

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



# Alternate English Language Proficiency Assessments for California 2021–22 Operational Field Test Technical Report

**Submitted May 1, 2023**

**By ETS**



**Contract # CN150012**

Table of Contents

[Chapter 1: Introduction 1](#_Toc131594524)

[1.1. Background 1](#_Toc131594525)

[1.2. Test Purposes 1](#_Toc131594526)

[1.3. Test Content and Design 2](#_Toc131594527)

[1.4. Intended Population 2](#_Toc131594528)

[1.5. Testing Windows and Estimated Testing Times 3](#_Toc131594529)

[1.6. Groups and Organizations Involved with the Alternate ELPAC 3](#_Toc131594533)

[1.6.1. California State Board of Education 3](#_Toc131594534)

[1.6.2. California Department of Education 4](#_Toc131594535)

[1.6.3. California Educators 4](#_Toc131594536)

[1.6.4. Contractors 4](#_Toc131594537)

[1.7. Systems Overview and Functionality 5](#_Toc131594538)

[1.7.1. Test Operations Management System 5](#_Toc131594539)

[1.7.2. Test Delivery System 6](#_Toc131594540)

[1.7.3. Practice and Training Tests 6](#_Toc131594541)

[1.7.4. California Educator Reporting System 7](#_Toc131594542)

[1.7.5. Test Results for California’s Assessments Website 7](#_Toc131594543)

[1.8. Overview of the Technical Report 7](#_Toc131594544)

[References 9](#_Toc131594545)

[Chapter 2: Overview of Alternate ELPAC Processes 10](#_Toc131594546)

[2.1. Item Development 10](#_Toc131594547)

[2.1.1. Selection of ELD Connectors for Development 10](#_Toc131594548)

[2.1.2. Item Format 10](#_Toc131594549)

[2.1.3. Task Type Specifications 11](#_Toc131594550)

[2.2. Test Assembly 11](#_Toc131594551)

[2.2.1. Test Blueprint 11](#_Toc131594552)

[2.2.2. Test Length 11](#_Toc131594553)

[2.3. Test Administration 12](#_Toc131594554)

[2.3.1. Test Security and Confidentiality 12](#_Toc131594555)

[2.3.2. Procedures to Maintain Standardization 12](#_Toc131594556)

[2.4. Fairness and Accessibility 13](#_Toc131594557)

[2.4.1. Overview 13](#_Toc131594558)

[2.4.2. Universal Tools, Designated Supports, and Accommodations 13](#_Toc131594559)

[2.4.3. Description of Differential Item Functioning Analyses 14](#_Toc131594560)

[2.5. Scores 14](#_Toc131594561)

[2.5.1. Estimating Ability Scores 14](#_Toc131594562)

[2.5.2. Score Reporting 14](#_Toc131594563)

[2.5.3. Aggregation Procedures 15](#_Toc131594564)

[2.6. Psychometric Analyses 15](#_Toc131594565)

[2.6.1. Description of the Classical Item Analyses 15](#_Toc131594566)

[2.6.2. Description of Item Response Theory Analyses 16](#_Toc131594567)

[References 17](#_Toc131594568)

[Chapter 3: Item Development and Review 18](#_Toc131594569)

[3.1. Overview 18](#_Toc131594570)

[3.1.1. Preparation 18](#_Toc131594571)

[3.1.2. Task Type Specifications 18](#_Toc131594572)

[3.1.3. Recruitment and Selection of Item Writers 19](#_Toc131594573)

[3.1.4. Item Writing by Contractors 21](#_Toc131594574)

[3.1.5. Item Writer Training 21](#_Toc131594575)

[3.2. ETS Item Review Process 22](#_Toc131594576)

[3.2.1. Overview 22](#_Toc131594577)

[3.2.2. ETS Content Review 23](#_Toc131594578)

[3.2.3. ETS Accessibility Review 23](#_Toc131594579)

[3.2.4. ETS Editorial Review 23](#_Toc131594580)

[3.2.5. ETS Sensitivity and Fairness Review 23](#_Toc131594581)

[3.3. California Department of Education Review 24](#_Toc131594582)

[3.4. California Educator Review 24](#_Toc131594583)

[3.4.1. California Educators as Content Experts 25](#_Toc131594584)

[3.4.2. Composition of Item Review Panels 25](#_Toc131594585)

[3.4.3. Meetings for Review of Alternate ELPAC Items 26](#_Toc131594586)

[3.5. Data Review Meeting 27](#_Toc131594587)

[References 28](#_Toc131594588)

[Chapter 4: Test Assembly 29](#_Toc131594589)

[4.1. Overview 29](#_Toc131594590)

[4.2. High-Level Test Design, Test Blueprint, and Form Assembly Specifications 29](#_Toc131594591)

[4.2.1. Test Blueprints 30](#_Toc131594592)

[4.2.2. High-Level Test Design 31](#_Toc131594593)

[4.2.3. Form Assembly Specifications 32](#_Toc131594594)

[4.3. Test Production Process 32](#_Toc131594595)

[4.3.1. Selection of Task Types and Items 33](#_Toc131594596)

[4.3.2. Test Forms 33](#_Toc131594597)

[4.3.3. Psychometric Criteria and Review 33](#_Toc131594598)

[4.3.4. Content Review of Forms 34](#_Toc131594599)

[4.3.5. California Department of Education Forms Review 34](#_Toc131594600)

[4.3.6. Configuration of the Test Delivery System 34](#_Toc131594601)

[4.3.7. Test Form Delivery 35](#_Toc131594602)

[References 36](#_Toc131594603)

[Chapter 5: Test Administration 37](#_Toc131594604)

[5.1. Overview 37](#_Toc131594605)

[5.2. Administration and Scoring Training 37](#_Toc131594606)

[5.2.1. Overview 37](#_Toc131594607)

[5.2.2. Goals 38](#_Toc131594608)

[5.2.3. Local Educational Agency Training Requirement 38](#_Toc131594609)

[5.2.4. Materials on the Moodle Training Site 39](#_Toc131594610)

[5.3. User Roles and Standardization 42](#_Toc131594611)

[5.3.1. Local Educational Agency ELPAC Coordinator 42](#_Toc131594612)

[5.3.2. Site ELPAC Coordinator 43](#_Toc131594613)

[5.3.3. Test Examiner 44](#_Toc131594614)

[5.3.4. Instructions for Test Administration 44](#_Toc131594615)

[5.4. Local Educational Agency Training 45](#_Toc131594616)

[5.4.1. Synchronous and Asynchronous Training 46](#_Toc131594617)

[5.4.2. Videos and Guides 46](#_Toc131594618)

[5.4.3. Training for Proper Identification and Assignment of Designated Supports and Accommodations 47](#_Toc131594619)

[5.4.4. Feedback for Continuous Improvement Survey 48](#_Toc131594620)

[5.5. Accessibility Resources 48](#_Toc131594621)

[5.5.1. Accessibility Resource Categories 48](#_Toc131594622)

[5.5.2. Identification and Selection 51](#_Toc131594623)

[5.5.3. Assignment 51](#_Toc131594624)

[5.5.4. Delivery of Embedded and Non-Embedded Resources to Students 55](#_Toc131594625)

[5.5.5. Usage of Designated Supports and Accommodations 55](#_Toc131594626)

[5.6. Practice and Training Tests 57](#_Toc131594627)

[5.7. Test Security and Confidentiality 57](#_Toc131594628)

[5.7.1. ETS’ Office of Testing Integrity 58](#_Toc131594629)

[5.7.2. Procedures to Maintain Standardization of Test Security 58](#_Toc131594630)

[5.7.3. Test Security Monitoring 59](#_Toc131594631)

[5.7.4. Security of Electronic Files Using a Firewall 60](#_Toc131594632)

[5.7.5. Transfer of Scores via Secure Data Exchange 60](#_Toc131594633)

[5.7.6. Data Management in the Secure Database 60](#_Toc131594634)

[5.7.7. Statistical Analysis on Secure Servers 60](#_Toc131594635)

[5.7.8. Student Confidentiality 61](#_Toc131594636)

[5.7.9. Student Test Results 61](#_Toc131594637)

[5.7.10. Security and Test Administration Incident Reporting System Process 62](#_Toc131594638)

[5.7.11. Appeals 63](#_Toc131594639)

[References 66](#_Toc131594640)

[Chapter 6: Standard Setting 67](#_Toc131594641)

[6.1. Background 67](#_Toc131594642)

[6.2. Performance Level Descriptors 67](#_Toc131594643)

[6.3. Standard Setting Methodology 68](#_Toc131594644)

[6.3.1. Modified and Extended Angoff Methods 68](#_Toc131594645)

[6.4. Standard Setting Procedures 68](#_Toc131594646)

[6.4.1. Panelists 68](#_Toc131594647)

[6.4.2. Materials 69](#_Toc131594648)

[6.4.3. Process 69](#_Toc131594649)

[6.5. Results of the Standard Setting 72](#_Toc131594650)

[References 73](#_Toc131594651)

[Chapter 7: Scoring and Reporting 74](#_Toc131594652)

[7.1. Student Test Scores 74](#_Toc131594653)

[7.1.1. Raw Scores 74](#_Toc131594654)

[7.1.2. Theta Scores 75](#_Toc131594655)

[7.1.3. Scale Scores for the Total Assessment 75](#_Toc131594656)

[7.1.4. Performance Levels 76](#_Toc131594657)

[7.2. Overview of Score Aggregation Procedures 76](#_Toc131594658)

[7.2.1. Student Score Distributions and Summary Statistics 77](#_Toc131594659)

[7.2.2. Demographic Student Group Summaries 78](#_Toc131594660)

[7.2.3. Student Group Distributions 79](#_Toc131594661)

[7.3. Reports Produced and Scores for Each Report 80](#_Toc131594662)

[7.3.1. Online Reporting 80](#_Toc131594663)

[7.3.2. Special Cases 81](#_Toc131594664)

[7.3.3. Types of Score Reports 81](#_Toc131594665)

[7.3.4. Criteria for Interpreting Test Scores 82](#_Toc131594666)

[7.3.5. Criteria for Interpreting Score Reports 82](#_Toc131594667)

[References 84](#_Toc131594668)

[Appendix 7.A: Raw and Scale Score Frequency Distributions 85](#_Toc131594669)

[Appendix 7.B: Means and Standard Deviations of Scale Scores and Percentage of Students in Each Performance Level by Demographic Student Group 106](#_Toc131594670)

[Appendix 7.C: Demographic Student Group Summaries 114](#_Toc131594671)

[Chapter 8: Psychometric Analyses 121](#_Toc131594672)

[8.1. Overview 121](#_Toc131594673)

[8.1.1. Summary of the Analyses 121](#_Toc131594674)

[8.1.2. Samples Used for the Analyses 122](#_Toc131594675)

[8.1.3. Test-Taking Rates 123](#_Toc131594676)

[8.2. Classical Item Analyses 124](#_Toc131594677)

[8.2.1. Classical Item Difficulty Indices (*p*-value and Average Item Score) 124](#_Toc131594678)

[8.2.2. Item-Total Correlation 125](#_Toc131594679)

[8.2.3. Distribution of Item Scores 126](#_Toc131594680)

[8.2.4. Omit Rates 126](#_Toc131594681)

[8.2.5. Distractor Analyses 126](#_Toc131594682)

[8.2.6. Summary of Classical Item Analysis Flagging Criteria 127](#_Toc131594683)

[8.2.7. Classical Item Analyses Results 128](#_Toc131594684)

[8.3. Differential Item Functioning Analyses 128](#_Toc131594685)

[8.3.1. Differential Item Functioning Procedure for Dichotomous Items 129](#_Toc131594686)

[8.3.2. Differential Item Functioning Procedure for Polytomous Items 130](#_Toc131594687)

[8.3.3. Classification 131](#_Toc131594688)

[8.3.4. Differential Item Functioning Analysis Results Summary 132](#_Toc131594689)

[8.4. Item Response Theory Analyses 132](#_Toc131594690)

[8.4.1. Item Response Theory Model 132](#_Toc131594691)

[8.4.2. Data Preparation 133](#_Toc131594692)

[8.4.3. Equating 133](#_Toc131594693)

[8.4.4. Parameter Estimates 135](#_Toc131594694)

[8.5. Scaling the Scores 136](#_Toc131594695)

[8.5.1. Inverse Test Characteristic Curve Procedure 137](#_Toc131594696)

[8.5.2. Transformation from Theta Scores to Scale Scores 137](#_Toc131594697)

[8.6. Response Time Analyses 139](#_Toc131594698)

[8.7. Dimensionality Study 140](#_Toc131594699)

[8.7.1. Overview 140](#_Toc131594700)

[8.7.2. Methods 141](#_Toc131594701)

[8.7.3. Results 147](#_Toc131594702)

[8.8. Reliability Analyses 148](#_Toc131594703)

[8.8.1. Reliability Measures 149](#_Toc131594704)

[8.8.2. Standard Error of Measurement 150](#_Toc131594705)

[8.8.3. Student Group Reliabilities 151](#_Toc131594706)

[8.8.4. Conditional Standard Errors of Measurement 151](#_Toc131594707)

[8.8.5. Decision Classification Analyses 153](#_Toc131594708)

[8.8.6. Interrater Agreement 155](#_Toc131594709)

[8.9. Validity Evidence 157](#_Toc131594710)

[8.9.1. Design of the Alternate ELPAC 158](#_Toc131594711)

[8.9.2. Content 159](#_Toc131594712)

[8.9.3. Response Processes 160](#_Toc131594713)

[8.9.4. Relationship to Other Variables 161](#_Toc131594714)

[8.9.5. Internal Structure 163](#_Toc131594715)

[8.9.6. English-Only Student Sample Validation Analysis 163](#_Toc131594716)

[References 165](#_Toc131594717)

[Accessibility Information 168](#_Toc131594718)

[Alternative Text for Equation 8.1 168](#_Toc131594719)

[Alternative Text for Equation 8.2 168](#_Toc131594720)

[Alternative Text for Equation 8.3 168](#_Toc131594721)

[Alternative Text for Equation 8.4 168](#_Toc131594722)

[Alternative Text for Equation 8.5 168](#_Toc131594723)

[Alternative Text for Equation 8.6 168](#_Toc131594724)

[Alternative Text for Equation 8.7 168](#_Toc131594725)

[Alternative Text for Equation 8.8 168](#_Toc131594726)

[Alternative Text for Equation 8.9 169](#_Toc131594727)

[Alternative Text for Equation 8.10 169](#_Toc131594728)

[Alternative Text for Equation 8.11 169](#_Toc131594729)

[Alternative Text for Equation 8.12 169](#_Toc131594730)

[Alternative Text for Equation 8.13 169](#_Toc131594731)

[Alternative Text for Equation 8.14 169](#_Toc131594732)

[Alternative Text for Equation 8.15 169](#_Toc131594733)

[Alternative Text for Equation 8.16 169](#_Toc131594734)

[Alternative Text for Equation 8.17 169](#_Toc131594735)

[Alternative Text for Equation 8.18 169](#_Toc131594736)

[Alternative Text for Equation 8.19 169](#_Toc131594737)

[Alternative Text for Equation 8.20 170](#_Toc131594738)

[Alternative Text for Equation 8.21 170](#_Toc131594739)

[Alternative Text for Equation 8.22 170](#_Toc131594740)

[Alternative Text for Equation 8.23 170](#_Toc131594741)

[Alternative Text for Equation 8.24 170](#_Toc131594742)

[Alternative Text for Equation 8.25 170](#_Toc131594743)

[Alternative Text for Equation 8.26 170](#_Toc131594744)

[Alternative Text for Equation 8.27 170](#_Toc131594745)

[Alternative Text for Equation 8.28 170](#_Toc131594746)

[Alternative Text for Equation 8.29 170](#_Toc131594747)

[Alternative Text for Equation 8.30 171](#_Toc131594748)

[Appendix 8.A: Classical Item Analyses Results 172](#_Toc131594749)

[Appendix 8.B: Differential Item Functioning Results 184](#_Toc131594750)

[Appendix 8.C: Item Response Theory Results 186](#_Toc131594751)

[Appendix 8.D: Response Time Results 207](#_Toc131594752)

[Appendix 8.E: Factor Loading 209](#_Toc131594753)

[Appendix 8.F: Reliability of Alternate ELPAC Performance and Decision   
Classification 224](#_Toc131594754)

[Appendix 8.G: Raw-to-Scale-Score Conversions 232](#_Toc131594755)

[Appendix 8.H: Interrater Reliability 243](#_Toc131594756)

[Chapter 9: Quality Control 246](#_Toc131594757)

[9.1. Quality Control of Item Development 246](#_Toc131594758)

[9.2. Quality Control of Test Assembly and Delivery 246](#_Toc131594759)

[9.3. Quality Control of Test Materials 247](#_Toc131594760)

[9.3.1. Test Administration Manuals 247](#_Toc131594761)

[9.3.2. Processing Test Materials 247](#_Toc131594762)

[9.4. Quality Control of Test Administration 248](#_Toc131594763)

[9.5. Quality Control of Scoring 248](#_Toc131594764)

[9.5.1. Machine-Scoring Procedures 248](#_Toc131594765)

[9.5.2. Rubric-Scored Item Scoring 249](#_Toc131594766)

[9.5.3. Development of Scoring Specifications 249](#_Toc131594767)

[9.6. Quality Control of Psychometric Processes 250](#_Toc131594768)

[9.6.1. Scoring Verification 250](#_Toc131594769)

[9.6.2. Psychometric Analyses 250](#_Toc131594770)

[9.7. Quality Control of Reporting 251](#_Toc131594771)

[9.7.1. Exclusion of Student Scores from Summary Reports 251](#_Toc131594772)

[9.8. Quality Control of End-to-End Testing 252](#_Toc131594773)

[9.8.1. Computer-based Assessments 252](#_Toc131594774)

[Reference 253](#_Toc131594775)

[Chapter 10: In-Test Survey 254](#_Toc131594776)

[10.1. Survey Design and Development 254](#_Toc131594777)

[10.2. Survey Administration 254](#_Toc131594778)

[10.3. Summary of Survey Responses 255](#_Toc131594779)

[Appendix 10.A: Student Survey Results 256](#_Toc131594780)

[Chapter 11: Continuous and Systematic Improvement 306](#_Toc131594781)

[11.1. Test Design 306](#_Toc131594782)

[11.2. Item Development and Item Pool Diversity 306](#_Toc131594783)

[11.3. Administration and Test Delivery 307](#_Toc131594784)

[11.3.1. Post-Test Survey 307](#_Toc131594785)

[11.3.2. Training and Communication 308](#_Toc131594786)

[11.3.3. Accessibility 308](#_Toc131594787)

[*11.3.4. Directions for Administration* 309](#_Toc131594788)

[11.4. Psychometric Analyses 309](#_Toc131594789)

[Reference 310](#_Toc131594790)

List of Tables

[Acronyms and Initialisms Used in the *Alternate English Language Proficiency Assessments for California Technical Report* xiv](#_Toc133500692)

[Table 1.1 Differences Between the Initial Alternate ELPAC and Summative Alternate   
ELPAC 1](#_Toc133500693)

[Table 3.1 Alternate ELPAC IWW and IRM Participant Qualifications by Meeting Type and Total 20](#_Toc133500694)

[Table 3.2 Status of Items After the 2020 Alternate ELPAC IRM 26](#_Toc133500695)

[Table 3.3 Status of Items After the 2021 Alternate ELPAC IRM 26](#_Toc133500696)

[Table 4.1 Kindergarten Through Grade Twelve Summative Alternate ELPAC Form Plan 32](#_Toc133500697)

[Table 5.1 Assignment of Accommodations and Designated Supports—Kindergarten   
Through Grade Two 52](#_Toc133500698)

[Table 5.2 Assignment of Accommodations and Designated Supports—Grade Three Through Grade Twelve 54](#_Toc133500699)

[Table 5.3 Assignment and Usage of Embedded Designated Supports 56](#_Toc133500700)

[Table 5.4 Types of Appeals 63](#_Toc133500701)

[Table 5.5 Number and Types of Incidents Submitted in STAIRS 64](#_Toc133500702)

[Table 5.6 Number of Appeals Approved in STAIRS in the 2021–22 Administration—All Grade Levels and Grade Spans 65](#_Toc133500703)

[Table 6.1 Alternate ELPAC General PLDs 68](#_Toc133500704)

[Table 6.2 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Kindergarten 71](#_Toc133500705)

[Table 6.3 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade One 71](#_Toc133500706)

[Table 6.4 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Two 71](#_Toc133500707)

[Table 6.5 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Three Through Five 71](#_Toc133500708)

[Table 6.6 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Six Through Eight 71](#_Toc133500709)

[Table 6.7 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Nine Through Twelve 72](#_Toc133500710)

[Table 6.8 SSPI’s Recommendations for the Proposed Performance Levels for the Alternate ELPAC 72](#_Toc133500711)

[Table 7.1 Reporting Scale Score Ranges for Each Performance Level by Grade Level 76](#_Toc133500712)

[Table 7.2 Mean and SD of Scale Scores by Grade Level 77](#_Toc133500713)

[Table 7.3 Percentage of Students in Each Performance Level by Grade Level 77](#_Toc133500714)

[Table 7.4 Demographic Student Groups to Be Reported 79](#_Toc133500715)

[Table 7.A.1 Raw Score Frequency Distribution for Kindergarten 85](#_Toc133500716)

[Table 7.A.2 Raw Score Frequency Distribution for Grade One 86](#_Toc133500717)

[Table 7.A.3 Raw Score Frequency Distribution for Grade Two 87](#_Toc133500718)

[Table 7.A.4 Raw Score Frequency Distribution for Grade Span Three Through Five, Form One 88](#_Toc133500719)

[Table 7.A.5 Raw Score Frequency Distribution for Grade Span Three Through Five, Form Two 89](#_Toc133500720)

[Table 7.A.6 Raw Score Frequency Distribution for Grade Span Six Through Eight,   
Form One 90](#_Toc133500721)

[Table 7.A.7 Raw Score Frequency Distribution for Grade Span Six Through Eight,   
Form Two 91](#_Toc133500722)

[Table 7.A.8 Raw Score Frequency Distribution for Grade Span Nine and Ten, Form One 92](#_Toc133500723)

[Table 7.A.9 Raw Score Frequency Distribution for Grade Span Nine and Ten, Form Two 93](#_Toc133500724)

[Table 7.A.10 Raw Score Frequency Distribution for Grade Span Eleven and Twelve, Form One 94](#_Toc133500725)

[Table 7.A.11 Raw Score Frequency Distribution for Grade Span Eleven and Twelve, Form Two 95](#_Toc133500726)

[Table 7.A.12 Scale Score Frequency Distribution for Kindergarten 96](#_Toc133500727)

[Table 7.A.13 Scale Score Frequency Distribution for Grade One 97](#_Toc133500728)

[Table 7.A.14 Scale Score Frequency Distribution for Grade Two 98](#_Toc133500729)

[Table 7.A.15 Scale Score Frequency Distribution for Grade Span Three Through Five 99](#_Toc133500730)

[Table 7.A.16 Scale Score Frequency Distribution for Grade Span Six Through Eight 100](#_Toc133500731)

