to be incorrect;
* all items met the statistical criteria, to the extent possible;
* distractors were plausible;
* multiple-choice, single-select (MCSS) item options were parallel in structure;
* language was grade-level appropriate;
* no more than three MCSS items in a row had the same key;
* all graphics were correct (copyright, spelling, relevance, etc.);
* there were no unintended mechanical errors in grammar, spelling, punctuation, and the like; and
* items adhered to the approved style guide.

Reviews were also conducted for functionality and sequencing during the user acceptance testing (UAT) process to ensure all items functioned as expected. Three cycles of UAT were conducted: the first by the test delivery system (TDS) vendor, the second by ETS, and the third by the CDE. CDE staff made a final quality check to ensure that all issues identified during UAT were resolved before the release of the operational assessment.

#### Quality Control of Test Assignment

Test assignment for the ELPAC, including the Initial Alternate ELPAC, is controlled by the Test Operations Management System (TOMS), using student demographic information received from the California Longitudinal Pupil Achievement Data System (CALPADS) (CDE, 2023b). The two systems are kept in sync during the testing window.

Eligible students were assigned to the general Initial ELPAC by default. For students eligible for the Initial Alternate ELPAC, local educational agencies (LEAs) logged on to TOMS and assigned students to take the alternate assessment, which automatically unassigned those students from taking the general Initial ELPAC. The CDE provided guidance to support LEAs in determining which students were eligible for the Initial Alternate ELPAC (CDE, 2023a).

The quality of test assignment for the Initial Alternate ELPAC was monitored and controlled through several strategies. TOMS enforced preconditions for eligibility for the Initial Alternate ELPAC by permitting assignment only for students with an Individuals with Disabilities Education Act (IDEA)11F[[4]](#footnote-5) indicator of “Yes” as sent by CALPADS.

TOMS prevented the prohibited “mixing and matching” of assessments—a student assigned to take an alternate assessment was automatically prevented from assignment to a general assessment. Additionally, students assigned to the Summative Alternate ELPAC would also automatically be assigned the California Alternate Assessments and vice versa.

### Quality Control of Test Materials

Brief descriptions of the types of materials used for and during testing appear in the following subsections.

#### Test Administration Manuals

ETS’ staff verified that test instruction manuals accurately matched the test materials and testing processes. Editors reviewed each document for spelling, grammar, accuracy, and adherence to CDE style. Each document was approved by the CDE before being published to the ELPAC website. Only nonsecure documents were posted to this website. Secure materials, such as the *Initial Alternate ELPAC Directions for Administration*, were made available to designated LEA staff through TOMS, which required a secure logon.

The manuals used in the administration of the Initial Alternate ELPAC are listed in subsection [*5.3.4 Instructions for Test Administration*](#_Instructions_for_Test_2).

#### Processing Test Materials

The ways in which materials associated with student testing were processed are described in subsequent subsections.

##### Computer-based Assessments

Computer-based tests submitted by students were transmitted from Cambium Assessment, Inc. (CAI) to ETS each day. Each system checked for the completeness of the student record and stopped records that were identified as having an error. (For example, the system would identify a test part that was missing a content registration ID, a unique identifier that matches the student’s opportunities.)

### Quality Control of Test Administration

The quality of test administration for the Initial Alternate ELPAC was monitored and controlled through several strategies.

A fully supported Outreach team that includes California Technical Assistance Center phone support and Success Agents supported all LEAs in the administration of the ELPAC. In addition to providing guidance and answering questions, the Outreach team regularly conducted campaigns on particular administration topics to ensure all LEAs understood correct test administration procedures. Outreach was guided by individuals who managed communications to LEAs; provided regional and web-based trainings; and hosted a website, [the](https://www.caaspp.org/) ELPAC website, that housed a full range of manuals, videos, and other instructional and support materials.

The quality of test administration was further managed through comprehensive rules and guidelines for maintaining the security and standardization of the ELPAC. LEAs received training on these topics and were provided tools for reporting security incidents and resolving testing discrepancies for specific testing sessions.

The ETS Office of Testing Integrity (OTI) reinforced the quality-control procedures for test administration, providing QA services for all testing programs managed by ETS. The detailed procedures the OTI developed and applied in quality control are described in subsection [*5.7.1 ETS’ Office of Testing Integrity*](#_ETS’_Office_of_2).