[Table 7.A.17 Scale Score Frequency Distribution for Grade Span Nine and Ten 102](#_Toc133500732)

[Table 7.A.18 Scale Score Frequency Distribution for Grade Span Eleven and Twelve 104](#_Toc133500733)

[Table 7.B.1 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Kindergarten 106](#_Toc133500734)

[Table 7.B.2 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade One 108](#_Toc133500735)

[Table 7.B.3 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Two 109](#_Toc133500736)

[Table 7.B.4 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Three Through Five 110](#_Toc133500737)

[Table 7.B.5 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Six Through Eight 111](#_Toc133500738)

[Table 7.B.6 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Nine and Ten 112](#_Toc133500739)

[Table 7.B.7 Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Eleven and Twelve 113](#_Toc133500740)

[Table 7.C.1 Demographic Summary for Students: Kindergarten 114](#_Toc133500741)

[Table 7.C.2 Demographic Summary for Students: Grade One 116](#_Toc133500742)

[Table 7.C.3 Demographic Summary for Students: Grade Two 117](#_Toc133500743)

[Table 7.C.4 Demographic Summary for Students: Grade Span Three Through Five 118](#_Toc133500744)

[Table 7.C.5 Demographic Summary for Students: Grade Span Six Through Eight 119](#_Toc133500745)

[Table 7.C.6 Demographic Summary for Students: Grade Span Nine Through Twelve 120](#_Toc133500746)

[Table 8.1 Alternate ELPAC Analyses Data Sources 122](#_Toc133500747)

[Table 8.2 The Test-Taking Rates by Grade Level 123](#_Toc133500748)

[Table 8.3 Summary Statistics for Classical Item Analysis 128](#_Toc133500749)

[Table 8.4 Student Groups for DIF Comparison 129](#_Toc133500750)

[Table 8.5 DIF Categories for Dichotomous Items 131](#_Toc133500751)

[Table 8.6 DIF Categories for Polytomous Items 132](#_Toc133500752)

[Table 8.7 Distribution of *b*-values 135](#_Toc133500753)

[Table 8.8 Slopes and Intercepts That Convert Theta Scores to Reporting Scale Scores 138](#_Toc133500754)

[Table 8.9 Item Distribution of the Alternate ELPAC Operational Field Test by Communication Mode 147](#_Toc133500755)

[Table 8.10 Dimensionality Evaluation 148](#_Toc133500756)

[Table 8.11 Correlations between the Communication Mode Scores 148](#_Toc133500757)

[Table 8.12 Test Reliability of the Total Scores 151](#_Toc133500758)

[Table 8.13 Scale Score CSEMs at the Reporting Score Range Thresholds 153](#_Toc133500759)

[Table 8.14 Decision Accuracy for Reaching a Performance Level 154](#_Toc133500760)

[Table 8.15 Decision Consistency for Reaching a Performance Level 155](#_Toc133500761)

[Table 8.16 Frequencies of Ratings 156](#_Toc133500762)

[Table 8.17 Correlation of Alternate ELPAC and CAA for ELA Scores 161](#_Toc133500763)

[Table 8.18 Correlation of Alternate ELPAC and CAA for Mathematics Scores 162](#_Toc133500764)

[Table 8.19 Correlation of Alternate ELPAC and CAA for Science Scores 162](#_Toc133500765)

[Table 8.20 Target Sample Size and the Number of Participants 164](#_Toc133500766)

[Table 8.A.1 Classical Item Statistics for Kindergarten 172](#_Toc133500767)

[Table 8.A.2 Classical Item Statistics for Grade One 174](#_Toc133500768)

[Table 8.A.3 Classical Item Statistics for Grade Two 176](#_Toc133500769)

[Table 8.A.4 Classical Item Statistics for Grade Span Three Through Five 178](#_Toc133500770)

[Table 8.A.5 Classical Item Statistics for Grade Span Six Through Eight 180](#_Toc133500771)

[Table 8.A.6 Classical Item Statistics for Grade Span Nine Through Twelve 182](#_Toc133500772)

[Table 8.B.1 Gender DIF Classifications Summary by Grade Level or Grade Span 184](#_Toc133500773)

[Table 8.B.2 Hispanic or Latino versus Non-Hispanic and Non-Latino DIF Classifications Summary by Grade Level or Grade Span 185](#_Toc133500774)

[Table 8.B.3 Items Exhibiting Significant DIF by Primary Disability Student Group 185](#_Toc133500775)

[Table 8.C.1 IRT Item Statistics, Kindergarten 186](#_Toc133500776)

[Table 8.C.2 IRT Item Statistics, Grade One 187](#_Toc133500777)

[Table 8.C.3 IRT Item Statistics, Grade Two 188](#_Toc133500778)

[Table 8.C.4 IRT Item Statistics, Grade Span Three Through Five 189](#_Toc133500779)

[Table 8.C.5 IRT Item Statistics, Grade Span Six Through Eight 191](#_Toc133500780)

[Table 8.C.6 IRT Item Statistics, Grade Span Nine Through Twelve 193](#_Toc133500781)

[Table 8.C.7 Item–Person Map for Kindergarten 195](#_Toc133500782)

[Table 8.C.8 Item–Person Map for Grade One 197](#_Toc133500783)

[Table 8.C.9 Item–Person Map for Grade Two 199](#_Toc133500784)

[Table 8.C.10 Item–Person Map for Grade Span Three Through Five 201](#_Toc133500785)

[Table 8.C.11 Item–Person Map for Grade Span Six Through Eight 203](#_Toc133500786)

[Table 8.C.12 Item–Person Map for Grade Span Nine Through Twelve 205](#_Toc133500787)

[Table 8.D.1 Total Testing Time (in Minutes) at Each Scale Score Interval 207](#_Toc133500788)

[Table 8.E.1 Standardized Factor Loading for Kindergarten 210](#_Toc133500789)

[Table 8.E.2 Standardized Factor Loading for Grade One 212](#_Toc133500790)

[Table 8.E.3 Standardized Factor Loading for Grade Two 214](#_Toc133500791)

[Table 8.E.4 Standardized Factor Loading for Grade Span Three Through Five 216](#_Toc133500792)

[Table 8.E.5 Standardized Factor Loading for Grade Span Six Through Eight 219](#_Toc133500793)

[Table 8.E.6 Standardized Factor Loading for Grade Span Nine Through Twelve 222](#_Toc133500794)

[Table 8.F.1 Reliability Estimates by Student Group for Kindergarten 224](#_Toc133500795)

[Table 8.F.2 Reliability Estimates by Student Group for Grade One 225](#_Toc133500796)

[Table 8.F.3 Reliability Estimates by Student Group for Grade Two 226](#_Toc133500797)

[Table 8.F.4 Reliability Estimates by Student Group for Grade Span Three Through Five 227](#_Toc133500798)

[Table 8.F.5 Reliability Estimates by Student Group for Grade Span Six Through Eight 228](#_Toc133500799)

[Table 8.F.6 Reliability Estimates by Student Group for Grade Span Nine and Ten 229](#_Toc133500800)

[Table 8.F.7 Reliability Estimates by Student Group for Grade Span Eleven and Twelve 230](#_Toc133500801)

[Table 8.F.8 Classification Accuracy and Consistency 231](#_Toc133500802)

[Table 8.G.1 Raw-to-Scale-Score Conversion Table for Kindergarten 232](#_Toc133500803)

[Table 8.G.2 Raw-to-Scale-Score Conversion Table for Grade One 233](#_Toc133500804)

[Table 8.G.3 Raw-to-Scale-Score Conversion Table for Grade Two 234](#_Toc133500805)

[Table 8.G.4 Raw-to-Scale-Score Conversion Table for Grade Span Three Through Five Form One 235](#_Toc133500806)

[Table 8.G.5 Raw-to-Scale-Score Conversion Table for Grade Span Three Through Five Form Two 236](#_Toc133500807)

[Table 8.G.6 Raw-to-Scale-Score Conversion Table for Grade Span Six Through Eight Form One 237](#_Toc133500808)

[Table 8.G.7 Raw-to-Scale-Score Conversion Table for Grade Span Six Through Eight Form Two 238](#_Toc133500809)

[Table 8.G.8 Raw-to-Scale-Score Conversion Table for Grade Span Nine and Ten Form   
One 239](#_Toc133500810)

[Table 8.G.9 Raw-to-Scale-Score Conversion Table for Grade Span Nine and Ten Form   
Two 240](#_Toc133500811)

[Table 8.G.10 Raw-to-Scale-Score Conversion Table for Grade Span Eleven and Twelve Form One 241](#_Toc133500812)

[Table 8.G.11 Raw-to-Scale-Score Conversion Table for Grade Span Eleven and Twelve Form Two 242](#_Toc133500813)

[Table 8.H.1 Interrater Reliability for Kindergarten 243](#_Toc133500814)

[Table 8.H.2 Interrater Reliability for Grade One 243](#_Toc133500815)

[Table 8.H.3 Interrater Reliability for Grade Two 244](#_Toc133500816)

[Table 8.H.4 Interrater Reliability for Grade Span Three Through Five 244](#_Toc133500817)

[Table 8.H.5 Interrater Reliability for Grade Span Six Through Eight 245](#_Toc133500818)

[Table 8.H.6 Interrater Reliability for Grade Span Nine Through Twelve 245](#_Toc133500819)

[Table 10.A.1 Responses to Question 1 for Kindergarten 257](#_Toc133500820)

[Table 10.A.2 Responses to Question 1 for Grade One 257](#_Toc133500821)

[Table 10.A.3 Responses to Question 1 for Grade Two 258](#_Toc133500822)

[Table 10.A.4 Responses to Question 1 for Grade Span Three Through Five 258](#_Toc133500823)

[Table 10.A.5 Responses to Question 1 for Grade Span Six Through Eight 259](#_Toc133500824)

[Table 10.A.6 Responses to Question 1 for Grade Span Nine and Ten 259](#_Toc133500825)

[Table 10.A.7 Responses to Question 1 for Grade Span Eleven and Twelve 260](#_Toc133500826)

[Table 10.A.8 Scale Score Summary by Response to English Language Proficiency Question for Kindergarten 260](#_Toc133500827)

[Table 10.A.9 Scale Score Summary by Response to English Language Proficiency Question for Grade One 261](#_Toc133500828)

[Table 10.A.10 Scale Score Summary by Response to English Language Proficiency Question for Grade Two 261](#_Toc133500829)

[Table 10.A.11 Scale Score Summary by Response to English Language Proficiency Question for Grade Span Three Through Five 262](#_Toc133500830)

[Table 10.A.12 Scale Score Summary by Response to English Language Proficiency Question for Grade Span Six Through Eight 262](#_Toc133500831)

[Table 10.A.13 Scale Score Summary by Response to English Language Proficiency Question for Grade Span Nine and Ten 263](#_Toc133500832)

[Table 10.A.14 Scale Score Summary by Response to English Language Proficiency Question for Grade Span Eleven and Twelve 263](#_Toc133500833)

[Table 10.A.15 Responses to Question 2 for Kindergarten 264](#_Toc133500834)

[Table 10.A.16 Responses to Question 2 for Grade One 264](#_Toc133500835)

[Table 10.A.17 Responses to Question 2 for Grade Two 264](#_Toc133500836)

[Table 10.A.18 Responses to Question 2 for Grade Span Three Through Five 265](#_Toc133500837)

[Table 10.A.19 Responses to Question 2 for Grade Span Six Through Eight 265](#_Toc133500838)

[Table 10.A.20 Responses to Question 2 for Grade Span Nine and Ten 265](#_Toc133500839)

[Table 10.A.21 Responses to Question 2 for Grade Span Eleven and Twelve 266](#_Toc133500840)

[Table 10.A.22 Responses to Question 3 for Kindergarten 266](#_Toc133500841)

[Table 10.A.23 Responses to Question 3 for Grade One 267](#_Toc133500842)

[Table 10.A.24 Responses to Question 3 for Grade Two 268](#_Toc133500843)

[Table 10.A.25 Responses to Question 3 for Grade Span Three Through Five 268](#_Toc133500844)

[Table 10.A.26 Responses to Question 3 for Grade Span Six Through Eight 269](#_Toc133500845)

[Table 10.A.27 Responses to Question 3 for Grade Span Nine and Ten 270](#_Toc133500846)

[Table 10.A.28 Responses to Question 3 for Grade Span Eleven and Twelve 271](#_Toc133500847)

[Table 10.A.29 Responses to Question 4 for Kindergarten 272](#_Toc133500848)

[Table 10.A.30 Responses to Question 4 for Grade One 272](#_Toc133500849)

[Table 10.A.31 Responses to Question 4 for Grade Two 273](#_Toc133500850)

[Table 10.A.32 Responses to Question 4 for Grade Span Three Through Five 273](#_Toc133500851)

[Table 10.A.33 Responses to Question 4 for Grade Span Six Through Eight 274](#_Toc133500852)

[Table 10.A.34 Responses to Question 4 for Grade Span Nine and Ten 274](#_Toc133500853)

[Table 10.A.35 Responses to Question 4 for Grade Span Eleven and Twelve 275](#_Toc133500854)

[Table 10.A.36 Responses to Question 5 for Kindergarten 275](#_Toc133500855)

[Table 10.A.37 Responses to Question 5 for Grade One 276](#_Toc133500856)

[Table 10.A.38 Responses to Question 5 for Grade Two 276](#_Toc133500857)

[Table 10.A.39 Responses to Question 5 for Grade Span Three Through Five 277](#_Toc133500858)

[Table 10.A.40 Responses to Question 5 for Grade Span Six Through Eight 277](#_Toc133500859)

[Table 10.A.41 Responses to Question 5 for Grade Span Nine and Ten 278](#_Toc133500860)

[Table 10.A.42 Responses to Question 5 for Grade Span Eleven and Twelve 278](#_Toc133500861)

[Table 10.A.43 Responses to Question 6 for Kindergarten 279](#_Toc133500862)

[Table 10.A.44 Responses to Question 6 for Grade One 280](#_Toc133500863)

[Table 10.A.45 Responses to Question 6 for Grade Two 281](#_Toc133500864)

[Table 10.A.46 Responses to Question 6 for Grade Span Three Through Five 282](#_Toc133500865)

[Table 10.A.47 Responses to Question 6 for Grade Span Six Through Eight 283](#_Toc133500866)

[Table 10.A.48 Responses to Question 6 for Grade Span Nine and Ten 284](#_Toc133500867)

[Table 10.A.49 Responses to Question 6 for Grade Span Eleven and Twelve 285](#_Toc133500868)

[Table 10.A.50 Responses to Question 7 for Kindergarten 286](#_Toc133500869)

[Table 10.A.51 Responses to Question 7 for Grade One 287](#_Toc133500870)

[Table 10.A.52 Responses to Question 7 for Grade Two 288](#_Toc133500871)

[Table 10.A.53 Responses to Question 7 for Grade Span Three Through Five 289](#_Toc133500872)

[Table 10.A.54 Responses to Question 7 for Grade Span Six Through Eight 290](#_Toc133500873)

[Table 10.A.55 Responses to Question 7 for Grade Span Nine and Ten 291](#_Toc133500874)

[Table 10.A.56 Responses to Question 7 for Grade Span Eleven and Twelve 292](#_Toc133500875)

[Table 10.A.57 Responses to Question 8 for Kindergarten 292](#_Toc133500876)

[Table 10.A.58 Responses to Question 8 for Grade One 293](#_Toc133500877)

[Table 10.A.59 Responses to Question 8 for Grade Two 293](#_Toc133500878)

[Table 10.A.60 Responses to Question 8 for Grade Span Three Through Five 293](#_Toc133500879)

[Table 10.A.61 Responses to Question 8 for Grade Span Six Through Eight 294](#_Toc133500880)

[Table 10.A.62 Responses to Question 8 for Grade Span Nine and Ten 294](#_Toc133500881)

[Table 10.A.63 Responses to Question 8 for Grade Span Eleven and Twelve 294](#_Toc133500882)

[Table 10.A.64 Responses to Question 9 for Kindergarten 295](#_Toc133500883)

[Table 10.A.65 Responses to Question 9 for Grade One 296](#_Toc133500884)

[Table 10.A.66 Responses to Question 9 for Grade Two 297](#_Toc133500885)

[Table 10.A.67 Responses to Question 9 for Grade Span Three Through Five 298](#_Toc133500886)

[Table 10.A.68 Responses to Question 9 for Grade Span Six Through Eight 299](#_Toc133500887)

[Table 10.A.69 Responses to Question 9 for Grade Span Nine and Ten 300](#_Toc133500888)

[Table 10.A.70 Responses to Question 9 for Grade Span Eleven and Twelve 301](#_Toc133500889)

[Table 10.A.71 Responses to Question 10 for Kindergarten 302](#_Toc133500890)

[Table 10.A.72 Responses to Question 10 for Grade One 302](#_Toc133500891)

[Table 10.A.73 Responses to Question 10 for Grade Two 303](#_Toc133500892)

[Table 10.A.74 Responses to Question 10 for Grade Span Three Through Five 303](#_Toc133500893)

[Table 10.A.75 Responses to Question 10 for Grade Span Six Through Eight 304](#_Toc133500894)

[Table 10.A.76 Responses to Question 10 for Grade Span Nine and Ten 304](#_Toc133500895)

[Table 10.A.77 Responses to Question 10 for Grade Span Eleven and Twelve 305](#_Toc133500896)

List of Figures

[Figure 6.1 Location of threshold students 69](#_Toc133488430)

[Figure 7.1 Percentage of students at each performance level 78](#_Toc133488431)

[Figure 8.1 The path diagram for the correlated two-factor model 142](#_Toc133488432)

[Figure 8.2 The path diagram for a bifactor model 143](#_Toc133488433)

[Figure 8.3 The path diagram for a one-factor model 144](#_Toc133488434)

[Figure 8.E.1 Factor loading comparison for kindergarten 209](#_Toc133488435)

[Figure 8.E.2 Factor loading comparison for grade one 211](#_Toc133488436)

[Figure 8.E.3 Factor loading comparison for grade two 213](#_Toc133488437)

[Figure 8.E.4 Factor loading comparison for grade span three through five 215](#_Toc133488438)

[Figure 8.E.5 Factor loading comparison for grade span six through eight 218](#_Toc133488439)

[Figure 8.E.6 Factor loading comparison for high school 221](#_Toc133488440)

Acronyms and Initialisms Used in the *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 |
| ALTRD | Assessment and Learning Technology Research & Development |
| APA | American Psychological Association |
| AST | Administration and Scoring Training |
| CA CCSS | California Common Core State Standards |
| CAA | California Alternate Assessment |
| 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* |
| CDE | California Department of Education |
| CDS | county/district/school |
| CERS | California Educator Reporting System |
| CFA | confirmatory factor analysis |
| CI | confidence interval |
| COVID-19 | novel coronavirus disease 2019 |
| CR | constructed response |
| CSEM | conditional standard error of measurement |
| DEI | Data Entry Interface |
| *DFA* | *Directions for Administration* |
| DIF | differential item functioning |
| DRM | data review meeting |
| *EC* | *Education Code* |
| ECV | explained common variance |
| EL | English learner |
| ELA | English language arts/literacy |
| ELP | English language proficiency |
| ELPAC | English Language Proficiency Assessments for California |
| EO | English only |
| eSKM | Enterprise Score Key Management |
| GPCM | generalized partial credit model |
| HLS | home language survey |
| HOSS | highest obtainable scale score |
| IEP | individualized education program |
| IFEP | initial fluent English proficient |

Table of Acronyms and Initialisms *(continuation one)*

|  |  |
| --- | --- |
| Term | Definition |
| IRM | item review meeting |
| IRT | item response theory |
| ISAAP | Individual Student Assessment Accessibility Profile |
| IWW | item writer workshop |
| K | kindergarten |
| K–2 | kindergarten through grade two |
| LDAA | locally determined alternate assessment |
| LEA | local educational agency |
| LOSS | lowest obtainable scale score |
| MC | multiple choice |
| MH | Mantel-Haenszel |
| MH-DIF | Mantel-Haenszel differential item functioning |
| MIRT | multidimensional item response theory |
| NCME | National Council on Measurement in Education |
| OTI | Office of Testing Integrity |
| PAR | Psychometric Analysis & Research |
| PLD | performance level descriptor |
| PPT | paper–pencil test |
| QA | quality assurance |
| QWK | quadratic-weighted kappa |
| RFEP | reclassified fluent English proficient |
| RPB | relative parameter bias |
| SBE | State Board of Education |
| SCOE | Sacramento County Office of Education |
| SD | standard deviation |
| SEM | standard error of measurement |
| SFTP | secure file transfer protocol |
| SMD | standardized mean difference |
| SR | selected response |
| SSID | Statewide Student Identifier |
| SSPI | State Superintendent of Public Instruction |
| SSR | Student Score Report |
| STAIRS | Security and Test Administration Incident Reporting System |
| TAG | Technical Advisory Group |
| TCC | test characteristic curve |
| TDS | test delivery system |
| TIF | test information function |
| TIP | Test Item Preview |
| TOMS | Test Operations Management System |

Table of Acronyms and Initialisms *(continuation two)*

|  |  |
| --- | --- |
| Term | Definition |
| UAT | user acceptance testing |
| USC | United States Code |

## Introduction

This chapter provides an introduction to the Alternate English Language Proficiency Assessments for California (ELPAC), including background information, purpose of the test, test content, intended population, testing windows, organizations and systems involved, and an overview of the technical report.

### Background

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” (CDE, 2022b) 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 administer a state test of ELP to eligible students in kindergarten through grade twelve (California Department of Education [CDE], 2022b). California *Education Code (EC)* Section 313(a) requires that the assessment of ELP be done upon initial enrollment and annually thereafter until the LEA reclassifies the student as fluent English proficient.

### Test Purposes

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. |

Table 1.1 *(continuation)*

|  |  |
| --- | --- |
| Initial Alternate ELPAC | Summative Alternate ELPAC |
| 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 test one time only. The Initial Alternate ELPAC is taken before the Summative Alternate ELPAC or Summative ELPAC, if the student is identified as an EL. | A student takes this test 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 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. 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 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 Alternate ELPAC is intended only for EL students, and potential EL students, in kindergarten through grade twelve (up to age twenty-one) who have been identified as having the most significant cognitive disabilities and who have been found eligible for alternate assessments by their IEP team.

In a typical year, students in kindergarten through grade twelve (up to age twenty-one) 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. However, as the Initial Alternate ELPAC was developed after the operational field test administration of the Summative Alternate ELPAC, the students who participated in the operational field test did not take the Initial Alternate ELPAC beforehand; rather, they took a locally determined alternate assessment or the general Initial ELPAC to verify their ELP.

Students in kindergarten through grade twelve (up to age twenty-one) who are classified as EL students and have been found eligible for alternate assessments by their IEP team must be administered the Summative Alternate ELPAC annually during the Summative Alternate ELPAC testing window until they are reclassified as fluent English proficient based on the CDE’s established guidelines for reclassification established by the California State Board of Education (SBE) (*EC*313[f]).

### Testing Windows and Estimated Testing Times

The Summative Alternate ELPAC testing window generally runs from February 1 through May 31 annually. However, because Initial Alternate ELPAC forms assembly was set to begin in March of 2022, the start of the 2021–22 field test administration testing window was pushed earlier, to November 1, 2021. This was done in an attempt to increase the number of tests completed by February 15, 2022, the date by which item analysis was to begin so that analysis could be completed in time for the items to be used operationally on the Initial Alternate ELPAC forms. The goal was to have as many tests completed by February 15 as possible to gather the most data. Because LEAs were heavily impacted by the school and business shutdowns that resulted from the novel coronavirus disease 2019 (COVID-19) pandemic, the testing window was extended through May 31, 2022, to ensure that all eligible students had an opportunity to participate in the test during the typical annual Summative Alternate ELPAC testing window.

The Alternate ELPAC is a computer-based, nonadaptive, untimed test 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 test may be administered over the course of multiple days. The estimated testing times for the Alternate ELPAC task types were posted by task type in the *Alternate ELPAC Operational Field Test Manual* (CDE, 2022a). Estimated testing times were provided for administration planning only.

### Groups and Organizations Involved with the Alternate ELPAC

#### California State Board of Education

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

In addition to adopting the rules and regulations for itself, its appointees, and California’s public schools, the SBE 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, 2022b).

#### 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 10,500 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, 2022c).

#### 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 Alternate ELPAC.

##### Primary Testing Contractor—ETS

The CDE and the SBE contract with ETS to develop, administer, and report the 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
* Supporting and training county offices of education, LEAs, and direct funded charter schools
* Constructing, producing, and controlling the quality of 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 (CERS)
* 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
* Developing high-quality items that are aligned to the 2012 ELD Standards via the ELD Connectors

##### 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;
* hosting and providing support for the DEI, the web browser–based application that, for the operational administration of the Alternate ELPAC, allows test examiners or data entry staff to enter second scores for the Alternate ELPAC.
* 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 set up the LEA’s test administration window; a test administrator or test examiner cannot download student reports. A description of user roles is explained more extensively in the *Alternate ELPAC Operational Field Test Manual* (CDE, 2022a).

#### 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 tests and monitor student testing;
* the Student Testing Interface, on which students take the test 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.

#### Practice and Training Tests

All California testing programs have practice and training tests to inform educators, parents/‌guardians, and students about the individual tests. The practice and training tests were provided to LEAs to prepare students and LEA staff for administration of the Alternate ELPAC. These tests simulated the experience of the Alternate ELPAC computer-based assessments. Unlike the summative assessments, the practice and training tests did not gauge student success on the operational test, or produce scores. Students, teachers, and the public could access them using a web browser, although accessing them through the secure browser permitted students to take the tests using the text-to-speech embedded accommodation and to test assistive technology.

The purposes of the training tests are to

* allow students and test examiners to quickly become familiar with the user interface and components of the TDS as well as with the process of starting and completing a testing session;
* allow students and test examiners to experience a grade-level assessment, grade-specific items and difficulty levels, and the format and structure of an operational assessment; and
* provide an opportunity for educators to assign embedded designated supports and accommodations and determine how they worked for their students prior to using the resources in an operational test setting.

#### California Educator Reporting System

CERS is the system used by LEAs to view preliminary student results from ELPAC testing. The primary purpose of CERS is to provide educators and administrators with access to timely test results data for individual students and groups of students.

CERS allows educators to view their students’ assessment results at the individual student level and at the aggregated level using grouping and other features. For example, educators can create customized groups from assigned student groups based on demographic information, performance level, or other characteristics of their choosing. The student results sent to CERS are appropriate for analysis of assessment results for use in informing instruction.

#### 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 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 special education services status) and for the total population.

### Overview of the Technical Report

This technical report addresses the characteristics of the 2021–22 administration of the 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 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 process of test assembly, including the content being measured and the content and psychometric criteria.
* [Chapter 5](#_Test_Administration) details the processes involved in the actual 2021–22 administration, with emphasis on efforts made to ensure the standardization of 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](#_Analyses_and_Results) summarizes the statistical procedures conducted for the 2021–22 operational field test administration. These analyses include
* classical item analyses,
* differential item functioning analyses,
* item response theory analyses, and
* dimensionality 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 Alternate ELPAC, including item development, test form development, test administration, scoring procedures, psychometric analysis, and score reporting.
* [Chapter 10](#_In-Test_Survey_1) describes the development and administration of an in-test survey for test examiners and the results from the analyses of the responses.
* [Chapter 11](#_Continuous_and_Systematic) describes analysis and administration processes and features targeted for improvement during future test administrations.

### References

California Department of Education. (2022a). *Alternate ELPAC operational field test manual.* Sacramento, CA: California Department of Education.

California Department of Education. (2022b, October). *English Language Proficiency Assessments for California (ELPAC).* California Department of Education website.

California Department of Education. (2022c, August). *Organization.* California Department of Education website.

California Department of Education. (2022d, October). *State Board of Education responsibilities.* California Department of Education website.

## Overview of Alternate ELPAC Processes

This chapter provides an overview of the processes implemented by ETS during a typical, full testing cycle for the 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 questions for the Alternate ELPAC in accordance with the *ETS* *Standards for Quality and Fairness* (2014).

#### Selection of ELD Connectors for Development

The 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 (California Department of Education [CDE], 2014 and 2019).

The development of the ELD Connectors began with a review of the Council of State School Officers (CCSSO) ELP Standards for 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 are ELD Connectors for each of the 2012 ELD Standards assessed on the ELPAC at each of the Alternate ELPAC grade levels and grade spans (i.e., kindergarten, grades one, two, three through five, six through eight, nine and ten, and eleven and twelve). The development of the ELD Connectors is consistent with the approach used for the California Alternate Assessments for English language art/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. The sample ELD Connectors were revised based on 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 done 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 Alternate ELPAC includes the following primary computer-based item formats:

* **Selected-response (SR) items—**Students are instructed to select one or more choices. Most Alternate ELPAC items have two or three options. The items are assigned one point and are machine-scored.
* **Constructed response (CR) 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 two points, one point, or zero points on these items.

#### Task Type Specifications

The *Alternate ELPAC Task Type Specifications* provide descriptions of the types of tasks contained within the assessment. Each task type is intended to measure specific 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.

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 2021–22 operational field test assessment was assembled in accordance with the Alternate ELPAC blueprint, which was approved by the SBE in May 2020 (CDE, 2020b).

The assembly began with the selection of seven task types, each with two or four items. The task types on the test appeared in sequential order according to the blueprint. Since the 2021–22 test administration was an operational field test, all items on the test were items that had not been previously administered and did not have previous statistics. There was a slight variance from the blueprint for the 2021–22 operational field test administration in that more items were field-tested than specified in the operational blueprint.

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

After test 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 collaboratively with the CDE to make revisions to the test forms until the CDE approved the test forms.

#### Test Blueprint

The Alternate ELPAC test blueprint is unique to each grade level or grade span. The test blueprint designates the task types and the sequence of task types to appear on a test. The blueprint also specifies the number of items and points for each task type and the total number of items and points on each test according to standards (CDE, 2020b). The blueprint for the Summative Alternate ELPAC was approved by the SBE in May 2020.

#### Test Length

The number of items in the Alternate ELPAC is the same across grade levels—the blueprint indicates there are 24 operational items and 6 embedded field test items, for a total of 30 items. For the 2021–22 operational field test administration, each student was administered 34 field test items as test developers embedded an overage of items and task types to allow for the possibility of item loss due to potentially poor statistical performance.

Refer to [*Chapter 4: Test Assembly*](#_Toc122102494) for more details on test form assembly.

### Test Administration

The 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 Alternate ELPAC can be found in [*Chapter 5: Test Administration*](#_Test_Administration).

#### Test Security and Confidentiality

All operational tests within the ELPAC System are secure. For the Alternate ELPAC administration, every person having access to test materials maintained the security and confidentiality of the tests. ETS’ internal Code of Ethics requires that all test information, including tangible materials (such as test questions and test results), confidential files, processes, and activities were kept secure. To ensure security for all tests 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_1) in [*Chapter 5: Test Administration*](#_Test_Administration).

In the pursuit of enforcing secure practices, ETS strives to safeguard the various processes involved in a test 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 Alternate ELPAC.

The 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 Alternate ELPAC administration include LEA ELPAC coordinators, site ELPAC coordinators, and test examiners. The responsibilities of each of the staff members are described in the *Alternate ELPAC Operational Field Test Manual* (CDE, 2022a).

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:

* ***Alternate ELPAC Operational Field Test Manual*—**This web-based manual provides test administration procedures and guidelines for LEA ELPAC coordinators and site ELPAC coordinators (CDE, 2022a). (Refer to [*5.3.4.2 Alternate ELPAC Operational Field Test Manual*](#_Alternate_ELPAC_Operational) in [chapter 5](#_Test_Administration) for more information.)
* ***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 setting up test administrations, adding and managing users, assigning tests, and configuring computer-based student test settings (CDE, 2022b). (Refer to [*5.3.4.3 CAASPP and ELPAC Test Operations Management System User Guide*](#_CAASPP_and_ELPAC_1) in [chapter 5](#_Test_Administration) for more information.)
* ***DFA*s—**These directions include test examiner directions and scripts for administering the tests. They contain grade-specific and form-specific information needed by the test examiners during test sessions.

### Fairness and Accessibility

Several procedures are in place to ensure that the Alternate ELPAC is fair and accessible to all test takers. 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.

#### Universal Tools, Designated Supports, and Accommodations

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, 2021).

**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 only on approval by the CDE.

Table 5.3 presents counts and percentages of students assigned designated supports, accommodations, and unlisted resources for the 2021–22 Alternate ELPAC administration. Table 5.3 was created using student demographic data in the end-of-year production data file updated on August 31, 2022.

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

#### Description of Differential Item Functioning Analyses

Differential item functioning (DIF) analyses are conducted to detect possible test bias by locating items for which one group of students performs significantly better than another group. DIF is a collection of statistical methods used to recognize whether performance varies across different groups of examinees (e.g., male vs. female or White vs. Black or African American). If an item performed differentially across student groups, even when students were matched on ability, the item may be measuring something other than the intended construct. Therefore, it is important to identify items flagged for DIF. Content experts and bias and sensitivity experts from diverse backgrounds reviewed these DIF-flagged items to determine the potential sources and meanings of performance differences. Refer to section [*8.3 Differential Item Functioning Analyses*](#_Differential_Item_Functioning) for additional information about DIF.

### Scores

Individual student scores were reported for the 2021–23 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).

#### Estimating Ability Scores

The item response theory (IRT) inverse test characteristic curve method (Stocking, 1996)—where the student’s ability value is estimated to be the value for which the expected number-correct score is equal to the student’s number-correct score—was used to estimate students’ overall ability parameters. For the purpose of reporting, students’ ability estimates (theta scores) were then expressed in three-digit scale scores by applying the appropriate linear transformation for each grade-level Alternate ELPAC.

Student performance on the reporting scale was designated into one of three levels:

1. Level 1—Novice English Learner
2. Level 2—Intermediate English Learner
3. Level 3—Fluent English Proficient

For information regarding score specifications and the establishment of score-reporting scales, refer to [*Chapter 7: Scoring and Reporting*](#_Scoring_and_Reporting)*.* For information regarding Alternate ELPAC performance levels, refer to [*Chapter 6: Standard Setting*](#_Appendix_4.A:_Demographic) for a description of the process used to set performance-level standards.

#### 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.

Alternate ELPAC scores can also be viewed through the California Educator Reporting System (CERS), a secure website that provided authorized users with interactive and cumulative online reports for scale scores and performance levels at the student, school, and LEA levels. CERS also provided individual score reports. Refer to subsection [*7.3.1 Online Reporting*](#_Online_Reporting_1) for details about TOMS and CERS and subsection [*7.3.3 Types of Score Reports*](#_Types_of_Score_1) for the content of each type of score report.

#### Aggregation Procedures

To provide meaningful results to interested educators, 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 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 [appendix 7.B](#_Appendix_7.B:_Means_1). In table 7.B.1 through table 7.B.7, students are grouped by demographic groups, including gender, ethnicity, English language fluency, special education service status, and economic status, as well as crosstab analysis for ethnicity and economic status. The tables show 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 tables as “N/A” when the number of students in the sample is 10 or fewer. Definitions for the demographic student groups included in these tables are provided in table 7.4.

### Psychometric Analyses

Psychometric analyses were conducted on the data from the Alternate ELPAC, including classical item analyses, DIF analyses, IRT calibration and linking, testing time analyses, and reliability analyses. The results of these analyses support understanding of item performances 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*](#_Psychometric_Analyses).

#### Description of the Classical Item Analyses

The psychometric analyses for the Alternate ELPAC data included classical item analyses and DIF analyses to evaluate the performance of the operational items and the embedded field test 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)*.* A description of the DIF analyses procedure is provided in section [*8.3 Differential Item Functioning Analyses*](#_Differential_Item_Functioning)*.*

#### 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 models and the procedures for the calibration and linking analyses are included in section [*8.4 Item Response Theory Analyses*](#_Item_Response_Theory).

### References

California Department of Education. (2014). *2012 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. (2021). *California assessment accessibility resources matrix*. California Department of Education website.

California Department of Education. (2022a). *Alternate ELPAC operational field test manual.* Sacramento, CA: California Department of Education.

California Department of Education. (2022b). *CAASPP and ELPAC Test Operations Management System user guide*. 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.

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

## Item Development and Review

This chapter discusses the detailed procedures of item development for the 2021–22 Alternate English Language Proficiency Assessments for California (ELPAC) administration.

### Overview

In partnership with the Sacramento County Office of Education (SCOE), ETS convened Alternate ELPAC item writer trainings and item review meetings (IRMs) to develop test items for the Alternate ELPAC. In addition, ETS trained a small group of experienced contractors and select California educators to draft Alternate ELPAC items. After the items went through ETS internal and California Department of Education (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. Some of these items were used as embedded field test items on the 2021–22 Alternate ELPAC operational field test. Additionally, an IRM was conducted in April 2021; some of those items will be used as embedded field test items on future Summative Alternate ELPAC forms for eventual operational use.

#### Preparation

To prepare for the 2021–22 Alternate ELPAC, several test design tasks were conducted. These tasks occurred prior to item development and test development tasks. The Alternate ELPAC test blueprint was reviewed, a high-level test design was developed, a pilot test using cognitive lab methodology was conducted, and 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 Alternate ELPAC, provide guidance to item writers, drive consistency and efficiency in item development, provide accessibility considerations, and provide the CDE with a reference guide to use while reviewing items. The 2021–22 Alternate ELPAC contained 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 aligned 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). Items may be linked to multiple standards.

The 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 domain. Because of the intent to assign an overall score and to provide students with the flexibility to use their individually preferred communication modes, 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 multiple choice (MC).

##### 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 MC or constructed response.

#### Recruitment and Selection of Item Writers

California educators were recruited through the Educator Opportunities Portal and email communications. To ensure broad representation, the CDE sent email messages announcing the opportunities to write items and to review items to various groups, including but not limited to:

* The CDE Assessment Spotlight listserv (This includes local educational agency [LEA] ELPAC coordinators, interest holders who subscribe, and Title III county leads.)
* Title III LEAs using the CDE Multilingual Updates Listserv
* The Bilingual Coordinators Network

Applications were solicited at various interest holder conferences and meetings, including but not limited to:

* California Association for Bilingual Education
* Bilingual Coordinators Network
* California Assessment Conference
* Special Education Local Plan Areas
* Special Education Administrators of County Offices

The email and letter directed applicants to submit an online application through the Educator Opportunities Portal. The application allowed California educators to apply for any number of the listed events. The information from the application was loaded into a database that was used for the review and selection process.

Applications were selected from current and retired California educators who met the following minimum qualifications:

* Possession of a bachelor’s or higher degree
* Expertise in language acquisition or experience teaching English learner (EL) students in kindergarten through grade twelve
* Experience teaching students with severe cognitive disabilities
* Knowledge of, and experience working with, the 2012 ELD Standards and the ELD Connectors
* Knowledge of, and experience working with, students with exceptionalities such as visual impairment, hearing impairment, deafness, or speech or language disorders

Additional desirable qualifications included the following:

* Specialized teaching certification in reading (e.g., Reading Certificate or Reading and Language Arts Specialist Certificate)
* Specialized teaching certification in special education
* Experience writing or reviewing items for standardized tests, especially tests for EL students in kindergarten through grade twelve
* Experience administering the California Alternate Assessments, general ELPAC, or Alternate ELPAC pilot

Selections were made to ensure representation from different cultural and linguistic groups, LEAs and county offices of education of various sizes, and different geographical regions of the state. ETS and SCOE made preliminary selections, which were reviewed by the CDE, adjusted as needed, and then approved. Fifteen educators participated in the June 2020 IRMs (Content Review Meeting and Bias and Sensitivity Meeting), and 10 educators were trained to write items during the item writer workshop (IWW) in June 2021.

Table 3.1 shows the self-reported educational qualifications, present occupation, and credentials of the individuals who participated in an Alternate ELPAC item writer training or IRM.

Table 3.1 Alternate ELPAC IWW and IRM Participant Qualifications by Meeting Type and Total

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Qualification Type | Qualification | 2020 IWW | 2020 IRM | 2021 IWW | 2021 IRM | Total |
| Occupation | Classroom Teacher, Teacher on Special Assignment, or Special Education Teacher | 9 | 8 | 7 | 13 | 37 |
| Occupation | EL or Literacy Coach | 5 | 3 | 0 | 0 | 8 |
| Occupation | Teacher of the Visually Impaired | 0 | 1 | 0 | 1 | 2 |
| Occupation | LEA or County Office Employee | 3 | 3 | 3 | 5 | 14 |
| Occupation | Speech Therapist/Pathologist | 1 | 0 | 0 | 0 | 1 |
| Highest Degree Earned | Bachelor’s Degree | 3 | 3 | 2 | 4 | 12 |
| Highest Degree Earned | Master’s Degree | 14 | 11 | 7 | 13 | 45 |
| Highest Degree Earned | Doctorate | 1 | 1 | 1 | 3 | 6 |
| K–12 Teaching Credential | Elementary Teaching (multiple subjects) | 6 | 4 | 3 | 3 | 16 |
| K–12 Teaching Credential | Secondary Teaching (single subject) | 4 | 3 | 1 | 5 | 13 |
| K–12 Teaching Credential | Special Education or Education Specialist | 7 | 10 | 9 | 13 | 39 |

Table 3.1 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Qualification Type | Qualification | 2020 IWW | 2020 IRM | 2021 IWW | 2021 IRM | Total |
| K–12 Teaching Credential | Reading Specialist | 0 | 1 | 0 | 0 | 1 |
| K–12 Teaching Credential | EL (Crosscultural, Language and Academic Development; Bilingual, Crosscultural, Language and Academic Development) | 7 | 7 | 3 | 0 | 17 |
| K–12 Teaching Credential | Administrative | 1 | 2 | 2 | 3 | 8 |
| K–12 Teaching Credential | Visual Impairment (student experience) | 4 | 3 | 3 | 6 | 16 |

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

SCOE contacted and invited the participants, as well as contacted the alternates as necessary (when confirmed participants cancelled and there was sufficient time to fill the opening). Once all participants were confirmed, SCOE notified those who were not selected.

#### Item Writing by Contractors

In 2020, 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 Alternate ELPAC, ETS recruited and trained educators in the 2012 ELD Standards.

The three primary goals for the training were to

1. provide teachers with knowledge, via professional development on writing items, 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.

### ETS Item Review Process

After items were drafted, ETS placed items 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).

All items were entered into the Item Banking Information System (IBIS) with corresponding artwork and metadata. Within IBIS, items received ETS internal content, fairness, and editorial reviews.

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 content-area 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, content-area assessment specialists conducted three reviews on items and stimuli. These assessment specialists verified thatthe items 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 *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 test
* 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 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.

#### 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 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 questions 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 and 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 awareness of, and 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 EL students with significant cognitive disabilities

##### Cognitive Laboratory

Under the direction of the CDE, ETS conducted a pilot using cognitive laboratory methodology across California in January 2020 (CDE, 2020a). One of the goals of the study was to collect evidence of the use of accessibility resources for test administration and determine whether the test design supports access.

Based on results of the study, minor adjustments to the task types were made, including providing considerations of alternate response options for expressive items to be more inclusive for students who are not verbal communicators. This includes students who are presymbolic communicators, who use few conventional symbols (words, signs, or pictures), or who use augmentative and alternative communication devices. Guidance on how to provide optional individualization has been added for rubric-scored expressive items that either allow for the use of real objects or provide picture card responses.

In addition, guidance was developed for increasing the number of objects and manipulatives for students who are presymbolic communicators. Guidance on how to provide optional individualization has been added for stories, passages, and items in which real objects or manipulatives could be provided to the student.

### 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 June. 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 have the option of making 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 for accepted with revisions, educators specify the revisions needed to improve the text or images and the reasons for the proposed revisions.

#### California Educators as Content Experts

During IRM meetings, 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 test positions, and are facilitated by ETS content experts. Typically, 15 to 20 educators are recruited to attend each meeting.

In the meetings, the item content, test examiner *Directions for Administration,* 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 2012 ELD Standards using the ELD Connectors. Meeting 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 level of the test taker. 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.

#### Composition of Item Review Panels

The group of California educators participating in the IRMs consisted of current and former teachers (some of whom had taught students eligible to take the Alternate ELPAC and others who were subject-matter experts), resource specialists, administrators, curriculum and content experts, and other education professionals. Minimum qualifications to be invited to participate were

* three or more years of teaching experience in kindergarten through grade twelve and in the relevant content area such as English language development or special education;
* bachelor’s or higher degree in a grade level or content area related to EL students and students with significant cognitive disabilities; and
* knowledge of, and experience with, the California 2012 ELD Standards via the ELD Connectors.

Preferred qualifications included:

* a special education credential,
* experience teaching students with more than one type of disability, and
* three to five years of experience as a special education teacher or school administrator with a special education credential.

School administrators; LEA, county content, or program specialists; or university educators had to meet the following qualifications to be invited to participate:

* Three or more years of experience as a school administrator; LEA, county content, or program specialist; or university instructor in a grade-specific area
* Bachelor’s or higher degree in a grade-specific area
* Knowledge of, and experience with, the California 2012 ELD Standards via the ELD Connectors

Every effort was made to ensure that each group of item reviewers included a wide representation of gender, geographic regions, and ethnic groups in California. Efforts also were made to ensure representation by members with experience serving California’s diverse special education population.

Item reviewers were recruited through an application process. Recommendations were solicited from LEAs and county offices of education as well as from the CDE. Applications were reviewed by ETS assessment directors, who confirmed that an applicant’s qualifications met the specified criteria. Applicants who met the criteria had their information forwarded to the CDE for further review and agreement before invitations to participate were distributed.

#### Meetings for Review of Alternate ELPAC Items

Table 3.2 provides the status of the items after the 2020 Alternate ELPAC IRM.