### Quality Control of Scoring

ETS conforms to high standards of quality and fairness when scoring assessments and reporting scores. These standards dictate that ETS provides accurate and understandable assessment results to the intended recipients. It is also ETS’ mission to provide appropriate guidelines for score interpretation and cautions about the limitations in the meaning and use of the test scores. Finally, ETS conducts analyses needed to ensure that the assessments are equitable for various demographic student groups.

#### Machine-Scoring Procedures

To ensure valid item-level scoring for the Initial Alternate ELPAC, quality-control procedures were employed by CAI, the ELPAC subcontractor responsible for providing the TDS and scoring machine-scorable items. CAI staff independently reviewed all Initial Alternate ELPAC forms by producing sample results for assessments. The sample results were compared with the answer keys for each form to confirm the accuracy of scoring keys. The scores for all applicable items were recorded. A final comparison of the test map to each computer-based form as configured in the UAT environment ensured that no changes to the form were introduced prior to operational deployment.

A real-time, quality-monitoring component was built into the TDS. After an assessment was administered to a student, the TDS passed the resulting data to the QA system. QA conducted a series of data integrity checks, ensuring, for example, that the record for each assessment contained information for each item, keys for MCSS items, score points in each item, and the total number of operational items. In addition, QA also checked to ensure that the test record contained no data from items that might have been invalidated.

Data passed directly from the Quality Monitoring System to the database of record, which served as the repository for all test information, and from which all test information was pulled and transmitted to ETS in a predetermined results format.

#### Rubric-Scored Item Scoring

The rubric-scored items in the Initial Alternate ELPAC were scored by local test examiners when they administered the assessment. Every LEA that had an eligible Initial Alternate ELPAC student in California was required to either complete the online LEA Certification course on the Moodle Training Site, Initial Alternate ELPAC Administration and Scoring Training, or coordinate with another certified LEA via a Memorandum of Understanding stating that the certified LEA would either provide test examiner training or provide a trained test examiner to perform the testing. For more information on the training of administration and scoring for the Initial Alternate ELPAC, refer to section [*5.2. Administration and Scoring Training*](#_Administration_and_Scoring_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 the reporting specifications approved by the CDE. These specifications contain detailed scoring procedures, along with the procedures for determining whether a student has attempted an assessment and whether that student’s response data should be included in the statistical analyses and calculations for computing summary data.

Prior to the test administration, ETS’ ALTD staff reviewed and verified the keys for each item. Then, these keys were provided to CAI for implementing machine scoring of the selected-response items. In addition, the student’s original response string was stored for data verification and auditing purposes. 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.

ETS’ scoring specifications for the Initial Alternate ELPAC were completed, approved, and checked well in advance of the receipt of student response data. These specifications contained detailed scoring procedures, as well as the procedures for determining whether a student attempted an assessment and whether that student’s response data should be included in the statistical analyses and computing summary data.

### Quality Control of Psychometric Processes

#### Scoring Verification

ETS developed two independent and parallel scoring structures to produce students’ scores: the Enterprise Score Key Management (eSKM) scoring system, which collected, scored, and delivered individual students’ scores to the ETS reporting system; and then the ETS PAR team computed individual student scores based on the same scoring specifications as described in subsection [*9.5.3 Development of Scoring Specifications*](#_Toc120784183). The scores from the two sources were then compared for internal quality control. Any differences in the scores were discussed and resolved. All scores complied with the ETS scoring specifications and passed the parallel scoring process. This ensured the quality and accuracy of scoring and supported the transfer of scores into TOMS, the database of the student records scoring.

#### Psychometric Analyses

The psychometric procedures for the Initial Alternate ELPAC were developed, reviewed, and approved prior to the receipt of student response data. The ETS psychometric team also developed specifications for each of the psychometric analyses performed. These specifications contain detailed descriptions of the analysis steps such as sample inclusion, analyses methods, and special handling of the data.

All psychometric analyses conducted at ETS underwent comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists and psychometric specifications were developed by members of the team for each of the statistical procedures performed on Initial Alternate ELPAC results data.

Once raw-to-scale-score conversion tables for each form were generated, psychometricians carried out quality-control checks on each scoring table to verify

* all possible raw scores for each form were included in the tables;
* the lowest obtainable scale score and the highest obtainable scale score matched the specifications for each grade level, respectively; and
* the threshold score for the performance level was correctly identified.