Table 3.2 Status of Items After the 2020 Alternate ELPAC IRM

|  |  |  |  |
| --- | --- | --- | --- |
| Grade Level or Grade Span | Accept As Is | Accept with Revisions | Rejected |
| Kindergarten | 14 | 98 | 0 |
| 1 | 27 | 85 | 1 |
| 2 | 55 | 50 | 0 |
| 3–5 | 39 | 67 | 0 |
| 6–8 | 50 | 59 | 0 |
| 9–12 | 77 | 70 | 0 |
| **Total:** | **262** | **429** | **1** |

Table 3.3 provides the status of the items after the 2021 Alternate ELPAC IRM.

Table 3.3 Status of Items After the 2021 Alternate ELPAC IRM

|  |  |  |  |
| --- | --- | --- | --- |
| Grade Level or Grade Span | Accept As Is | Accept with Revisions | Rejected |
| Kindergarten | 8 | 12 | 0 |
| 1 | 6 | 14 | 0 |
| 2 | 5 | 15 | 0 |
| 3–5 | 9 | 11 | 0 |
| 6–8 | 6 | 14 | 0 |
| 9–12 | 11 | 29 | 0 |
| **Total:** | **45** | **95** | **0** |

### Data Review Meeting

After the items were administered to students, ETS prepared the statistically flagged items and the associated statistics for review by the CDE and California educators. The data review meeting (DRM) for the 2021–22 Alternate ELPAC operational field test administration occurred from March 17 through March 18, 2022.

Review materials included items with their statistical data and statistical flags based on the respective administration’s item analyses along with annotated comment sheets for use by reviewers. Educators who were part of the DRM were assigned a training video in Upskill—a centralized, online location for training materials—giving them an overview of what is involved in a DRM as well as an understanding of the statistical measures used to review the items. This was followed by ETS conducting an introductory training at the beginning of the meeting to highlight any issues and to serve as a statistical refresher.

Reviewers then made decisions about which items should be included in the item bank for future operational forms assembly. Reviewers could determine whether items should be accepted as-is or rejected. Educators could also offer suggestions for rejected items to be revised, field-tested again, and put through another round of item analysis in a future review cycle. ETS psychometric and content staff were available to reviewers throughout this process.

ETS content staff facilitated the meeting and ensured that all educators gave input on whether there were any reasons to be concerned about the content of the flagged items. ETS psychometricians provided training on the interpretation of item statistics and responded to questions about the item statistics during the item discussion. The DRM participants reviewed the content and statistics of each item and then made a recommendation to accept or reject an item.

ETS content staff recorded the participants’ recommendations and comments regarding the flagged items. The feedback was referenced when working with the CDE to reconcile educator feedback and to make a final decision on whether to include particular flagged items in the operational pool. One item out of 252 items was rejected as a result of the DRM.

### References

California Department of Education. (2014). *2012 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.

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 2021–22 Alternate English Language Proficiency Assessments for California (ELPAC) administration.

### Overview

ETS assessment specialists assembled the Summative Alternate ELPAC operational field test, 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 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 assembly process is described in the field test specifications (CDE, 2020c).

The operational field test was a census field test, meaning all students eligible for the Summative Alternate ELPAC were required to take the test. The field test specifications detail the content characteristics, psychometric characteristics, and number of items to be used on the 2021–22 Alternate ELPAC operational field test. ETS created the test development specifications that the CDE reviewed and approved.

After the field test specifications were approved, ETS assessment specialists assembled the tests according to the specifications. 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.

### High-Level Test Design, Test Blueprint, and Form Assembly Specifications

The 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 test measures what it is intended to measure. It was also designed using universal design principles to ensure that it would be accessible to the testing population. All items and tasks are 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.

#### Test Blueprints

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 type and domains assessed by each task type (a given task type, which will assess both receptive and expressive domains),
* 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 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 results led to modifications of the Alternate ELPAC test blueprint. For example, all but one of the piloted task types was retained, which exceeded estimations. In addition, test design features were added to increase the accessibility of expressive (constructed response) items. Optional individualization was added to the test 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 six embedded field test items. The SBE approved the Alternate ELPAC test blueprints in May 2020, which was prior to the start of the operational field test administration of the Summative Alternate ELPAC on November 1, 2021.

The test blueprint provided information about the number of receptive and expressive items and points that were administered per task type within each grade level and grade span. The test blueprint also provided 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 that there was a moderate or partial match between the standard and the item in terms of language knowledge, skills, and abilities.

#### High-Level Test Design

##### Test Design Principles

Three principles guided the design of the Alternate ELPAC. The principles are based on discussions with, and feedback from, various interest holder groups and local educational agencies (LEAs) as well as the Alternate ELPAC Test Design Advisory Team. The guiding principles were as follows:

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.

##### Communication Modes

The 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 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). Each 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 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 Alternate ELPAC is administered one-on-one: one test examiner assessing one student at a time. LEAs are told to ensure the test examiner is someone who is familiar with the student and the student’s preferred communication mode(s). This facilitates a primary test-design feature: allowing a student to have the test administered in a way that provides access via an individually preferred communication mode(s). Communication modes include, but are not limited to, verbal communication and communication via sign language, eye gaze, facial expressions, gestures, a picture exchange system, and 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.

#### Form Assembly Specifications

The Summative Alternate ELPAC operational field test consisted of one test form for kindergarten, grade one, and grade two, and two test forms per grade span (grade span three through five, grade span six through eight, grade span nine and ten, and grade span eleven and twelve). Each test form contained 24 items, described in table 4.1, and 10 embedded field test items.

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 summative form.

Table 4.1 Kindergarten Through Grade Twelve Summative Alternate ELPAC Form 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** |

### Test Production Process

The high-level test design for the Alternate ELPAC describes the Alternate ELPAC as a computer-based, linear test (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 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. Such a blueprint design allows for the potential to include exit points after task types targeted to a low linguistic complexity level are presented. Exit points help to minimize the degree to which students at the early stages of language development are required to respond to items that are beyond their level of English proficiency.

#### Selection of Task Types and Items

From the eligible item pool, test 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

Each grade level and grade span of the 2021–22 Summative Alternate ELPAC had two forms created. Each form met the blueprint and contained an overage of items to account for the possible loss of items or task types. In October, just before the operational field test administration, the number of forms for kindergarten, grade one, and grade two were reduced to one because of unexpectedly low student registration numbers in those three grade levels.

Since the results of the operational field test item analyses would inform which items would be used in subsequent test administrations, the decision was made to reduce the number of forms for these grade levels from two forms to one form. This would help ensure that the n-counts for these grade levels would be sufficient in supporting item analysis outcomes. Though efforts were made to estimate the number of test takers for the Alternate ELPAC ahead of the operational field test administration, students in kindergarten through grade two had not been previously identified for alternate assessments, as the content assessments begin in grade three.

Table 4.1 in subsection [*4.2.3 Form Assembly Specifications*](#_Form_Assembly_Specifications) provides an overview of the number of operational items and points by task type and domain. The information contained in table 4.1 is true for all grade levels and grade spans for the Alternate ELPAC (kindergarten, grade one, grade two, grade span three through five, grade span six through eight, grade span nine and ten, and grade span eleven and twelve).

#### Psychometric Criteria and Review

All items for spring 2021–22 operational field test forms for each grade level and grade span were newly developed field-tested items. There were no statistics to review for the field test items when the forms were assembled. ETS psychometricians reviewed and confirmed that each test form was consistent with the field test specifications.

The following criteria will be used to review future operational forms:

* Forms align with the Alternate ELPAC test blueprints.
* 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. Such items need to be reviewed by the DIF panel and confirmed that they are not biased against any student group.
* Item response theory *b*-parameter estimates are within the range of -4.0 to +4.0, prior to the application of the vertical scale.
* Forms should have average item difficulty that is similar to the historical forms and can provide sufficient information around the threshold scores for the performance levels.

#### Content Review of Forms

After psychometric approval, the proposed assessment underwent two additional content reviews and one editorial review. The content reviewers were test 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
* Verification of coverage of the 2012 ELD Standards
* 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 test delivery system (TDS). Unlike other stages of the test production process, this stage must occur prior to every administration of the 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 looked 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 California Department of Education (CDE). The UAT review provided the CDE with an opportunity to interact with the exact test that would be administered to the students. The CDE had to approve the Alternate ELPAC UAT before the test 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 tests;
* the Student Testing Interface, on which students take the 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). *2012 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. (2020c). *Specifications for the Alternate English Language Proficiency Assessments for California operational field test.* [Unpublished manuscript.] Sacramento, CA: California Department of Education.

## Test Administration

This chapter details the processes involved in the administration of the 2021–22 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 testing window for the 2021–22 administration of the Alternate ELPAC operational field test was scheduled to take place from November 1, 2021, through February 15, 2022, but was extended to May 31, 2022, the typical close of the summative testing window. This extension allowed LEAs additional time to ensure that all eligible students had an opportunity to participate in the test, as required by state and federal law. Specific test administration schedules within that window were determined locally, pursuant to the *California Code of Regulations,* Title 5 (5 *CCR*), Section 313(d)(2). Refer to [*1.5 Testing Windows and Estimated Testing Times*](#_Testing_Windows_and) for additional information about the Alternate ELPAC testing window.

The Alternate ELPAC operational field test did not allow remote testing, as it uses the one-on-one administration model that all other California alternate assessments use, which requires the test 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 Alternate ELPAC Administration and Scoring Training (AST) training team created a completely virtual training model for the 2021–‍22 administration of the Alternate ELPAC operational field test.

Every local educational agency (LEA) that had an eligible Alternate ELPAC student in California was required to complete the online LEA Certification course on the Moodle Training Site (Moodle). The alternative 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 test. 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 in the administration of the Alternate ELPAC operational field test. Access to the Moodle training site 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 the Moodle training site cannot be accessed. Second, it was restricted because the Moodle keys are stored on a secure website that can only be accessed by logging on to the site using access codes provided only to LEA ELPAC coordinators.

The Moodle training site contained all resources needed to conduct a local test examiner training, such as training videos and PowerPoint training presentations. All aspects of the training courses and materials were developed for the first time using training materials and experiences learned from the pilot.

The following items list high-level assumptions, for the 2021–22 Alternate ELPAC training:

* All LEA ELPAC coordinators from LEAs that had an eligible Alternate ELPAC student in California 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 the operational field test. LEAs were permitted to have their test examiners complete the online training course or they provided their own local training.

#### Goals

The five goals of the 2021–22 Alternate ELPAC operational field test AST were to

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

#### Local Educational Agency Training Requirement

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

##### Certification of Training

There were 1,483 LEA certifications issued, and although only those LEAs with eligible EL students with the most significant cognitive disabilities were required to have the LEA ELPAC coordinator or a designee complete certification, many LEAs without eligible students underwent the training. Reasons for this include that there were some LEA ELPAC coordinators who did not know that, without eligible students, they were not required to participate in the training; and others who were curious to learn about the Alternate ELPAC.

As this was the first time requiring LEAs to complete a certificate of training for the Alternate ELPAC, the Sacramento County Office of Education (SCOE) identified LEAs that reported students taking a locally determined alternate assessment (LDAA) in lieu of the Summative ELPAC during the previous school year to target their AST communications. Any additional LEAs that registered eligible students for the operational field test but were not originally identified on the list were added and those LEAs were also expected to complete the LEA certification. Of the 706 LEAs on the final list, 673 LEAs completed the LEA certification, a completion rate of 95 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 also provided the 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.

##### Monitoring Test Examiner Calibration

Each LEA has a unique user group on the Moodle Training Site, 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’ calibration 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 a password-protected, online platform where developed course materials were made available to ELPAC coordinators, trainers, and test examiners. The Moodle site provided California LEAs with necessary training resources to train test examiners who would administer the Alternate ELPAC.

Training binders were produced and shipped directly to addresses provided by each LEA. All LEAs also had access to the contents of the training binders in the form of PDFs provided on the secure Moodle site.

The Moodle site contained all resources needed to conduct an LEA test examiner training session, including all the materials available in the AST binder and training videos. LEA trainers downloaded materials or ordered extra training binders to prepare for their training sessions and shared access to the site with the test examiners within the LEA. Test examiners used the site to review training materials and complete their individual certifications.

Training materials included the following:

* *Draft Alternate ELPAC Operational Field Test DFA*—Draft *DFAs* were made available for staff to use while watching training videos and to use during local trainings. These *DFAs* were “for training purposes only” and watermarked as “DRAFT.” Once the final *DFAs* were posted in the Test Operations Management System (TOMS), the draft *DFAs* were removed, and staff who still needed to train were directed to TOMS to retrieve the *DFAs*. The appendices of the *DFA* included picture cards.
* Alternate ELPAC Operational Field Test AST Test Item Previews (TIPs)—The TIPs are form- and grade level– or grade span–specific item books that are a rendering of what the test delivery platform will look like. They were provided to LEAs so test examiners could see the items, examine the *DFA* with appendices, and prepare for individualized administration for their students. They were provided in the binder and made available on Moodle prior to the opening of the test delivery system (TDS) for test administration.
* *Alternate ELPAC Training Test DFAs*—LEAs were strongly encouraged to use the training tests and the training test *DFAs* (kindergarten through grade five and grades six through twelve) in local training as an activity to familiarize test examiners with the test delivery platform, test item types, and accessibility resources.

##### Training Module Content

Training courses were divided into five modules. The LEA Certification course included all five modules, while the Test Examiner Certification course included only the first four modules. The module sections for the training materials included content and documents as described in this subsection.

The LEA Certification course was divided into five modules:

* Module 1—Alternate ELPAC Overview
* Module 2—Operational Field Test Overview
* Module 3—Test Administration
* Module 4—Scoring
* Module 5—Training at the Local Level

Modules were as follows:

* **Module 1—Alternate ELPAC Overview**

Module content included the following:

* Timeline
* Eligibility (including the Alternate Assessment Decision Confirmation Worksheet)
* Standards and Connectors
* Understanding of what is measured
* Understanding of how the assessment is administered, arranged by task type
* Three levels of linguistic complexity
* Domain integration to allow individually preferred communication modes
* Task types and item types
* Grade levels and grade spans tested
* Overview of the Initial and Summative Alternate ELPAC
* Score reports

Materials used included the following:

* Printed slides with lines for notetaking
* Resource video: Introduction to the Alternate ELPAC (for parents/guardians) in English and Spanish
* **Module 2—Operational Field Test Overview:**

Module content included the following:

* Operational field test overview and purpose
* The presence of two different test forms
* Introduction to second scoring of rubric-scored items
* English Only student study

Materials used included the following:

* Printed slides with lines for notetaking
* *Draft Alternate ELPAC Operational Field Test DFA* for training purposes only for the test form assigned to the LEA, either Form 1 or Form 2
* **Module 3—Test Administration:**

Module content included the following:

* Review of the *DFA* instructions regarding testing environment
* Alternate text and alternate presentation
* Picture cards
* Technical requirements and resources
* Test settings
* Individualization
* Preparation of augmentative and alternative communication devices and other accessibility resources
* Review of the *DFA* item content
* Review of engagement strategies
* Review of stopping policy
* Survey questions during and following the administration

Materials used included the following:

* Printed slides with lines for notetaking
* Alternate ELPAC Operational Field Test AST TIP
* *Alternate ELPAC Training Test DFAs,* kindergarten through grade five and grades six through twelve
* *Draft Alternate ELPAC Operational Field Test DFA* for training purposes only for the test form assigned to the LEA, either Form 1 or Form 2
* Alternative presentation and Optional Individualization materials
* Appendices of the field test *DFA* for the test form assigned
* Resource video: Overview of Test Administration
* **Module 4—Scoring:**

Module content included the following:

* Review of rubrics and videos of administrations
* Second scoring of rubric-scored items

Materials used included the following:

* Printed slides with lines for notetaking
* Individual videos covering each rubric: modeling rubric, opinion rubric, and comprehensiveness rubric
* **Module 5—Training at the Local Level:**

Module content included the following:

* Anticipated testing time
* Guidelines for who needs to be trained
* Training options and resources
* Test preparation and individualization time

Materials used included the following:

* Printed slides with lines for notetaking

##### Training Videos

Each module had its own training video created and posted to Moodle; these were made available with the other training materials. The topics were the same as described for the training modules. Module four was composed of three videos that covered the administration and scoring of rubric-scored items. There was one video for each of the three rubric-scored item types. Originally, these videos were intended to feature test examiners administering the Alternate ELPAC to students; however, because of the pandemic, this became impossible. In lieu of actual students, actors were used, as well as short vignettes where test examiners described students they had taught previously and how they would individualize administration of the Alternate ELPAC for those students. Both the vignettes and footage using actors were used to demonstrate how test examiners were to administer and score the rubric-scored items.

### User Roles and Standardization

The test administration procedures were designed so that the tests 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 2021–22 school year. LEAs include public school districts, 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 *California Code of Regulations,* Title 5 (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;
* filing a report of a testing incident in STAIRS;
* 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;
* ensuring that all site ELPAC coordinators and test examiners were trained and certified to administer the Alternate ELPAC;
* 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 test site 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 Alternate ELPAC, as appropriate;
* coordinating and verifying the correction of student data errors in the California Longitudinal Pupil Achievement Data System;
* 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 will administer the Alternate ELPAC.

A test examiner must sign 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 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 test 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 *Directions for Administration (DFAs)* for the Alternate ELPAC (CDE, 2022a).

#### Instructions for Test Administration

##### *Directions for Administration*

Test examiners were required to use the *DFAs*, housed securely in TOMS, to administer tests to all eligible students. There was one *DFA* for each form for each grade-level and grade-span test. Test examiners could only access the *DFAs* for the form(s) their LEA was assigned to. LEA ELPAC coordinators could access all *DFAs*, regardless of form assignment, so that if a student moved to an LEA assigned to a different form from the previous LEA during testing, the student could continue testing and the test examiner could use the appropriate *DFA*.

##### *Alternate ELPAC Operational Field Test Manual*

The *Alternate ELPAC Operational Field Test Manual* (CDE, 2022a) 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 test and systems associated with the test.

##### *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, 2022e):

* **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.

* ***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, 2022d).
* ***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, 2022c).
* ***CAASPP and ELPAC Accessibility Guide*—**This manual provided descriptions of the accessibility features for computer-based tests as well as information about supported hardware and software requirements for administering tests to students using accessibility resources, including those with a braille accommodation using Job Access With Speech (JAWS®) (software) or a braille embosser (hardware) (CDE, 2022b).

### Local Educational Agency Training

Each year, ETS, in collaboration with the CDE and its Assessment Validity and Outreach contractor, the Sacramento County Office of Education (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. Training was also livestreamed on YouTube so that educators still had access if a particular training reached registration capacity.

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 job or home 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 a “workday” for coordinators where support staff were available throughout the day, allowing coordinators to join as needed and get customized support. SCOE continued to offer assessment update meetings intended to provide LEA coordinators with regular updates about California’s assessment system. 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 tests 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 and guardians to help them understand the assessment’s purpose or how to read the Student Score Report (SSR). 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 Quick Reference Guides and Videos 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 2021–22 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 510,000 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 2021–22.

### 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 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 2021–22 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 content area.

The CDE’s *California Assessment Accessibility Resources Matrix* (Accessibility Matrix) (CDE, 2021b) 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 2021–22 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
* Highlighter
* Line reader (grades three through twelve)
* Oral clarification of test directions by the test examiner in English
* Scratch paper

##### Designated Supports

Designated supports were available to all students 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 2021–22 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
* 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
* Color contrast
* Color overlay
* Magnification
* Masking
* Medical supports
* Noise buffers
* Print-on-demand
* 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. 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 2021–22 Alternate ELPAC administration.

###### Embedded

Because the 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
* Speech-to-text
* Word processor

##### 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, 2021b). During the 2021–22 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 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 with an asterisk and a footnote that the test 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 Alternate ELPAC operational field test.