After all quality-control steps were completed and any differences were resolved, one final inspection of scoring tables was conducted prior to uploading the tables to eSKM for score reporting.

### Quality Control of Reporting

To ensure the quality of Initial Alternate ELPAC results, for both individual student and summary reports, three general areas were evaluated:

1. Comparison of report formats with input sources from the CDE-approved samples
2. Validation of the report data through quality-control checks performed by ETS’ Data Quality Services and Center of Reporting & Scoring Services teams, as well as running of all Student Score Reports (SSRs) through ETS’ patented QC Interrogator software, which compares elements of the SSR to acceptable values to identify errors and is used in conjunction with human review to detect errors on every score report batch as part of quality-control procedures
3. Proofreading of the quality-control and production reports by the CDE and ETS prior to making reports available to the LEA for download in TOMS as well as via the LEA’s student information system

All reports were required to include a single, accurate LEA code, 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. The CDE provided a revised LEA Master File to ETS throughout the year as updates became available.

After the reports were validated in accordance with the CDE’s requirements, a set of reports representing all possible grade levels and reporting outcomes was provided to the CDE and ETS for review and approval. Electronic reports were sent on the actual report template to the CDE. The CDE and ETS reviewed and approved the reports after a thorough examination.

Upon the CDE’s approval of the reports generated for the quality-control LEAs, ETS proceeded with the first batch of report production. The reviewed set of reports incorporated CDE-selected LEAs and provided the final check prior to generating all reports and making them available electronically for download in TOMS and for student information systems through an application programming interface.

### Quality Control of End-to-End Testing

ETS conducted end-to-end testing prior to the start of the test administration. The purpose of this testing is to verify that all systems, processes, and resources were ready for the operational administration. Once released from processing, 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.

#### Computer-based Assessments

ETS employed a number of strategies to verify ongoing systems performance, including monitoring of system availability and system usage. Time was allotted for UAT to confirm that the systems met requirements and to make identified corrections before final deployment. To accomplish system acceptance and sign-off, ETS deployed systems to a staging area, which mirrors the final production environment, for operational testing and UAT. Final approval by the CDE triggered final deployment of the system.

To begin the quality-control process for end-to-end testing of the administration, the ETS program and resolutions teams prepared by entering responses in computer-based assessments for all grade levels and grade spans. These responses were entered for fictitious students in selected schools and across several LEAs. Each student’s assessment was completed with responses that were all correct, all incorrect, and combinations of correct and incorrect. These response combinations were the expected results across performance levels and score ranges. The responses were sent for processing, including for system quality control of computer-based assessments.

Once released from processing, 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 on the basis of the fictitious students when 100 percent quality control was demonstrated by ETS’ Resolution staff.

### References

California Department of Education. (2023a). *Alternate assessment decision-making tool for California.* Sacramento, CA: California Department of Education.

California Department of Education. (2023b). *CAASPP and ELPAC Test Operations Management System user guide*. Sacramento, CA: California Department of Education.

Educational Testing Service. (2014). *ETS Standards for Quality and Fairness*. Princeton, NJ: Educational Testing Service.

## Threshold Score Validation Study

ETS conducted a threshold score validation study to provide validity evidence for the performance-level cut scores produced by the Initial Alternate English Language Proficiency Assessments for California (ELPAC). The study was designed to use a contrasting-groups method to elicit ratings from test takers’ teachers about their English learner (EL) status based on their in-class performance. These ratings would be compared with the categorization into EL versus non-EL groups based on test takers’ total test scores on the Initial Alternate ELPAC. However, since only 26 grade one students, and fewer than 10 students at any grade level above grade one, had taken the Initial Alternate ELPAC at the time the study was carried out, the methodology could only be applied using kindergarten data. The study found that 83 percent of teachers agreed or strongly agreed with the assessment’s categorization of test takers as EL students.

Details describing the background, statistical method, study participants, and results, as well as a discussion, can be found in the *Initial Alternate ELPAC Threshold Score Validation Study Final Report* (California Department of Education, 2023).

### Reference

California Department of Education. (2023). *Initial Alternate English Language Proficiency Assessments for California threshold score validation study final report*[Unpublished report]. Sacramento, CA: California Department of Education.