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, including students with disabilities and English learner (EL) students. The CDE Accessibility Matrix (CDE, 2021b) 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 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 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 Alternate ELPAC was provided in the *Alternate ELPAC Accessibility and Accommodations Guidelines* (CDE, 2021a)*.* It was developed by using the California Alternate Assessment accessibility guidelines in conjunction with the considerations for 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 *Guidelines* and the *2021–22* CAASPP and ELPAC Accessibility Guide (CDE, 2022b), 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 [*1.7 Systems Overview and Functionality*](#_Systems_Overview_and_2) in *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).)
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 Accommodations—Additional Instructional Supports and Resources for Alternate Assessments | 12 | 1% | 13 | 1% | 15 | 1% |
| Non-Embedded Accommodations—Alternate Response Options | 59 | 5% | 47 | 4% | 40 | 3% |
| Non-Embedded Accommodations—ASL or Manually Coded English | 2 | 0% | 3 | 0% | 4 | 0% |
| Non-Embedded Accommodations—Breaks | 97 | 7% | 97 | 8% | 89 | 7% |

Table 5.1 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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 Accommodations—Scribe | 32 | 2% | 31 | 2% | 45 | 4% |
| Non-Embedded Accommodation—Unlisted Resources | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Supports—Color Contrast | 7 | 1% | 5 | 0% | 2 | 0% |
| Embedded Designated Supports—Masking | 38 | 3% | 24 | 2% | 25 | 2% |
| Embedded Designated Supports—Mouse Pointer (Size and Color) | 34 | 3% | 38 | 3% | 22 | 2% |
| Embedded Designated Supports—Permissive Mode | 7 | 1% | 4 | 0% | 2 | 0% |
| Embedded Designated Supports—Print Size | 13 | 1% | 8 | 1% | 8 | 1% |
| Embedded Designated Supports—Streamline | 20 | 2% | 10 | 1% | 18 | 1% |
| Embedded Designated Supports—Turn Off Any Universal Tool(s) | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Supports—Color Contrast | 4 | 0% | 5 | 0% | 2 | 0% |
| Non-Embedded Designated Supports—Color Overlay | 2 | 0% | 1 | 0% | 0 | N/A |
| Non-Embedded Designated Supports—Designated Interface Assistant | 14 | 1% | 8 | 1% | 12 | 1% |
| Non-Embedded Designated Supports—Magnification | 5 | 0% | 12 | 1% | 8 | 1% |
| Non-Embedded Designated Supports—Masking | 28 | 2% | 26 | 2% | 23 | 2% |
| Non-Embedded Designated Supports—Medical Supports | 1 | 0% | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Supports—Noise Buffers | 33 | 3% | 29 | 2% | 39 | 3% |
| Non-Embedded Designated Supports—Print on demand | 1 | 0% | 2 | 0% | 2 | 0% |
| Non-Embedded Designated Supports—Read-Aloud Items | 40 | 3% | 61 | 5% | 60 | 5% |
| Non-Embedded Designated Supports—Separate Setting | 113 | 9% | 101 | 8% | 89 | 7% |
| Non-Embedded Designated Supports—Simplified Test Directions | 128 | 10% | 122 | 10% | 106 | 9% |
| Non-Embedded Designated Supports—Translated Test Directions (including ASL) | 14 | 1% | 19 | 2% | 16 | 1% |
| **Total Students Tested:** | **1,297** | **N/A** | **1,266** | **N/A** | **1,219** | **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 | Grades 6-8: % of Total Tested | Grades 9–12: N | Grades 9–12: % of Total Tested |
| Non-Embedded Accommodations—Additional Instructional Supports and Resources for Alternate Assessments | 63 | 1% | 72 | 2% | 127 | 3% |
| Non-Embedded Accommodations—Alternate Response Options | 127 | 3% | 110 | 3% | 202 | 4% |
| Non-Embedded Accommodations—ASL or Manually Coded English | 1 | 0% | 19 | 1% | 22 | 0% |
| Non-Embedded Accommodations—Breaks | 291 | 7% | 256 | 7% | 275 | 5% |
| Non-Embedded Accommodations—Scribe | 183 | 4% | 141 | 4% | 153 | 3% |
| Non-Embedded Accommodation—Unlisted Resources | 0 | N/A | 0 | N/A | 0 | N/A |
| Embedded Designated Supports—Color Contrast | 14 | 0% | 18 | 0% | 11 | 0% |
| Embedded Designated Supports—Masking | 83 | 2% | 80 | 2% | 43 | 1% |
| Embedded Designated Supports—Mouse Pointer (Size and Color) | 90 | 2% | 71 | 2% | 94 | 2% |
| Embedded Designated Supports—Permissive Mode | 16 | 0% | 32 | 1% | 8 | 0% |
| Embedded Designated Supports—Print Size | 35 | 1% | 67 | 2% | 62 | 1% |
| Embedded Designated Supports—Streamline | 62 | 1% | 68 | 2% | 38 | 1% |
| Embedded Designated Supports—Turn Off Any Universal Tool(s) | 0 | N/A | 0 | N/A | 0 | N/A |
| Non-Embedded Designated Supports—Color Contrast | 15 | 0% | 22 | 1% | 13 | 0% |
| Non-Embedded Designated Supports—Color Overlay | 3 | 0% | 15 | 0% | 5 | 0% |
| Non-Embedded Designated Supports—Designated Interface Assistant | 55 | 1% | 56 | 2% | 58 | 1% |

Table 5.2 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accessibility Resource | Grade Span 3–5: N | Grade Span 3–5: % of Total Tested | Grade Span 6–8: N | Grades 6-8: % of Total Tested | Grades 9–12: N | Grades 9–12: % of Total Tested |
| Non-Embedded Designated Supports—Magnification | 21 | 0% | 42 | 1% | 78 | 2% |
| Non-Embedded Designated Supports—Masking | 64 | 2% | 45 | 1% | 30 | 1% |
| Non-Embedded Designated Supports—Medical Supports | 3 | 0% | 2 | 0% | 0 | N/A |
| Non-Embedded Designated Supports—Noise Buffers | 169 | 4% | 138 | 4% | 149 | 3% |
| Non-Embedded Designated Supports—Print on demand | 3 | 0% | 2 | 0% | 13 | 0% |
| Non-Embedded Designated Supports—Read-Aloud Items | 319 | 8% | 293 | 8% | 308 | 6% |
| Non-Embedded Designated Supports—Separate Setting | 458 | 11% | 392 | 11% | 333 | 7% |
| Non-Embedded Designated Supports—Simplified Test Directions | 498 | 12% | 415 | 11% | 415 | 8% |
| Non-Embedded Designated Supports—Translated Test Directions (including ASL) | 52 | 1% | 50 | 1% | 42 | 1% |
| **Total Students Tested:** | **4,231** | **N/A** | **3,624** | **N/A** | **5,032** | **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 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 and PPT assessments. These resources are not part of the technology platform for the computer-administered 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*](#_Toc121842306) for a detailed description of the accessibility resources available to students taking the 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 STAIRS incident could be submitted to reset the test 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 test so the student could be retested without the accommodation or designated support.

After schools and LEAs assigned eligible students to accommodations or designated supports, Cambium Assessment, Inc.’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 test. However, because there are no embedded accommodations for the Alternate ELPAC, there was no usage of accommodations to capture or report.

Table 5.3 reports the number of students who, based on the availability of data, were assigned and used a designated support at least once during test administration.

Types of designated supports included in table 5.3 are as follows:

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

Table 5.3 presents the assignment and usage of embedded designated supports for each grade level or grade span.

Table 5.3 Assignment and Usage of Embedded Designated Supports

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Accessibility Resource | Students Assigned | Students Used | Percentage Used |
| All | Any Tracked Resource | 315 | 19 | 6.03 |
| All | Embedded Masking | 293 | 10 | 3.41 |
| All | Non-Embedded Print-on-Demand | 23 | 9 | 39.13 |
| Kindergarten | Any Tracked Resource | 39 | 4 | 10.26 |
| Kindergarten | Embedded Masking | 38 | 4 | 10.53 |
| Kindergarten | Non-Embedded Print-on-Demand | 1 | 0 | 0.00 |
| 1 | Any Tracked Resource | 26 | 3 | 11.54 |
| 1 | Embedded Masking | 24 | 2 | 8.33 |
| 1 | Non-Embedded Print-on-Demand | 2 | 1 | 50.00 |
| 2 | Any Tracked Resource | 27 | 1 | 3.70 |
| 2 | Embedded Masking | 25 | 1 | 4.00 |
| 2 | Non-Embedded Print-on-Demand | 2 | 0 | 0.00 |
| 3–5 | Any Tracked Resource | 86 | 4 | 4.65 |
| 3–5 | Embedded Masking | 83 | 3 | 3.61 |
| 3–5 | Non-Embedded Print-on-Demand | 3 | 1 | 33.33 |
| 6–8 | Any Tracked Resource | 82 | 2 | 2.44 |
| 6–8 | Embedded Masking | 80 | 0 | 0.00 |
| 6–8 | Non-Embedded Print-on-Demand | 2 | 2 | 100.00 |

Table 5.3 *(continuation)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Accessibility Resource | Students Assigned | Students Used | Percentage Used |
| 9–12 | Any Tracked Resource | 55 | 5 | 9.09 |
| 9–12 | Embedded Masking | 43 | 0 | 0.00 |
| 9–12 | Non-Embedded Print-on-Demand | 13 | 5 | 38.46 |

### Practice and Training Tests

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

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 Alternate ELPAC. These tests more closely simulate the 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 Alternate ELPAC operational field test. 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 operational 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 tests, 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 tests that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI), which is described in the next subsection.

All tests within the ELPAC System, as well as the confidentiality of student information, should be protected to ensure the validity, reliability, and fairness of the results. As stated in *Standard 7.9* (AERA, APA, & NCME, 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 *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

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

The OTI helps prevent misconduct on the part of students and administrators, detects potential misconduct through empirically established indicators, and resolves situations involving misconduct in a fair and balanced way that reflects the laws and professional standards governing the integrity of testing. In its pursuit of enforcing secure testing practices, the OTI strives to safeguard the various processes involved in a test development and administration cycle. For the Alternate ELPAC, those processes included the following:

* Test 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, 2022a).

The following measures ensured the security of the ELPAC:

* LEA ELPAC coordinators and site ELPAC coordinators must have electronically signed and submitted a “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* 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 the ELPAC and the CAASPP 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 sub section [*1.7 Systems Overview and Functionality*](#_Systems_Overview_and_2) 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 currently uses secure file transfer protocol (SFTP) and encryption for all data file transfers; test data is never sent via email. SFTP is a method for 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.

#### Data Management in the Secure Database

ETS currently maintains a secure database to house all student demographic data and assessment results. Information associated with each student has a database relationship to the LEA, school, and grade codes as the data is collected during 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, 2022e).

#### 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 Alternate ELPAC:

* Individual student results for computer-based assessments in the California Educator Reporting System (CERS)
* Individual SSRs (electronic)
* Internet reports—available on the CDE Test Results for California’s Assessments website—aggregated by content area and 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. ETS 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 the BS 25999-2 standard for business continuity and conducts disaster recovery exercises annually. ETS routinely backs up all data to either disks through deduplication or to tapes, all of which are stored off site.

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

* 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 a test 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 tests, which, in turn, compromise the reliability and validity of test results (CDE, 2022c). Whether intentional or unintentional, failure by staff or students to comply with security rules constitutes a test security incident. Test security incidents have impacts on scoring and affect students’ performance on the test.

LEA ELPAC coordinators and site ELPAC coordinators ensured that all test security and summative 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 test. If an Appeal to a student’s test 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, 2022c).

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, 2022c).

Types of Appeals available during the 2021–22 ELPAC administration are described in table 5.4.

Table 5.4 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 tests that were reset previously. |

##### 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 test, test security, or test validity. 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 test or impact test security or test validity. 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 test. 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 telephone. 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.

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.

#### 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, 2022e). Types of Appeals available during the 2021–22 ELPAC administration are described in table 5.4.

The total number of incidents reported in STAIRS for the Alternate ELPAC operational field test was 481. The number of STAIRS incidents that required an Appeal was 210. The most common Appeal type was Reset, and the second most common was Re-open. These counts include incidents that were in draft form, pending, or partially approved. Note that in table 5.5, because some SSIDs were submitted with multiple Appeal types, and some Appeal types were submitted with multiple SSIDs, the numbers will not add up to the numbers within this paragraph.

Table 5.5 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 Statewide Student Identifiers (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 5.5 Number and Types of Incidents Submitted in STAIRS

|  |  |  |  |
| --- | --- | --- | --- |
| Description | Appeal Type | Number of Incidents | Total Number of SSIDs Submitted |
| Accessibility Issue | Reset | 20 | 22 |
| Accidentally Submitted Test | Re-open | 25 | 30 |
| Administered Incorrect Assessment | Reset or No Appeal | 150 | 265 |
| Administration Error | No Appeal | 49 | 68 |
| Data Entry Issue | Reset, No Appeal, or Re-open | 77 | 106 |
| Incorrect SSID Used | Reset or No Appeal | 6 | 7 |
| Irregularity Flag Submitted in Error | No Appeal | 0 | 0 |
| Other Issues | No Appeal | 0 | 0 |
| Restore from Reset | Restore | 1 | 1 |
| Student Cheating or Accessing Unauthorized Devices | No Appeal | 0 | 0 |
| Student Disruption | No Appeal | 1 | 1 |

Table 5.6 provides the counts of approved Appeals.

Table 5.6 Number of Appeals Approved in STAIRS in the 2021–22 Administration—All Grade Levels and Grade Spans

|  |  |  |
| --- | --- | --- |
| Appeal Type | Number of Incidents | Total Number of Appeals |
| Reset | 140 | 223 |
| Re-open | 28 | 31 |
| Restore | 1 | 1 |
| No Appeal | 163 | 247 |

### References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

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

California Department of Education. (2021b). *California assessment accessibility resources matrix*. Sacramento, CA: California Department of Education.

California Department of Education. (2022a). *Alternate ELPAC operational field test manual.* Sacramento, CA: California Department of Education.

California Department of Education. (2022b). *CAASPP and ELPAC accessibility guide*. Sacramento, CA: California Department of Education.

California Department of Education. (2022c). *CAASPP and ELPAC security incidents and Appeals procedure guide.* Sacramento, CA: California Department of Education.

California Department of Education. (2022d). *CAASPP and ELPAC technical specifications and configuration guide*. Sacramento, CA: California Department of Education.

California Department of Education. (2022e). *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.

## Standard Setting

This chapter summarizes the standard setting process through which Alternate English Language Proficiency Assessments for California (ELPAC) performance levels and threshold scores were recommended. Included are an overview of the standard setting methodology, a summary of the standard setting procedure, the description of the performance level descriptors (PLDs), and the results. Detailed standard setting information for the Alternate ELPAC is described in the *Standard Setting Technical Report for the Alternate ELPAC* (California Department of Education [CDE], 2022).

### Background

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

The Alternate ELPAC aligns with English language development (ELD) standards set out in the 2012 *California English Language Development Standards: Kindergarten Through Grade 12* (2012 ELD Standards) (CDE, 2014) via the ELD Connectors (CDE, 2019a). The ELD Connectors offer a reduction in the depth, breadth, and complexity of the standards, as appropriate for students with the most significant cognitive disabilities. In addition to being directed by the 2012 ELD Standards, the ELD Connectors were substantially informed by the English language proficiency (ELP) level descriptors in the Council of Chief State School Officers’ *English Language Proficiency (ELP) Standards for English Learners with Significant Cognitive Disabilities* (2019).

Prior to the standard setting, Alternate ELPAC general PLDs were presented and approved at the May 2019 SBE meeting, as listed in the *Proposed High-Level Test Design for the Alternate ELPAC* (CDE, 2019b). These general, or policy, PLDs convey the degree of a student’s ELP using three levels: Novice English Learner (EL) (Level 1), Intermediate EL (Level 2), and Fluent English Proficient (Level 3).

ETS conducted standard setting workshops in 2022, to collect judgments on the 2021–22 operational field test items of the Alternate ELPAC. The Modified Angoff and Extended Angoff standard setting methods were applied to each grade-level or grade-span test. These methods allowed for the collection of panelist judgments for each item administered in 2021–22, thereby providing flexibility in the development of threshold scores for reporting the overall score for the Initial Alternate and Summative Alternate ELPAC. Refer to section [*6.3 Standard Setting Methodology*](#_Standard_Setting_Methodology) for more information about the Angoff methods.

### Performance Level Descriptors

The Alternate ELPAC reports three performance levels—Level 1 through Level 3—defined by the general PLDs. In fall 2020, range PLDs were reviewed by California educators and approved by the CDE. Range PLDs further describe what students at each performance level know and can do. As one of the most critical parts of the standard setting process, participants referred to the range PLDs to define the threshold PLDs, which is the set of knowledge, skills, and abilities expected of threshold students who are at the entry point of each performance level for each grade level or grade span assessed.

Table 6.1 provides a description of the three general PLDs, with Level 3 being the highest level of performance and Level 1 being the lowest (CDE, 2019b).

Table 6.1 Alternate ELPAC General PLDs

|  |  |
| --- | --- |
| Level | Description |
| Level 3, Fluent English Proficient | Students at this level have **sufficient** English language proficiency. They may need **occasional** linguistic support to enable them to access adapted grade-level content in English. |
| Level 2, Intermediate EL | Students at this level have **moderate** English language proficiency. They may need **frequent** linguistic support to enable them to access adapted grade-level content in English. |
| Level 1, Novice EL | Students at this level have **minimal** English language proficiency. They need **substantial** linguistic support to enable them to access adapted grade-level content in English. |

### Standard Setting Methodology

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

#### Modified and Extended Angoff Methods

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

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

### Standard Setting Procedures

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

#### Panelists

Prior to standard setting, a diverse group, representative of California educators familiar with instructing this student population, were recruited to participate as panelists in the standard setting sessions. Panelists were assigned to one of six panels of educators; each panel focused on one grade level or grade span. The panels included a range of California educators who were familiar with the ELD Connectors, were engaged in the daily instruction of EL students with the most significant cognitive disabilities, and were familiar with the diverse group of students eligible to take the Alternate ELPAC. It was important to include teachers working with these students, as these teachers provided a perspective on the knowledge, skills, and learning goals for the students taking the Alternate ELPAC.

Participating educators included representatives from across regions in California (north, south, and central) and across gender, race, and ethnic categories. Fifty-three panelists participated; all educators indicated experience working with the 2012 ELD Standards, and 47 indicated experience working with the ELD Connectors.

#### Materials

Panelists were provided with a letter describing the purpose and procedure of the standard setting workshop along with a preworkshop assignment, instructions, a notetaking form, and the links to both the general PLDs and the Alternate ELPAC training materials. During the workshop, panelists received Alternate ELPAC operational field test materials, judgment data collection materials, and standard setting evaluation forms. Materials used during the workshop were provided using remote tools, such as the ETS Content Review Tool and Zoom web conference software. The detailed procedure regarding securing the operational assessment materials is described in the *Standard Setting Technical Report for the Alternate ELPAC* (CDE, 2022).

#### Process

Prior to making judgments, panelists reviewed and discussed the test and the SBE-approved PLDs, and the range PLDs for each level, and then developed threshold student definitions as a group. Two threshold student definitions were developed, one for Level 2 and one for Level 3. For example, a threshold Level 2 student is a student at the beginning of Level 2; this student definition represents the knowledge and skills of the lowest-performing Level 2 student and differentiates this student from the knowledge and skills of the highest-performing Level 1 student.

Figure 6.1 shows where threshold students are located. This graph shows figures representing students in the three levels. The arrow pointing to the leftmost figure in the Level 2 group indicates that this is the threshold Level 2 student. The arrow pointing to the leftmost figure in the Level 3 group indicates that this is the threshold Level 3 student.

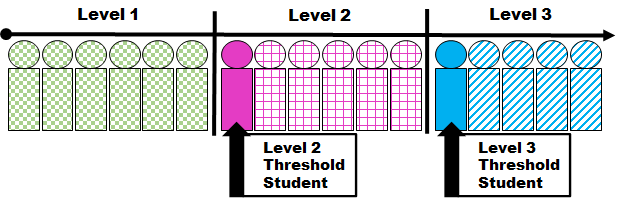


Figure 6.1 Location of threshold students

For each of the performance levels, panelists made two rounds of judgments on operational items. Round 1 judgments were made independently, without discussion. For items scored as 1-point items, the Modified Angoff method was used; for 2-point items, the Extended Angoff method was used (Cizek & Bunch, 2007; Hambleton & Plake, 1995). The 1-point items included discrete item types such as multiple-choice selected response. The 2-point rubric-scored items included expressive responses with scoring rules that indicated how a student will obtain a score of 0, 1, or 2. Panelists were asked to first review the definition of the threshold student, then review the item and its key or scoring rubric, and then make their judgment.

Panelists’ judgments were collected, analyzed, summarized, and shared with the panel. Panelists received their own item-level and total-score judgments as well as panel summary data. As part of the post–Round 1 feedback, panelists reviewed the individual judgments in the context of the range of judgments across the panel, and the facilitator shared feedback on the similarities and differences of the panel judgments on the Alternate ELPAC items. Panelists discussed with the other panelists the rationales for the independent judgments.

The feedback and discussion from the Round 1 judgment data then informed Round 2 judgments. Panelists were invited to make changes to their item-level judgments and enter changes into the Round 2 judgment form. After this round, panelists’ judgments were collected, analyzed, summarized, and shared with the panel as the final panel recommendation. The final recommended threshold scores were based on the mean of panelists’ judgments.

After completing a final evaluation survey, participants from each of the six grade-level and grade-span tests were asked to participate in a vertical articulation meeting to evaluate the process and recommendations across all six pairs of threshold score judgments. Panelists discussed the similarities and differences in the Level 3 threshold student definitions. Specifically, representatives in adjacent grade levels or grade spans described the progression of expectations as the grade levels or grade spans increased (e.g., panelists from grades one and two; panelists from grade two and grade span three through five). There was general agreement that the threshold student definitions for entering Level 3 appropriately described the expectations as the grade levels or grade spans increased.

As part of the standard setting process, ETS created a standard setting score scale specifically to report threshold scores based on standard setting; this is different from the reporting scale for the Alternate ELPAC. The impact data results presented in table 6.2 through table 6.7 were calculated after the calibration and scaling analyses were conducted and are presented in the standard setting scale score metric. Although the impact data was not available during the workshop, the data was considered by the CDE and the Technical Advisory Group (TAG) prior to the presentation of the State Superintendent of Public Instruction (SSPI) recommendations. The CDE and the TAG analyzed the standard setting panel’s judgments and refined the threshold scores for consistency across the Alternate ELPAC grade levels and grade spans tested.

Table 6.2 through table 6.7 present the panel-recommended threshold scores on the standard setting scale and the projected distribution of 2021–22 students based on Round 2 recommendations for each grade level or grade span. Note that percentages might not equal 100 because of rounding.

Table 6.2 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Kindergarten

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

Table 6.3 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade One

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

Table 6.4 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Two

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

Table 6.5 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Three Through Five

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

Table 6.6 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Six Through Eight

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

Table 6.7 Projected Distribution of 2021–22 Students Based on Round 2 Recommendations: Grade Span Nine Through Twelve

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

### Results of the Standard Setting

The SBE approved the recommendation of the final threshold scores for the Alternate ELPAC. The recommendations of the SSPI are presented in table 6.8. The table shows the percentage of students statewide that would be placed at this performance level based on the results of the 2021–22 Alternate ELPAC administration. Also shown in this table is the percentage of students statewide that would be at or above this level on the basis of the results of the 2021–22 administration. The SSPI recommendations included separate threshold scores for grade spans nine and ten and eleven and twelve, to mimic the reporting of performance levels in the Summative ELPAC. Finally, the standard setting threshold score is the minimum standard setting scale score needed to achieve a performance level on the 2021–‍22 administration of tests. Note that threshold score scales were generated solely for the standard setting process; reporting scales were developed to report scores on the Student Score Report and public reporting.

Table 6.8 SSPI’s Recommendations for the Proposed Performance Levels for the Alternate ELPAC

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade | Percent of Students in Level 1 | Percent at or Above in Level 1 | Percent of Students in Level 2 | Standard Setting Scale Threshold Score for Level 2 | Percent at or Above in Level 2 | Percent of Students in Level 3 | Standard Setting Scale Threshold Score for Level 3 |
| K | 39.6 | 100 | 45.0 | 254 | 60.4 | 15.4 | 275 |
| 1 | 33.3 | 100 | 49.0 | 250 | 66.7 | 17.8 | 271 |
| 2 | 40.5 | 100 | 41.8 | 252 | 59.5 | 17.7 | 272 |
| 3–5 | 27.6 | 100 | 44.8 | 244 | 72.4 | 27.7 | 268 |
| 6–8 | 22.7 | 100 | 36.7 | 237 | 77.3 | 40.6 | 261 |
| 9–10 | 27.7 | 100 | 39.5 | 239 | 72.3 | 32.7 | 267 |
| 11–12 | 28.1 | 100 | 36.9 | 239 | 71.9 | 35.1 | 267 |

The reporting scale score range for each performance level at different grade levels is presented in table 7.1 in [chapter 7](#_Scoring_and_Reporting). The performance threshold score for each level is the lower bound of each scale score range. The scale score ranges do not change from year to year. Once established, they remain unchanged from administration to administration until such time that new performance standards are adopted. Table 7.3 presents the percentage of students meeting each proficiency level in the 2021–22 administration of the Alternate ELPAC.

### References

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

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

California Department of Education. (2014). *2012 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*. Sacramento, California: California Department of Education.

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

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

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

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

## Scoring and Reporting

To determine individual students’ scores for the 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 Alternate ELPAC. Student score distribution summaries and summaries by student group are also presented here.

### Student Test Scores

Overall scores were reported for the Alternate ELPAC at the individual student level. To obtain overall scores, the ability (theta) scores needed to be estimated.

Prior to the test administration, ETS Assessment and Learning Technology Research & 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. ETS’ Enterprise Score Key Management (eSKM) system collected and calculated individual students’ overall scores (e.g., total raw scores). ETS’ Psychometric Analysis & Research (PAR) team conducted a series of psychometric analyses such as calibration, equating, and scaling using individual item scores of the test samples and produced the raw-to-scale-score conversion tables based on all psychometric analyses. When the conversion tables were implemented, eSKM produced the scale score and performance levels using the reporting score ranges for students who completed the test.