## Continuous and Systematic Improvement

The first operational administration of the Initial Alternate English Language Proficiency Assessments for California (ELPAC) occurred in 2022–23. Continuous efforts have been made to improve the assessments. This chapter summarizes accomplishments and ongoing improvements for the Initial Alternate ELPAC as well as strategies to implement possible future improvements.

### 2022–23 Feedback for Continuous Improvement Survey

The ELPAC program annually solicits feedback from educators through the Feedback for Continuous Improvement Survey. Local educational agency (LEA) and test site staff, as well as test administrators and test examiners, were invited to participate in the 2022–23 Feedback for Continuous Improvement Survey. Its goal was to highlight successes and identify areas for improvement. A total of 3,869 survey respondents participated in this survey for the 2022–23 administration. The California Department of Education (CDE) and ETS use key recommendations from educators to implement positive changes in the following administration year.

Educators provided valuable feedback for potential improvements to the future administration of the California Assessment of Student Performance and Progress (CAASPP) and the ELPAC by reporting some lessons they learned in 2022–23. Based on those lessons and suggestions for improvement, the *CAASPP and ELPAC Feedback for Continuous Improvement Survey and Focus Groups Report* (CDE, 2023) presents recommendations for the CDE, with the goal of enhancing the administrative support provided to LEAs and schools for future CAASPP and ELPAC test administrations. Refer also to subsection [*5.4.4 Feedback for Continuous Improvement Survey*](#_Feedback_for_Continuous_1) for assessment-specific results.

#### Recommendations for Improvement

In response to the LEA feedback, ETS and the CDE will continue to clarify and expand the use of universal tools, designated supports, and accommodations in daily instruction and on assessments to address respondents’ confusion regarding the assignment and use of embedded accessibility resources.

### Test Design

The Alternate ELPAC was designed to provide a consistent, standardized measurement of English language proficiency (ELP) across California for students with the most significant cognitive disabilities as determined by the student’s individualized educational program team. A guiding principle for development was that the assessment be designed to ensure that the intended test-taking population is able to demonstrate ELP.

The Alternate ELPAC Post-Test Survey had a set of questions that asked respondents the level of agreement around the guiding principles of the Alternate ELPAC test design (CDE, 2023). In all cases, the majority of respondents agreed or strongly agreed that the assessment met each guiding principle. The feedback shows evidence that educators found that the test design was successful at achieving the guiding principles.

### Test Delivery

#### Changes to the Test Administrator Interface

The Test Administrator Interface will be updated to a cleaner, more user-friendly appearance. This will include a new functionality that allows the test examiner to pin information for specific students to the top of the screen for monitoring.

#### Changes to Ending the Assessment in the Test Delivery System

The process for ending the assessment will be streamlined. After the last question is presented, students will select [**Next**] (instead of [**End Test**]) to reach the review screen, which will include the [**Submit Test**] button.

### Student Score Reports Redesign

Redesigned Student Score Reports (SSRs) will be made available; SSR formats are PDF and HTML. For an HTML SSR, an LEA or parent or student portal vendor will provide a link to a parent/guardian.

Additionally, SSRs for the ELPAC will be available in Arabic.

### Accessibility Resources

Like all ELPAC assessments, the Initial Alternate ELPAC is administered using the test delivery system (TDS) created by Cambium Assessment, Inc. As such, implementation of new computer-based universal tools, designated supports, and accommodations are aligned with the TDS.

The following change will be implemented during the 2023–24 Initial Alternate ELPAC administration:

* The definition of the non-embedded medical supports designated support will be updated to mention “Bluetooth hearing aids.”

### Reference

California Department of Education. (2023). *Alternate ELPAC post-test survey summary report* [Unpublished report]. California Department of Education.

1. Data for 2022–23 was retrieved from the *CalEdFacts* web page on the CDE website. [↑](#footnote-ref-2)
2. This definition was retrieved from the CDE CALPADS web page on the CDE website. [↑](#footnote-ref-3)
3. This technical report is based on the versions of the Accessibility Matrix that was available during the 2022–23 ELPAC administration. [↑](#footnote-ref-4)
4. The IDEA is the primary federal program that authorizes state and local aid for special education and related services for children with disabilities. [↑](#footnote-ref-5)