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. ETS’ PAR team also computed individual student scores based on the same data files using a statistical analysis software. The scores from the two systems 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 items answered correctly and the number of a rubric-based item score points obtained. These 24 items from each form were selected from 34 operational field-tested items for each form.

The number and percentage of students at each raw score and the total test score are reported in [appendix 7.A](#_Appendix_6.A:_Raw). Table 7.A.1 through table 7.A.11 present the raw score frequency distributions for the scores for each grade level or grade span. For grades three and above, the raw score frequency tables are presented for each form, since the 24 operational items are not identical across the two forms.

#### Theta Scores

All the items presented to the student were calibrated onto the theta scale; the student’s raw score could be transformed into an ability (theta) estimate by using the item response theory (IRT) inverse test characteristic curve (TCC) method (Stocking, 1996). With this method, the student’s estimated ability is the ability value at which the expected raw score is equal to the student’s raw score. Refer to subsection [*8.4.3 Equating*](#_Equating_1) for equating procedures and the IRT inverse TCC method. The Alternate ELPAC test for each grade level and grade span has its own theta scale.

When a conversion table from the raw score to theta score was created for each form, the estimated ability (theta) score of each individual student could be obtained from the conversion table. The theta score could later be transformed into a scale score through a linear transformation. Refer to subsection [*8.5.2 Transformation from Theta Scores to Scale Scores*](#_Transformation_from_Theta_1) for more information. Refer to [appendix 8.G](#_Appendix_7.G_=) for the raw-to-scale-score conversion tables.

Raw scores are not directly comparable from administration to administration and from form to form because each raw score is based on a set of items that may differ in difficulty. Instead, student performance on the Summative Alternate ELPAC is reported in terms of scale scores that express student proficiency in terms of a constant metric. Thus, a scale score of 350 for grade one in one administration represents the same level of proficiency as 350 for grade one in another administration even though each scale score may represent a different raw score. Similarly, for grade levels or grade spans with more than one form, the same scale score represents the same level of proficiency even though it may correspond to different raw scores on the two forms.

#### Scale Scores for the Total Assessment

The following requirements were used to develop and define the Summative Alternate ELPAC reporting scale ranges, as described in subsection [*8.5 Scaling the Scores*](#_Toc120784038):

1. Each scale score has three digits (e.g., 320, 551, or 780), where the first digit is indicative of the grade level or grade span being reported. The latter two 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. Score ranges are grade level– and grade span–specific. For example, the possible scale scores would be 201 to 299 for kindergarten with the lowest obtainable scale score (LOSS) at 201 and the highest obtainable scale score (HOSS) at 299. For grade span three through five, this range is 501 to 599 with a LOSS of 501 and a HOSS of 599, and so on for the other grade levels or grade spans. For grade span eleven and twelve, the scale ranges from 801 to 899 with a LOSS of 801 and a HOSS of 899.
3. Each threshold score on the scale is the same from year to year. Also, across the grade levels, the last two digits corresponding to the Level 2—Intermediate English Learner (EL) and Level 3—Fluent English Proficient threshold scores are the same (refer to subsection [*6.2 Performance Level Descriptors*](#_Performance_Level_Descriptors)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.12 through table 7.A.18.

#### Performance Levels

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

* **Level 1—Novice English Learner:** The student is beginning to develop the English skills needed to communicate and learn in school.
* **Level 2—Intermediate English Learner:** The student can sometimes use English to communicate and learn in school.
* **Level 3—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], 2022a).

Scale score ranges for each performance level are presented in table 7.1.

Table 7.1 Reporting Scale Score Ranges for Each Performance Level by Grade Level

|  |  |  |  |
| --- | --- | --- | --- |
| Grade Level | Level 1 | Level 2 | Level 3 |
| Kindergarten | 201ꟷ243 | 244ꟷ259 | 260ꟷ299 |
| 1 | 301ꟷ343 | 344ꟷ359 | 360ꟷ399 |
| 2 | 401ꟷ443 | 444ꟷ459 | 460ꟷ499 |
| 3 | 501ꟷ543 | 544ꟷ559 | 560ꟷ599 |
| 4 | 501ꟷ543 | 544ꟷ559 | 560ꟷ599 |
| 5 | 501ꟷ543 | 544ꟷ559 | 560ꟷ599 |
| 6 | 601ꟷ643 | 644ꟷ659 | 660ꟷ699 |
| 7 | 601ꟷ643 | 644ꟷ659 | 660ꟷ699 |
| 8 | 601ꟷ643 | 644ꟷ659 | 660ꟷ699 |
| 9 | 701ꟷ743 | 744ꟷ759 | 760ꟷ799 |
| 10 | 701ꟷ743 | 744ꟷ759 | 760ꟷ799 |
| 11 | 801ꟷ843 | 844ꟷ859 | 860ꟷ899 |
| 12 | 801ꟷ843 | 844ꟷ859 | 860ꟷ899 |

### Overview of Score Aggregation Procedures

To provide meaningful results to the interest holders, test scores for a given grade level and content area 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) for more information.

#### Student Score Distributions and Summary Statistics

Table 7.2 presents the means and standard deviations (SDs) of overall scale scores by grade level.

Table 7.2 Mean and SD of Scale Scores by Grade Level

|  |  |  |  |
| --- | --- | --- | --- |
| Grade Level | Number of Students Tested | Overall Scale Score Mean | Overall Scale Score SD |
| Kindergarten | 1,297 | 242 | 19 |
| 1 | 1,266 | 345 | 19 |
| 2 | 1,219 | 444 | 19 |
| 3 | 1,475 | 549 | 17 |
| 4 | 1,388 | 550 | 18 |
| 5 | 1,368 | 553 | 18 |
| 6 | 1,252 | 655 | 22 |
| 7 | 1,258 | 656 | 23 |
| 8 | 1,114 | 658 | 23 |
| 9 | 902 | 752 | 21 |
| 10 | 849 | 753 | 22 |
| 11 | 896 | 855 | 22 |
| 12 | 2,385 | 853 | 23 |

The means and SDs of scale scores for the overall test are also presented by student group. These results are in [appendix 7.B](#_Appendix_7.B:_Means_1), in table 7.B.1 through table 7.B.7.

The percentage of students in each performance level for the overall test is presented in table 7.3. Note that numbers might not add to 100 because of rounding.

Table 7.3 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 | 556 | 43 | 544 | 42 | 197 | 15 |
| 1 | 489 | 39 | 568 | 45 | 209 | 17 |
| 2 | 537 | 44 | 474 | 39 | 208 | 17 |
| 3 | 465 | 32 | 685 | 46 | 325 | 22 |
| 4 | 392 | 28 | 640 | 46 | 356 | 26 |
| 5 | 331 | 24 | 592 | 43 | 445 | 33 |
| 6 | 322 | 26 | 470 | 38 | 460 | 37 |
| 7 | 305 | 24 | 462 | 37 | 491 | 39 |
| 8 | 235 | 21 | 385 | 35 | 494 | 44 |

Table 7.3 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Level 1 N | Level 1 % | Level 2 N | Level 2 % | Level 3 N | Level 3 % |
| 9 | 237 | 26 | 406 | 45 | 259 | 29 |
| 10 | 248 | 29 | 339 | 40 | 262 | 31 |
| 11 | 228 | 25 | 352 | 39 | 316 | 35 |
| 12 | 695 | 29 | 928 | 39 | 762 | 32 |

Figure 7.1, which is derived from the data in table 7.3, presents the percentage of students at each performance level by grade level. The percentage of students in Level 3 increases as grade level increases until grade nine, at which point the percentage drops from 44 percent in grade eight to 29 percent in grade nine. In all high school grades, the percentage of students in Level 3 remains between 29 percent and 35 percent. Beginning with grade five, the percentage of students in Level 1 is always the smallest of the three performance levels.

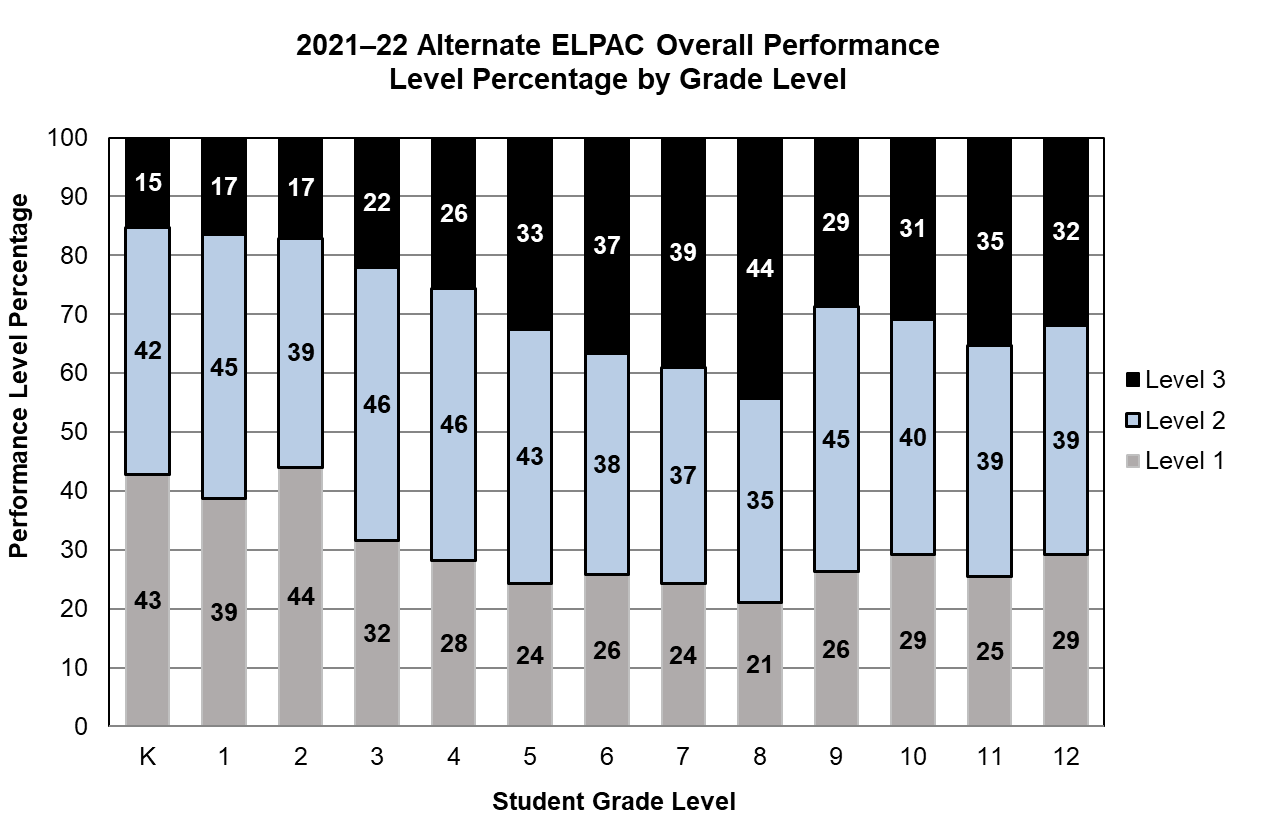


Figure 7.1 Percentage of students at each performance level

#### Demographic Student Group Summaries

Table 7.C.1 through table 7.C.6 in [appendix 7.C](#_Appendix_6.C:_Demographic) 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 2021–22 operational field test administration of the Alternate ELPAC.

The tables in [appendix 7.C](#_Appendix_6.C:_Demographic) group students by demographic characteristics, including gender, ethnicity, economic status, primary disabilities, length of enrollment in US schools, and migrant status. The list of student groups is presented in table 7.4.

Table 7.4 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** | * 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 |

#### Student Group Distributions

Table 7.B.1 through table 7.B.7 in [appendix 7.B](#_Appendix_6.B:_Means) show consistent patterns among test takers. For all grade levels and grade spans, male students accounted for 64 to 73 percent of the test takers, and 73 to 78 percent of the students taking the Alternate ELPAC were Hispanic or Latino. Students with autism accounted for 27 to 59 percent of the students tested, while students with intellectual disabilities accounted for between 23 to 51 percent of the students tested. Autism was the most common primary disability for students testing in kindergarten through grade span three through five, and intellectual disability was the most common primary disability in the higher grade spans.

### Reports Produced and Scores for Each Report

The tests 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

In addition to TOMS, another California online reporting system was used during the 2021–‍22 administration: the California Educator Reporting System (CERS).

TOMS communicated with CERS, which provided authorized users with interactive and cumulative online reports for the Alternate ELPAC at the student, school, group, and LEA levels. CERS provided preliminary score data for each administered test available in the reporting system.

Based on the ELPAC reporting requirements, CERS provided the preliminary summative reports containing information outlining student knowledge and skills. CERS also permitted access to individual score reports, which provided preliminary score data for each administered test available in the reporting system. The online aggregated reports were available to be downloaded in PDF, Excel, and comma-separated value formats.

CERS was the primary source for LEA staff to analyze ELPAC results at the LEA, school, grade, classroom, or customized group level. CERS provided these reports, which can be downloaded and used to inform instruction. LEA staff with TOMS logon credentials could enter CERS through the ELPAC website to access student assessment results.

#### 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.

#### 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.

##### Student Score Reports

The Summative 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 Alternate ELPAC performance levels: Fluent English Proficient, Intermediate EL, and Novice EL.

LEAs had three 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

The LEA ELPAC coordinator could forward the appropriate reports to test sites. In the case of a locally printed Alternate ELPAC SSR, the LEA sent the printed report(s) to the child’s parent/guardian. 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*](#_Toc121842306).

For the 2021–22 test administration, SSRs were made available to the LEAs in English, Spanish, Filipino, Chinese (Traditional), and Vietnamese. An SSR in a supported language was created if the student’s primary language as reported in the California Longitudinal 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.

Preliminary student scores and aggregations were also available to LEAs prior to the release of final reports via electronic reporting, using CERS. This website permitted LEAs to view preliminary results data for all tests taken.

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

#### 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 test 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 a content area 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. 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 band 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 or the *2021–22 ELPAC Post-Test Guide* (CDE, 2022a).

### References

California Department of Education. (2022a). *ELPAC post-test guide: Technical information for student score reports for ELPAC LEA and test site coordinators and research specialists.* Sacramento, CA: California Department of Education.

California Department of Education. (2022b). *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 and Scale Score Frequency Distributions

Table 7.A. Raw Score Frequency Distribution for Kindergarten

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 125 | 9.64 | 9.64 |
| 1 | 66 | 5.09 | 14.73 |
| 2 | 39 | 3.01 | 17.73 |
| 3 | 42 | 3.24 | 20.97 |
| 4 | 50 | 3.86 | 24.83 |
| 5 | 38 | 2.93 | 27.76 |
| 6 | 44 | 3.39 | 31.15 |
| 7 | 46 | 3.55 | 34.70 |
| 8 | 31 | 2.39 | 37.09 |
| 9 | 39 | 3.01 | 40.09 |
| 10 | 36 | 2.78 | 42.87 |
| 11 | 47 | 3.62 | 46.49 |
| 12 | 48 | 3.70 | 50.19 |
| 13 | 50 | 3.86 | 54.05 |
| 14 | 35 | 2.70 | 56.75 |
| 15 | 54 | 4.16 | 60.91 |
| 16 | 56 | 4.32 | 65.23 |
| 17 | 54 | 4.16 | 69.39 |
| 18 | 58 | 4.47 | 73.86 |
| 19 | 46 | 3.55 | 77.41 |
| 20 | 46 | 3.55 | 80.96 |
| 21 | 50 | 3.86 | 84.81 |
| 22 | 39 | 3.01 | 87.82 |
| 23 | 30 | 2.31 | 90.13 |
| 24 | 26 | 2.00 | 92.14 |
| 25 | 26 | 2.00 | 94.14 |
| 26 | 23 | 1.77 | 95.91 |
| 27 | 16 | 1.23 | 97.15 |
| 28 | 14 | 1.08 | 98.23 |
| 29 | 12 | 0.93 | 99.15 |
| 30 | 5 | 0.39 | 99.54 |
| 31 | 6 | 0.46 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 86 | 6.79 | 6.79 |
| 1 | 37 | 2.92 | 9.72 |
| 2 | 23 | 1.82 | 11.53 |
| 3 | 33 | 2.61 | 14.14 |
| 4 | 22 | 1.74 | 15.88 |
| 5 | 22 | 1.74 | 17.61 |
| 6 | 30 | 2.37 | 19.98 |
| 7 | 31 | 2.45 | 22.43 |
| 8 | 45 | 3.55 | 25.99 |
| 9 | 50 | 3.95 | 29.94 |
| 10 | 41 | 3.24 | 33.18 |
| 11 | 69 | 5.45 | 38.63 |
| 12 | 42 | 3.32 | 41.94 |
| 13 | 55 | 4.34 | 46.29 |
| 14 | 62 | 4.90 | 51.18 |
| 15 | 62 | 4.90 | 56.08 |
| 16 | 68 | 5.37 | 61.45 |
| 17 | 55 | 4.34 | 65.80 |
| 18 | 47 | 3.71 | 69.51 |
| 19 | 51 | 4.03 | 73.54 |
| 20 | 60 | 4.74 | 78.28 |
| 21 | 37 | 2.92 | 81.20 |
| 22 | 29 | 2.29 | 83.49 |
| 23 | 37 | 2.92 | 86.41 |
| 24 | 38 | 3.00 | 89.42 |
| 25 | 28 | 2.21 | 91.63 |
| 26 | 30 | 2.37 | 94.00 |
| 27 | 25 | 1.97 | 95.97 |
| 28 | 21 | 1.66 | 97.63 |
| 29 | 13 | 1.03 | 98.66 |
| 30 | 17 | 1.34 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 77 | 6.32 | 6.32 |
| 1 | 35 | 2.87 | 9.19 |
| 2 | 39 | 3.20 | 12.39 |
| 3 | 18 | 1.48 | 13.86 |
| 4 | 32 | 2.63 | 16.49 |
| 5 | 20 | 1.64 | 18.13 |
| 6 | 34 | 2.79 | 20.92 |
| 7 | 31 | 2.54 | 23.46 |
| 8 | 39 | 3.20 | 26.66 |
| 9 | 41 | 3.36 | 30.02 |
| 10 | 45 | 3.69 | 33.72 |
| 11 | 32 | 2.63 | 36.34 |
| 12 | 49 | 4.02 | 40.36 |
| 13 | 45 | 3.69 | 44.05 |
| 14 | 48 | 3.94 | 47.99 |
| 15 | 46 | 3.77 | 51.76 |
| 16 | 50 | 4.10 | 55.87 |
| 17 | 54 | 4.43 | 60.30 |
| 18 | 51 | 4.18 | 64.48 |
| 19 | 47 | 3.86 | 68.33 |
| 20 | 48 | 3.94 | 72.27 |
| 21 | 44 | 3.61 | 75.88 |
| 22 | 42 | 3.45 | 79.33 |
| 23 | 44 | 3.61 | 82.94 |
| 24 | 41 | 3.36 | 86.30 |
| 25 | 33 | 2.71 | 89.01 |
| 26 | 42 | 3.45 | 92.45 |
| 27 | 35 | 2.87 | 95.32 |
| 28 | 21 | 1.72 | 97.05 |
| 29 | 25 | 2.05 | 99.10 |
| 30 | 11 | 0.90 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Three Through Five, Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 88 | 4.10 | 4.10 |
| 1 | 33 | 1.54 | 5.63 |
| 2 | 33 | 1.54 | 7.17 |
| 3 | 19 | 0.88 | 8.05 |
| 4 | 29 | 1.35 | 9.40 |
| 5 | 24 | 1.12 | 10.52 |
| 6 | 37 | 1.72 | 12.24 |
| 7 | 45 | 2.09 | 14.34 |
| 8 | 55 | 2.56 | 16.90 |
| 9 | 45 | 2.09 | 18.99 |
| 10 | 64 | 2.98 | 21.97 |
| 11 | 71 | 3.31 | 25.28 |
| 12 | 67 | 3.12 | 28.40 |
| 13 | 73 | 3.40 | 31.80 |
| 14 | 62 | 2.89 | 34.68 |
| 15 | 83 | 3.86 | 38.55 |
| 16 | 79 | 3.68 | 42.23 |
| 17 | 80 | 3.72 | 45.95 |
| 18 | 73 | 3.40 | 49.35 |
| 19 | 65 | 3.03 | 52.37 |
| 20 | 95 | 4.42 | 56.80 |
| 21 | 47 | 2.19 | 58.99 |
| 22 | 53 | 2.47 | 61.45 |
| 23 | 77 | 3.58 | 65.04 |
| 24 | 67 | 3.12 | 68.16 |
| 25 | 74 | 3.45 | 71.60 |
| 26 | 84 | 3.91 | 75.51 |
| 27 | 105 | 4.89 | 80.40 |
| 28 | 98 | 4.56 | 84.96 |
| 29 | 116 | 5.40 | 90.36 |
| 30 | 115 | 5.35 | 95.72 |
| 31 | 92 | 4.28 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Three Through Five, Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 78 | 3.74 | 3.74 |
| 1 | 26 | 1.25 | 4.99 |
| 2 | 33 | 1.58 | 6.58 |
| 3 | 32 | 1.54 | 8.11 |
| 4 | 29 | 1.39 | 9.51 |
| 5 | 27 | 1.30 | 10.80 |
| 6 | 32 | 1.54 | 12.34 |
| 7 | 34 | 1.63 | 13.97 |
| 8 | 49 | 2.35 | 16.32 |
| 9 | 53 | 2.54 | 18.87 |
| 10 | 57 | 2.74 | 21.60 |
| 11 | 71 | 3.41 | 25.01 |
| 12 | 57 | 2.74 | 27.75 |
| 13 | 94 | 4.51 | 32.26 |
| 14 | 84 | 4.03 | 36.29 |
| 15 | 74 | 3.55 | 39.85 |
| 16 | 87 | 4.18 | 44.02 |
| 17 | 74 | 3.55 | 47.58 |
| 18 | 58 | 2.78 | 50.36 |
| 19 | 96 | 4.61 | 54.97 |
| 20 | 88 | 4.22 | 59.19 |
| 21 | 76 | 3.65 | 62.84 |
| 22 | 82 | 3.94 | 66.78 |
| 23 | 94 | 4.51 | 71.29 |
| 24 | 82 | 3.94 | 75.23 |
| 25 | 96 | 4.61 | 79.84 |
| 26 | 75 | 3.60 | 83.44 |
| 27 | 73 | 3.50 | 86.94 |
| 28 | 95 | 4.56 | 91.50 |
| 29 | 82 | 3.94 | 95.44 |
| 30 | 58 | 2.78 | 98.22 |
| 31 | 37 | 1.78 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Six Through Eight, Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 69 | 3.88 | 3.88 |
| 1 | 24 | 1.35 | 5.23 |
| 2 | 30 | 1.69 | 6.92 |
| 3 | 18 | 1.01 | 7.93 |
| 4 | 21 | 1.18 | 9.11 |
| 5 | 16 | 0.90 | 10.01 |
| 6 | 25 | 1.41 | 11.42 |
| 7 | 16 | 0.90 | 12.32 |
| 8 | 16 | 0.90 | 13.22 |
| 9 | 21 | 1.18 | 14.40 |
| 10 | 30 | 1.69 | 16.09 |
| 11 | 28 | 1.57 | 17.66 |
| 12 | 39 | 2.19 | 19.85 |
| 13 | 50 | 2.81 | 22.67 |
| 14 | 42 | 2.36 | 25.03 |
| 15 | 57 | 3.21 | 28.23 |
| 16 | 53 | 2.98 | 31.21 |
| 17 | 51 | 2.87 | 34.08 |
| 18 | 56 | 3.15 | 37.23 |
| 19 | 44 | 2.47 | 39.71 |
| 20 | 54 | 3.04 | 42.74 |
| 21 | 53 | 2.98 | 45.73 |
| 22 | 46 | 2.59 | 48.31 |
| 23 | 52 | 2.92 | 51.24 |
| 24 | 54 | 3.04 | 54.27 |
| 25 | 67 | 3.77 | 58.04 |
| 26 | 72 | 4.05 | 62.09 |
| 27 | 106 | 5.96 | 68.05 |
| 28 | 116 | 6.52 | 74.58 |
| 29 | 214 | 12.04 | 86.61 |
| 30 | 238 | 13.39 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Six Through Eight, Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 81 | 4.39 | 4.39 |
| 1 | 18 | 0.98 | 5.36 |
| 2 | 18 | 0.98 | 6.34 |
| 3 | 11 | 0.60 | 6.93 |
| 4 | 18 | 0.98 | 7.91 |
| 5 | 17 | 0.92 | 8.83 |
| 6 | 17 | 0.92 | 9.75 |
| 7 | 26 | 1.41 | 11.16 |
| 8 | 19 | 1.03 | 12.19 |
| 9 | 25 | 1.35 | 13.54 |
| 10 | 23 | 1.25 | 14.79 |
| 11 | 35 | 1.90 | 16.68 |
| 12 | 48 | 2.60 | 19.28 |
| 13 | 61 | 3.30 | 22.59 |
| 14 | 44 | 2.38 | 24.97 |
| 15 | 64 | 3.47 | 28.44 |
| 16 | 68 | 3.68 | 32.12 |
| 17 | 59 | 3.20 | 35.32 |
| 18 | 64 | 3.47 | 38.79 |
| 19 | 65 | 3.52 | 42.31 |
| 20 | 61 | 3.30 | 45.61 |
| 21 | 67 | 3.63 | 49.24 |
| 22 | 67 | 3.63 | 52.87 |
| 23 | 89 | 4.82 | 57.69 |
| 24 | 82 | 4.44 | 62.13 |
| 25 | 71 | 3.85 | 65.98 |
| 26 | 108 | 5.85 | 71.83 |
| 27 | 103 | 5.58 | 77.41 |
| 28 | 110 | 5.96 | 83.37 |
| 29 | 141 | 7.64 | 91.01 |
| 30 | 166 | 8.99 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Nine and Ten, Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 24 | 2.79 | 2.79 |
| 1 | 14 | 1.63 | 4.42 |
| 2 | 10 | 1.16 | 5.58 |
| 3 | 8 | 0.93 | 6.51 |
| 4 | 11 | 1.28 | 7.79 |
| 5 | 12 | 1.40 | 9.19 |
| 6 | 7 | 0.81 | 10.00 |
| 7 | 9 | 1.05 | 11.05 |
| 8 | 10 | 1.16 | 12.21 |
| 9 | 14 | 1.63 | 13.84 |
| 10 | 12 | 1.40 | 15.23 |
| 11 | 18 | 2.09 | 17.33 |
| 12 | 22 | 2.56 | 19.88 |
| 13 | 28 | 3.26 | 23.14 |
| 14 | 21 | 2.44 | 25.58 |
| 15 | 22 | 2.56 | 28.14 |
| 16 | 30 | 3.49 | 31.63 |
| 17 | 29 | 3.37 | 35.00 |
| 18 | 20 | 2.33 | 37.33 |
| 19 | 27 | 3.14 | 40.47 |
| 20 | 32 | 3.72 | 44.19 |
| 21 | 38 | 4.42 | 48.60 |
| 22 | 29 | 3.37 | 51.98 |
| 23 | 25 | 2.91 | 54.88 |
| 24 | 40 | 4.65 | 59.53 |
| 25 | 35 | 4.07 | 63.60 |
| 26 | 40 | 4.65 | 68.26 |
| 27 | 69 | 8.02 | 76.28 |
| 28 | 58 | 6.74 | 83.02 |
| 29 | 77 | 8.95 | 91.98 |
| 30 | 69 | 8.02 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Nine and Ten, Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 52 | 5.84 | 5.84 |
| 1 | 13 | 1.46 | 7.30 |
| 2 | 11 | 1.23 | 8.53 |
| 3 | 6 | 0.67 | 9.20 |
| 4 | 5 | 0.56 | 9.76 |
| 5 | 10 | 1.12 | 10.89 |
| 6 | 9 | 1.01 | 11.90 |
| 7 | 7 | 0.79 | 12.68 |
| 8 | 9 | 1.01 | 13.69 |
| 9 | 25 | 2.81 | 16.50 |
| 10 | 17 | 1.91 | 18.41 |
| 11 | 15 | 1.68 | 20.09 |
| 12 | 25 | 2.81 | 22.90 |
| 13 | 23 | 2.58 | 25.48 |
| 14 | 19 | 2.13 | 27.61 |
| 15 | 19 | 2.13 | 29.74 |
| 16 | 20 | 2.24 | 31.99 |
| 17 | 34 | 3.82 | 35.80 |
| 18 | 32 | 3.59 | 39.39 |
| 19 | 36 | 4.04 | 43.43 |
| 20 | 25 | 2.81 | 46.24 |
| 21 | 40 | 4.49 | 50.73 |
| 22 | 41 | 4.60 | 55.33 |
| 23 | 28 | 3.14 | 58.47 |
| 24 | 38 | 4.26 | 62.74 |
| 25 | 37 | 4.15 | 66.89 |
| 26 | 47 | 5.27 | 72.17 |
| 27 | 42 | 4.71 | 76.88 |
| 28 | 55 | 6.17 | 83.05 |
| 29 | 59 | 6.62 | 89.67 |
| 30 | 92 | 10.33 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Eleven and Twelve, Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 74 | 4.80 | 4.80 |
| 1 | 21 | 1.36 | 6.16 |
| 2 | 15 | 0.97 | 7.14 |
| 3 | 15 | 0.97 | 8.11 |
| 4 | 20 | 1.30 | 9.41 |
| 5 | 13 | 0.84 | 10.25 |
| 6 | 18 | 1.17 | 11.42 |
| 7 | 31 | 2.01 | 13.43 |
| 8 | 18 | 1.17 | 14.60 |
| 9 | 33 | 2.14 | 16.74 |
| 10 | 38 | 2.47 | 19.21 |
| 11 | 30 | 1.95 | 21.16 |
| 12 | 34 | 2.21 | 23.36 |
| 13 | 37 | 2.40 | 25.76 |
| 14 | 29 | 1.88 | 27.64 |
| 15 | 33 | 2.14 | 29.79 |
| 16 | 40 | 2.60 | 32.38 |
| 17 | 47 | 3.05 | 35.43 |
| 18 | 36 | 2.34 | 37.77 |
| 19 | 32 | 2.08 | 39.84 |
| 20 | 38 | 2.47 | 42.31 |
| 21 | 44 | 2.86 | 45.17 |
| 22 | 41 | 2.66 | 47.83 |
| 23 | 54 | 3.50 | 51.33 |
| 24 | 68 | 4.41 | 55.74 |
| 25 | 72 | 4.67 | 60.42 |
| 26 | 81 | 5.26 | 65.67 |
| 27 | 108 | 7.01 | 72.68 |
| 28 | 130 | 8.44 | 81.12 |
| 29 | 149 | 9.67 | 90.79 |
| 30 | 142 | 9.21 | 100.00 |

Table 7.A. Raw Score Frequency Distribution for Grade Span Eleven and Twelve, Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Frequency | Percent | Cumulative Percent |
| 0 | 76 | 4.37 | 4.37 |
| 1 | 30 | 1.72 | 6.09 |
| 2 | 23 | 1.32 | 7.41 |
| 3 | 26 | 1.49 | 8.91 |
| 4 | 13 | 0.75 | 9.66 |
| 5 | 13 | 0.75 | 10.40 |
| 6 | 23 | 1.32 | 11.72 |
| 7 | 20 | 1.15 | 12.87 |
| 8 | 30 | 1.72 | 14.60 |
| 9 | 26 | 1.49 | 16.09 |
| 10 | 27 | 1.55 | 17.64 |
| 11 | 28 | 1.61 | 19.25 |
| 12 | 38 | 2.18 | 21.44 |
| 13 | 47 | 2.70 | 24.14 |
| 14 | 42 | 2.41 | 26.55 |
| 15 | 35 | 2.01 | 28.56 |
| 16 | 47 | 2.70 | 31.26 |
| 17 | 58 | 3.33 | 34.60 |
| 18 | 51 | 2.93 | 37.53 |
| 19 | 69 | 3.97 | 41.49 |
| 20 | 50 | 2.87 | 44.37 |
| 21 | 57 | 3.28 | 47.64 |
| 22 | 56 | 3.22 | 50.86 |
| 23 | 65 | 3.74 | 54.60 |
| 24 | 75 | 4.31 | 58.91 |
| 25 | 66 | 3.79 | 62.70 |
| 26 | 100 | 5.75 | 68.45 |
| 27 | 84 | 4.83 | 73.28 |
| 28 | 114 | 6.55 | 79.83 |
| 29 | 123 | 7.07 | 86.90 |
| 30 | 228 | 13.10 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Kindergarten

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 201 | 125 | 9.64 | 9.64 |
| 214 | 66 | 5.09 | 14.73 |
| 222 | 39 | 3.01 | 17.73 |
| 227 | 42 | 3.24 | 20.97 |
| 230 | 50 | 3.86 | 24.83 |
| 233 | 38 | 2.93 | 27.76 |
| 235 | 44 | 3.39 | 31.15 |
| 237 | 46 | 3.55 | 34.70 |
| 239 | 31 | 2.39 | 37.09 |
| 241 | 39 | 3.01 | 40.09 |
| 242 | 36 | 2.78 | 42.87 |
| 244 | 47 | 3.62 | 46.49 |
| 245 | 48 | 3.70 | 50.19 |
| 247 | 50 | 3.86 | 54.05 |
| 248 | 35 | 2.70 | 56.75 |
| 249 | 54 | 4.16 | 60.91 |
| 251 | 56 | 4.32 | 65.23 |
| 252 | 54 | 4.16 | 69.39 |
| 253 | 58 | 4.47 | 73.86 |
| 255 | 46 | 3.55 | 77.41 |
| 256 | 46 | 3.55 | 80.96 |
| 258 | 50 | 3.86 | 84.81 |
| 260 | 39 | 3.01 | 87.82 |
| 261 | 30 | 2.31 | 90.13 |
| 263 | 26 | 2.00 | 92.14 |
| 265 | 26 | 2.00 | 94.14 |
| 268 | 23 | 1.77 | 95.91 |
| 271 | 16 | 1.23 | 97.15 |
| 274 | 14 | 1.08 | 98.23 |
| 279 | 12 | 0.93 | 99.15 |
| 287 | 5 | 0.39 | 99.54 |
| 299 | 6 | 0.46 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade One

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 301 | 86 | 6.79 | 6.79 |
| 311 | 37 | 2.92 | 9.72 |
| 319 | 23 | 1.82 | 11.53 |
| 324 | 33 | 2.61 | 14.14 |
| 328 | 22 | 1.74 | 15.88 |
| 331 | 22 | 1.74 | 17.61 |
| 334 | 30 | 2.37 | 19.98 |
| 336 | 31 | 2.45 | 22.43 |
| 338 | 45 | 3.55 | 25.99 |
| 340 | 50 | 3.95 | 29.94 |
| 341 | 41 | 3.24 | 33.18 |
| 343 | 69 | 5.45 | 38.63 |
| 345 | 42 | 3.32 | 41.94 |
| 346 | 55 | 4.34 | 46.29 |
| 347 | 62 | 4.90 | 51.18 |
| 349 | 62 | 4.90 | 56.08 |
| 350 | 68 | 5.37 | 61.45 |
| 352 | 55 | 4.34 | 65.80 |
| 353 | 47 | 3.71 | 69.51 |
| 355 | 51 | 4.03 | 73.54 |
| 356 | 60 | 4.74 | 78.28 |
| 358 | 37 | 2.92 | 81.20 |
| 359 | 29 | 2.29 | 83.49 |
| 361 | 37 | 2.92 | 86.41 |
| 363 | 38 | 3.00 | 89.42 |
| 365 | 28 | 2.21 | 91.63 |
| 368 | 30 | 2.37 | 94.00 |
| 372 | 25 | 1.97 | 95.97 |
| 376 | 21 | 1.66 | 97.63 |
| 384 | 13 | 1.03 | 98.66 |
| 399 | 17 | 1.34 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade Two

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 401 | 77 | 6.32 | 6.32 |
| 409 | 35 | 2.87 | 9.19 |
| 417 | 39 | 3.20 | 12.39 |
| 422 | 18 | 1.48 | 13.86 |
| 425 | 32 | 2.63 | 16.49 |
| 428 | 20 | 1.64 | 18.13 |
| 430 | 34 | 2.79 | 20.92 |
| 432 | 31 | 2.54 | 23.46 |
| 434 | 39 | 3.20 | 26.66 |
| 436 | 41 | 3.36 | 30.02 |
| 438 | 45 | 3.69 | 33.72 |
| 439 | 32 | 2.63 | 36.34 |
| 441 | 49 | 4.02 | 40.36 |
| 442 | 45 | 3.69 | 44.05 |
| 444 | 48 | 3.94 | 47.99 |
| 445 | 46 | 3.77 | 51.76 |
| 447 | 50 | 4.10 | 55.87 |
| 448 | 54 | 4.43 | 60.30 |
| 450 | 51 | 4.18 | 64.48 |
| 451 | 47 | 3.86 | 68.33 |
| 453 | 48 | 3.94 | 72.27 |
| 455 | 44 | 3.61 | 75.88 |
| 457 | 42 | 3.45 | 79.33 |
| 459 | 44 | 3.61 | 82.94 |
| 461 | 41 | 3.36 | 86.30 |
| 463 | 33 | 2.71 | 89.01 |
| 466 | 42 | 3.45 | 92.45 |
| 470 | 35 | 2.87 | 95.32 |
| 475 | 21 | 1.72 | 97.05 |
| 484 | 25 | 2.05 | 99.10 |
| 499 | 11 | 0.90 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade Span Three Through Five

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 501 | 166 | 3.92 | 3.92 |
| 515 | 59 | 1.39 | 5.32 |
| 522 | 66 | 1.56 | 6.88 |
| 526 | 51 | 1.21 | 8.08 |
| 529 | 58 | 1.37 | 9.45 |
| 532 | 51 | 1.21 | 10.66 |
| 534 | 69 | 1.63 | 12.29 |
| 535 | 34 | 0.80 | 13.09 |
| 536 | 45 | 1.06 | 14.16 |
| 537 | 104 | 2.46 | 16.62 |
| 539 | 98 | 2.32 | 18.93 |
| 540 | 121 | 2.86 | 21.79 |
| 541 | 71 | 1.68 | 23.47 |
| 542 | 71 | 1.68 | 25.15 |
| 543 | 124 | 2.93 | 28.08 |
| 544 | 167 | 3.95 | 32.03 |
| 545 | 146 | 3.45 | 35.48 |
| 546 | 83 | 1.96 | 37.44 |
| 547 | 153 | 3.62 | 41.05 |
| 548 | 167 | 3.95 | 45.00 |
| 549 | 147 | 3.47 | 48.48 |
| 551 | 123 | 2.91 | 51.38 |
| 552 | 191 | 4.51 | 55.90 |
| 553 | 135 | 3.19 | 59.09 |
| 555 | 129 | 3.05 | 62.14 |
| 556 | 159 | 3.76 | 65.89 |
| 557 | 67 | 1.58 | 67.48 |
| 558 | 94 | 2.22 | 69.70 |
| 559 | 156 | 3.69 | 73.39 |
| 561 | 180 | 4.25 | 77.64 |
| 563 | 180 | 4.25 | 81.90 |
| 565 | 73 | 1.73 | 83.62 |
| 566 | 98 | 2.32 | 85.94 |
| 568 | 95 | 2.25 | 88.18 |
| 570 | 116 | 2.74 | 90.92 |
| 572 | 82 | 1.94 | 92.86 |
| 576 | 115 | 2.72 | 95.58 |
| 579 | 58 | 1.37 | 96.95 |
| 599 | 129 | 3.05 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade Span Six Through Eight

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 601 | 150 | 4.14 | 4.14 |
| 613 | 42 | 1.16 | 5.30 |
| 620 | 48 | 1.32 | 6.62 |
| 624 | 29 | 0.80 | 7.42 |
| 627 | 39 | 1.08 | 8.50 |
| 629 | 16 | 0.44 | 8.94 |
| 630 | 17 | 0.47 | 9.41 |
| 631 | 25 | 0.69 | 10.10 |
| 632 | 17 | 0.47 | 10.57 |
| 633 | 16 | 0.44 | 11.01 |
| 634 | 26 | 0.72 | 11.73 |
| 635 | 35 | 0.97 | 12.69 |
| 636 | 21 | 0.58 | 13.27 |
| 637 | 25 | 0.69 | 13.96 |
| 638 | 53 | 1.46 | 15.42 |
| 639 | 28 | 0.77 | 16.20 |
| 640 | 74 | 2.04 | 18.24 |
| 641 | 48 | 1.32 | 19.56 |
| 642 | 111 | 3.06 | 22.63 |
| 643 | 42 | 1.16 | 23.79 |
| 644 | 101 | 2.79 | 26.57 |
| 645 | 117 | 3.23 | 29.80 |
| 647 | 119 | 3.28 | 33.08 |
| 648 | 115 | 3.17 | 36.26 |
| 649 | 108 | 2.98 | 39.24 |
| 651 | 119 | 3.28 | 42.52 |
| 652 | 114 | 3.15 | 45.67 |
| 654 | 113 | 3.12 | 48.79 |
| 655 | 52 | 1.43 | 50.22 |
| 656 | 67 | 1.85 | 52.07 |
| 657 | 143 | 3.95 | 56.02 |
| 659 | 149 | 4.11 | 60.13 |
| 661 | 143 | 3.95 | 64.07 |
| 664 | 214 | 5.91 | 69.98 |
| 667 | 103 | 2.84 | 72.82 |
| 668 | 116 | 3.20 | 76.02 |
| 671 | 110 | 3.04 | 79.06 |

Table 7.A.16 *(continuation)*

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 675 | 214 | 5.91 | 84.96 |
| 677 | 141 | 3.89 | 88.85 |
| 699 | 404 | 11.15 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade Span Nine and Ten

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 701 | 76 | 4.34 | 4.34 |
| 713 | 13 | 0.74 | 5.08 |
| 714 | 14 | 0.80 | 5.88 |
| 720 | 11 | 0.63 | 6.51 |
| 721 | 10 | 0.57 | 7.08 |
| 724 | 6 | 0.34 | 7.42 |
| 725 | 8 | 0.46 | 7.88 |
| 727 | 5 | 0.29 | 8.17 |
| 728 | 11 | 0.63 | 8.79 |
| 729 | 10 | 0.57 | 9.37 |
| 730 | 12 | 0.69 | 10.05 |
| 731 | 9 | 0.51 | 10.57 |
| 732 | 7 | 0.40 | 10.97 |
| 733 | 7 | 0.40 | 11.36 |
| 734 | 18 | 1.03 | 12.39 |
| 735 | 10 | 0.57 | 12.96 |
| 736 | 25 | 1.43 | 14.39 |
| 737 | 31 | 1.77 | 16.16 |
| 738 | 27 | 1.54 | 17.70 |
| 739 | 43 | 2.46 | 20.16 |
| 740 | 45 | 2.57 | 22.73 |
| 741 | 28 | 1.60 | 24.33 |
| 742 | 19 | 1.09 | 25.41 |
| 743 | 40 | 2.28 | 27.70 |
| 744 | 42 | 2.40 | 30.10 |
| 745 | 64 | 3.66 | 33.75 |
| 746 | 61 | 3.48 | 37.24 |
| 747 | 56 | 3.20 | 40.43 |
| 748 | 52 | 2.97 | 43.40 |
| 750 | 72 | 4.11 | 47.52 |
| 751 | 79 | 4.51 | 52.03 |
| 752 | 57 | 3.26 | 55.28 |
| 754 | 63 | 3.60 | 58.88 |
| 755 | 40 | 2.28 | 61.17 |
| 756 | 37 | 2.11 | 63.28 |
| 757 | 35 | 2.00 | 65.28 |
| 758 | 47 | 2.68 | 67.96 |
| 759 | 40 | 2.28 | 70.25 |

Table 7.A.17 *(continuation)*

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 760 | 42 | 2.40 | 72.64 |
| 762 | 69 | 3.94 | 76.58 |
| 764 | 55 | 3.14 | 79.73 |
| 766 | 58 | 3.31 | 83.04 |
| 769 | 59 | 3.37 | 86.41 |
| 771 | 77 | 4.40 | 90.81 |
| 799 | 161 | 9.19 | 100.00 |

Table 7.A. Scale Score Frequency Distribution for Grade Span Eleven and Twelve

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 801 | 150 | 4.57 | 4.57 |
| 813 | 30 | 0.91 | 5.49 |
| 814 | 21 | 0.64 | 6.13 |
| 820 | 23 | 0.70 | 6.83 |
| 821 | 15 | 0.46 | 7.28 |
| 824 | 26 | 0.79 | 8.08 |
| 825 | 15 | 0.46 | 8.53 |
| 827 | 13 | 0.40 | 8.93 |
| 828 | 20 | 0.61 | 9.54 |
| 829 | 13 | 0.40 | 9.94 |
| 830 | 13 | 0.40 | 10.33 |
| 831 | 23 | 0.70 | 11.03 |
| 832 | 18 | 0.55 | 11.58 |
| 833 | 20 | 0.61 | 12.19 |
| 834 | 61 | 1.86 | 14.05 |
| 835 | 18 | 0.55 | 14.60 |
| 836 | 26 | 0.79 | 15.39 |
| 837 | 60 | 1.83 | 17.22 |
| 838 | 66 | 2.01 | 19.23 |
| 839 | 68 | 2.07 | 21.30 |
| 840 | 81 | 2.47 | 23.77 |
| 841 | 37 | 1.13 | 24.90 |
| 842 | 42 | 1.28 | 26.18 |
| 843 | 64 | 1.95 | 28.13 |
| 844 | 80 | 2.44 | 30.57 |
| 845 | 98 | 2.99 | 33.56 |
| 846 | 98 | 2.99 | 36.54 |
| 847 | 105 | 3.20 | 39.74 |
| 848 | 82 | 2.50 | 42.24 |
| 850 | 95 | 2.90 | 45.14 |
| 851 | 100 | 3.05 | 48.19 |
| 852 | 106 | 3.23 | 51.42 |
| 854 | 129 | 3.93 | 55.35 |
| 855 | 68 | 2.07 | 57.42 |
| 856 | 66 | 2.01 | 59.43 |
| 857 | 72 | 2.19 | 61.63 |
| 858 | 100 | 3.05 | 64.68 |
| 859 | 81 | 2.47 | 67.14 |

Table 7.A.18 *(continuation)*

|  |  |  |  |
| --- | --- | --- | --- |
| Scale Score | Frequency | Percent | Cumulative Percent |
| 860 | 84 | 2.56 | 69.70 |
| 862 | 108 | 3.29 | 73.00 |
| 864 | 114 | 3.47 | 76.47 |
| 866 | 130 | 3.96 | 80.43 |
| 869 | 123 | 3.75 | 84.18 |
| 871 | 149 | 4.54 | 88.72 |
| 899 | 370 | 11.28 | 100.00 |

### Appendix 7.B: Means and Standard Deviations of Scale Scores and Percentage of Students in Each Performance Level by Demographic Student Group

**Note:** In table 7.B.1 through table 7.B.7, to protect student privacy, when the number of students in a student group is 10 or fewer, the summary statistics of scale scores and percentage of students in each performance level are not reported, and are replaced by “N/A.”

Table 7.B. 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 | 1,297 | 241.80 | 19 | 43 | 42 | 15 |
| Male | 944 | 241.37 | 20 | 44 | 41 | 15 |
| Female | 352 | 242.94 | 19 | 40 | 44 | 16 |
| Nonbinary | 1 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A | N/A | N/A | N/A |
| Asian | 227 | 241.85 | 18 | 43 | 43 | 14 |
| Native Hawaiian or Other Pacific Islander | 4 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 20 | 246.95 | 23 | 45 | 25 | 30 |
| Hispanic or Latino | 943 | 241.91 | 20 | 42 | 42 | 15 |
| Black or African American | 8 | N/A | N/A | N/A | N/A | N/A |
| White | 68 | 240.90 | 19 | 46 | 40 | 15 |
| Two or more races | 25 | 240.56 | 22 | 44 | 44 | 12 |
| Intellectual disability | 300 | 239.56 | 18 | 51 | 40 | 9 |
| Hearing impairment | 4 | N/A | N/A | N/A | N/A | N/A |
| Speech or language impairment | 59 | 256.80 | 15 | 10 | 56 | 34 |
| Visual impairment | 2 | 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 | 31 | 240.23 | 20 | 42 | 48 | 10 |
| Other health impairment | 56 | 241.07 | 21 | 34 | 48 | 18 |
| Specific learning disability | 5 | 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 | 67 | 226.54 | 21 | 76 | 22 | 1 |

Table 7.B.1 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| Autism | 771 | 242.88 | 19 | 40 | 43 | 17 |
| Traumatic brain injury | 2 | 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 | 385 | 240.26 | 20 | 45 | 41 | 15 |
| Economically disadvantaged | 912 | 242.45 | 19 | 42 | 43 | 15 |
| In US schools less than 12 months | 1,046 | 240.87 | 20 | 45 | 41 | 14 |
| In US schools 12 months or more | 223 | 245.90 | 18 | 35 | 44 | 20 |
| Duration unknown | 28 | 244.11 | 20 | 36 | 46 | 18 |
| Migrant education | 6 | N/A | N/A | N/A | N/A | N/A |
| Not migrant education | 1,291 | 241.78 | 19 | 43 | 42 | 15 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade One

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 1,266 | 345.29 | 19 | 39 | 45 | 17 |
| Male | 888 | 345.87 | 19 | 38 | 44 | 18 |
| Female | 378 | 343.93 | 19 | 41 | 46 | 13 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A | N/A | N/A | N/A |
| Asian | 202 | 346.66 | 20 | 35 | 45 | 20 |
| Native Hawaiian or Other Pacific Islander | 3 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 18 | 345.28 | 15 | 39 | 56 | 6 |
| Hispanic or Latino | 936 | 345.26 | 19 | 39 | 45 | 16 |
| Black or African American | 17 | 332.88 | 23 | 59 | 29 | 12 |
| White | 69 | 345.71 | 17 | 42 | 43 | 14 |
| Two or more races | 19 | 341.11 | 17 | 42 | 53 | 5 |
| Intellectual disability | 365 | 343.17 | 17 | 44 | 47 | 9 |
| Hearing impairment | 7 | N/A | N/A | N/A | N/A | N/A |
| Speech or language impairment | 41 | 359.39 | 16 | 7 | 41 | 51 |
| Visual impairment | 2 | N/A | N/A | N/A | N/A | N/A |
| Emotional impairment | 2 | N/A | N/A | N/A | N/A | N/A |
| Orthopedic impairment | 29 | 341.41 | 25 | 52 | 38 | 10 |
| Other health impairment | 63 | 346.30 | 20 | 35 | 43 | 22 |
| Specific learning disability | 4 | N/A | N/A | N/A | N/A | N/A |
| Deaf-blindness | 1 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 58 | 329.43 | 22 | 66 | 33 | 2 |
| Autism | 691 | 346.75 | 18 | 36 | 45 | 19 |
| Traumatic brain injury | 2 | N/A | N/A | N/A | N/A | N/A |
| Not classified | 1 | N/A | N/A | N/A | N/A | N/A |
| Not economically disadvantaged | 340 | 346.22 | 20 | 35 | 47 | 18 |
| Economically disadvantaged | 926 | 344.95 | 19 | 40 | 44 | 16 |
| In US schools less than 12 months | 65 | 337.97 | 20 | 57 | 38 | 5 |
| In US schools 12 months or more | 1,178 | 345.79 | 19 | 37 | 46 | 17 |
| Duration unknown | 23 | 340.65 | 21 | 52 | 22 | 26 |
| Migrant education | 11 | 345.64 | 14 | 36 | 45 | 18 |
| Not migrant education | 1,255 | 345.29 | 19 | 39 | 45 | 16 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Two

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 1,219 | 443.54 | 19 | 44 | 39 | 17 |
| Male | 856 | 444.66 | 19 | 42 | 38 | 19 |
| Female | 363 | 440.89 | 19 | 48 | 40 | 12 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A | N/A | N/A | N/A |
| Asian | 191 | 442.07 | 21 | 48 | 34 | 19 |
| Native Hawaiian or Other Pacific Islander | 8 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 21 | 442.43 | 24 | 48 | 33 | 19 |
| Hispanic or Latino | 931 | 443.86 | 19 | 43 | 40 | 17 |
| Black or African American | 9 | N/A | N/A | N/A | N/A | N/A |
| White | 44 | 444.91 | 25 | 41 | 32 | 27 |
| Two or more races | 13 | 446.00 | 7 | 46 | 46 | 8 |
| Intellectual disability | 414 | 441.49 | 17 | 50 | 39 | 12 |
| Hearing impairment | 11 | 445.82 | 18 | 36 | 45 | 18 |
| Speech or language impairment | 39 | 456.00 | 14 | 15 | 56 | 28 |
| Visual impairment | 2 | 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 | 28 | 433.32 | 22 | 61 | 32 | 7 |
| Other health impairment | 57 | 443.56 | 19 | 40 | 39 | 21 |
| Specific learning disability | 5 | N/A | N/A | N/A | N/A | N/A |
| Deaf-blindness | 1 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 67 | 432.64 | 21 | 63 | 28 | 9 |
| Autism | 593 | 445.87 | 20 | 40 | 39 | 21 |
| Traumatic brain injury | 2 | 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 | 321 | 444.13 | 20 | 43 | 39 | 18 |
| Economically disadvantaged | 898 | 443.33 | 19 | 45 | 39 | 17 |
| In US schools less than 12 months | 15 | 435.60 | 29 | 73 | 13 | 13 |
| In US schools 12 months or more | 1,193 | 443.65 | 19 | 44 | 39 | 17 |
| Duration unknown | 11 | 442.55 | 24 | 64 | 18 | 18 |
| Migrant education | 8 | N/A | N/A | N/A | N/A | N/A |
| Not migrant education | 1,211 | 443.48 | 19 | 44 | 39 | 17 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Three Through Five

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 4,231 | 550.61 | 18 | 28 | 45 | 27 |
| Male | 2,909 | 550.99 | 18 | 28 | 45 | 28 |
| Female | 1,322 | 549.77 | 18 | 29 | 46 | 24 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 8 | N/A | N/A | N/A | N/A | N/A |
| Asian | 611 | 548.31 | 18 | 30 | 50 | 20 |
| Native Hawaiian or Other Pacific Islander | 14 | 542.29 | 26 | 50 | 36 | 14 |
| Filipino | 91 | 546.82 | 16 | 37 | 51 | 12 |
| Hispanic or Latino | 3,248 | 551.22 | 18 | 27 | 44 | 28 |
| Black or African American | 31 | 551.48 | 17 | 35 | 42 | 23 |
| White | 168 | 550.21 | 19 | 28 | 44 | 28 |
| Two or more races | 60 | 550.38 | 18 | 28 | 50 | 22 |
| Intellectual disability | 1,663 | 549.88 | 16 | 29 | 49 | 23 |
| Hearing impairment | 23 | 556.65 | 15 | 13 | 48 | 39 |
| Speech or language impairment | 78 | 561.97 | 12 | 5 | 38 | 56 |
| Visual impairment | 5 | N/A | N/A | N/A | N/A | N/A |
| Emotional impairment | 8 | N/A | N/A | N/A | N/A | N/A |
| Orthopedic impairment | 108 | 543.41 | 24 | 36 | 37 | 27 |
| Other health impairment | 204 | 557.31 | 17 | 16 | 42 | 42 |
| Specific learning disability | 152 | 570.11 | 13 |  | 18 | 82 |
| Deaf-blindness | 0 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 198 | 531.45 | 21 | 63 | 32 | 5 |
| Autism | 1,786 | 550.76 | 17 | 28 | 47 | 24 |
| Traumatic brain injury | 6 | 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 | 1,046 | 548.67 | 18 | 31 | 48 | 21 |
| Economically disadvantaged | 3,185 | 551.24 | 18 | 27 | 45 | 28 |
| In US schools less than 12 months | 46 | 538.46 | 21 | 52 | 33 | 15 |
| In US schools 12 months or more | 4,153 | 550.75 | 18 | 28 | 45 | 27 |
| Duration unknown | 32 | 549.06 | 17 | 31 | 53 | 16 |
| Migrant education | 50 | 550.14 | 16 | 24 | 52 | 24 |
| Not migrant education | 4,181 | 550.61 | 18 | 28 | 45 | 27 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Six Through Eight

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 3,624 | 656.35 | 23 | 24 | 36 | 40 |
| Male | 2,368 | 656.95 | 23 | 23 | 36 | 41 |
| Female | 1,256 | 655.22 | 23 | 25 | 37 | 39 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 8 | N/A | N/A | N/A | N/A | N/A |
| Asian | 483 | 654.16 | 21 | 26 | 40 | 34 |
| Native Hawaiian or Other Pacific Islander | 10 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 61 | 657.02 | 18 | 11 | 54 | 34 |
| Hispanic or Latino | 2,837 | 656.94 | 23 | 23 | 35 | 41 |
| Black or African American | 29 | 658.03 | 27 | 28 | 24 | 48 |
| White | 162 | 652.15 | 25 | 30 | 36 | 34 |
| Two or more races | 34 | 653.53 | 23 | 29 | 32 | 38 |
| Intellectual disability | 1,709 | 655.53 | 21 | 23 | 40 | 36 |
| Hearing impairment | 20 | 658.05 | 21 | 15 | 40 | 45 |
| Speech or language impairment | 42 | 671.55 | 17 | 0 | 29 | 71 |
| Visual impairment | 9 | N/A | N/A | N/A | N/A | N/A |
| Emotional impairment | 8 | N/A | N/A | N/A | N/A | N/A |
| Orthopedic impairment | 80 | 647.28 | 25 | 40 | 31 | 29 |
| Other health impairment | 173 | 666.05 | 23 | 12 | 29 | 60 |
| Specific learning disability | 227 | 677.26 | 16 | 0 | 16 | 84 |
| Deaf-blindness | 4 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 197 | 629.77 | 24 | 70 | 17 | 13 |
| Autism | 1,148 | 656.60 | 21 | 23 | 40 | 37 |
| Traumatic brain injury | 7 | 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 | 829 | 653.58 | 23 | 29 | 37 | 35 |
| Economically disadvantaged | 2,795 | 657.17 | 22 | 22 | 36 | 41 |
| In US schools less than 12 months | 28 | 645.04 | 22 | 46 | 29 | 25 |
| In US schools 12 months or more | 3,572 | 656.42 | 23 | 24 | 37 | 40 |
| Duration unknown | 24 | 658.92 | 28 | 29 | 17 | 54 |
| Migrant education | 59 | 661.49 | 18 | 14 | 32 | 54 |
| Not migrant education | 3,565 | 656.26 | 23 | 24 | 36 | 40 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Nine and Ten

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 1,751 | 752.46 | 21 | 28 | 43 | 30 |
| Male | 1,127 | 753.11 | 21 | 26 | 42 | 32 |
| Female | 624 | 751.28 | 22 | 31 | 43 | 26 |
| Nonbinary | 0 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 1 | N/A | N/A | N/A | N/A | N/A |
| Asian | 231 | 749.59 | 19 | 29 | 46 | 25 |
| Native Hawaiian or Other Pacific Islander | 9 | N/A | N/A | N/A | N/A | N/A |
| Filipino | 35 | 748.63 | 22 | 31 | 43 | 26 |
| Hispanic or Latino | 1,362 | 753.19 | 22 | 27 | 42 | 31 |
| Black or African American | 9 | N/A | N/A | N/A | N/A | N/A |
| White | 78 | 753.47 | 21 | 24 | 40 | 36 |
| Two or more races | 26 | 738.15 | 25 | 58 | 27 | 15 |
| Intellectual disability | 892 | 752.02 | 19 | 27 | 48 | 25 |
| Hearing impairment | 12 | 758.08 | 14 | 0 | 58 | 42 |
| Speech or language impairment | 17 | 761.71 | 12 | 0 | 47 | 53 |
| Visual impairment | 10 | N/A | N/A | N/A | N/A | N/A |
| Emotional impairment | 14 | 767.71 | 15 | 0 | 29 | 71 |
| Orthopedic impairment | 40 | 738.18 | 26 | 58 | 23 | 20 |
| Other health impairment | 59 | 763.19 | 25 | 12 | 34 | 54 |
| Specific learning disability | 94 | 774.12 | 19 | 2 | 23 | 74 |
| Deaf-blindness | 0 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 125 | 732.49 | 22 | 67 | 27 | 6 |
| Autism | 479 | 753.32 | 20 | 25 | 44 | 30 |
| Traumatic brain injury | 8 | N/A | N/A | N/A | N/A | N/A |
| Not classified | 1 | N/A | N/A | N/A | N/A | N/A |
| Not economically disadvantaged | 422 | 748.70 | 20 | 34 | 43 | 24 |
| Economically disadvantaged | 1,329 | 753.65 | 22 | 26 | 43 | 32 |
| In US schools less than 12 months | 10 | N/A | N/A | N/A | N/A | N/A |
| In US schools 12 months or more | 1,734 | 752.51 | 21 | 28 | 43 | 30 |
| Duration unknown | 7 | N/A | N/A | N/A | N/A | N/A |
| Migrant education | 11 | 754.36 | 18 | 18 | 64 | 18 |
| Not migrant education | 1,740 | 752.45 | 21 | 28 | 42 | 30 |

Table 7.B. Mean and SD of Scale Scores and Percentage of Students in Each Performance Level by Student Group, Grade Span Eleven and Twelve

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Student Group | Number of Students Tested | Scale Score Mean | Scale Score SD | Performance Level 1 | Performance Level 2 | Performance Level 3 |
| All | 3,281 | 853.52 | 23 | 28 | 39 | 33 |
| Male | 2,131 | 853.99 | 23 | 28 | 38 | 34 |
| Female | 1,149 | 852.69 | 22 | 29 | 40 | 31 |
| Nonbinary | 1 | N/A | N/A | N/A | N/A | N/A |
| American Indian or Alaska Native | 3 | N/A | N/A | N/A | N/A | N/A |
| Asian | 455 | 851.39 | 21 | 31 | 41 | 28 |
| Native Hawaiian or Other Pacific Islander | 14 | 855.57 | 17 | 29 | 36 | 36 |
| Filipino | 69 | 857.83 | 27 | 23 | 38 | 39 |
| Hispanic or Latino | 2,572 | 853.93 | 23 | 27 | 39 | 34 |
| Black or African American | 26 | 851.88 | 21 | 35 | 35 | 31 |
| White | 122 | 849.95 | 19 | 31 | 46 | 23 |
| Two or more races | 20 | 857.05 | 27 | 35 | 25 | 40 |
| Intellectual disability | 1,668 | 853.59 | 20 | 26 | 44 | 30 |
| Hearing impairment | 32 | 856.09 | 19 | 19 | 50 | 31 |
| Speech or language impairment | 13 | 868.54 | 19 | 0 | 38 | 62 |
| Visual impairment | 16 | 843.31 | 26 | 56 | 19 | 25 |
| Emotional impairment | 18 | 863.50 | 17 | 11 | 17 | 72 |
| Orthopedic impairment | 130 | 847.55 | 27 | 38 | 35 | 27 |
| Other health impairment | 102 | 865.72 | 24 | 10 | 35 | 55 |
| Specific learning disability | 122 | 875.93 | 18 | 2 | 17 | 81 |
| Deaf-blindness | 4 | N/A | N/A | N/A | N/A | N/A |
| Multiple disabilities | 194 | 829.87 | 24 | 70 | 20 | 10 |
| Autism | 960 | 854.57 | 22 | 28 | 38 | 34 |
| Traumatic brain injury | 21 | 856.71 | 21 | 14 | 43 | 43 |
| Not classified | 1 | N/A | N/A | N/A | N/A | N/A |
| Not economically disadvantaged | 918 | 850.92 | 23 | 32 | 39 | 29 |
| Economically disadvantaged | 2,363 | 854.53 | 22 | 26 | 39 | 34 |
| In US schools less than 12 months | 6 | N/A | N/A | N/A | N/A | N/A |
| In US schools 12 months or more | 3,249 | 853.60 | 23 | 28 | 39 | 33 |
| Duration unknown | 26 | 842.73 | 17 | 42 | 46 | 12 |
| Migrant education | 30 | 858.97 | 13 | 0 | 60 | 40 |
| Not migrant education | 3,251 | 853.47 | 23 | 28 | 39 | 33 |

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

**Notes for table 7.C.1 through table 7.C.6:**

* The student group “All” represents all students who took a test.
* The *Number Tested* column contains the number of students in each demographic student group who took the test.
* The *Number Analyzed* column contains the number of students included in item analyses after data cleaning rules were applied:
* *Number Analyzed*—Students who met the attemptedness criteria, which is defined as having responded to at least one item in the expressive and one item in the receptive communication modes.
* The *Percent Tested* and *Percent Analyzed* columns are both relative to the *Number Registered* count.

Table 7.C. **Demographic Summary for Students: Kindergarten**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 1,431 | 1,336 | 93.36 | 1,297 | 90.64 |
| Male | 1,033 | 971 | 94.00 | 944 | 91.38 |
| Female | 397 | 364 | 91.69 | 352 | 88.66 |
| Nonbinary | 1 | 1 | 100.00 | 1 | 100.00 |
| American Indian or Alaska Native | 2 | 2 | 100.00 | 2 | 100.00 |
| Asian | 241 | 232 | 96.27 | 227 | 94.19 |
| Native Hawaiian or Other Pacific Islander | 4 | 4 | 100.00 | 4 | 100.00 |
| Filipino | 21 | 21 | 100.00 | 20 | 95.24 |
| Hispanic or Latino | 1,051 | 972 | 92.48 | 943 | 89.72 |
| Black or African American | 12 | 10 | 83.33 | 8 | 66.67 |
| White | 72 | 70 | 97.22 | 68 | 94.44 |
| Two or more races | 28 | 25 | 89.29 | 25 | 89.29 |
| Intellectual disability | 329 | 309 | 93.92 | 300 | 91.19 |
| Hearing impairment | 4 | 4 | 100.00 | 4 | 100.00 |
| Speech or language impairment | 64 | 60 | 93.75 | 59 | 92.19 |
| Visual impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 36 | 35 | 97.22 | 31 | 86.11 |
| Other health impairment | 71 | 59 | 83.10 | 56 | 78.87 |

Table 7.C.1 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| Specific learning disability | 5 | 5 | 100.00 | 5 | 100.00 |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 83 | 70 | 84.34 | 67 | 80.72 |
| Autism | 835 | 790 | 94.61 | 771 | 92.34 |
| Traumatic brain injury | 2 | 2 | 100.00 | 2 | 100.00 |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 420 | 390 | 92.86 | 385 | 91.67 |
| Economically disadvantaged | 1,011 | 946 | 93.57 | 912 | 90.21 |
| In US schools less than 12 months | 1,155 | 1,078 | 93.33 | 1,046 | 90.56 |
| In US schools 12 months or more | 244 | 229 | 93.85 | 223 | 91.39 |
| Duration unknown | 32 | 29 | 90.63 | 28 | 87.50 |
| Migrant education | 7 | 6 | 85.71 | 6 | 85.71 |
| Not migrant education | 1,424 | 1,330 | 93.40 | 1,291 | 90.66 |

Table 7.C. **Demographic Summary for Students: Grade One**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 1,401 | 1,301 | 92.86 | 1,266 | 90.36 |
| Male | 985 | 914 | 92.79 | 888 | 90.15 |
| Female | 416 | 387 | 93.03 | 378 | 90.87 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 2 | 2 | 100.00 | 2 | 100.00 |
| Asian | 227 | 208 | 91.63 | 202 | 88.99 |
| Native Hawaiian or Other Pacific Islander | 3 | 3 | 100.00 | 3 | 100.00 |
| Filipino | 20 | 18 | 90.00 | 18 | 90.00 |
| Hispanic or Latino | 1,032 | 964 | 93.41 | 936 | 90.70 |
| Black or African American | 18 | 17 | 94.44 | 17 | 94.44 |
| White | 75 | 69 | 92.00 | 69 | 92.00 |
| Two or more races | 24 | 20 | 83.33 | 19 | 79.17 |
| Intellectual disability | 403 | 378 | 93.80 | 365 | 90.57 |
| Hearing impairment | 9 | 7 | 77.78 | 7 | 77.78 |
| Speech or language impairment | 43 | 41 | 95.35 | 41 | 95.35 |
| Visual impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Emotional impairment | 2 | 2 | 100.00 | 2 | 100.00 |
| Orthopedic impairment | 39 | 32 | 82.05 | 29 | 74.36 |
| Other health impairment | 71 | 64 | 90.14 | 63 | 88.73 |
| Specific learning disability | 4 | 4 | 100.00 | 4 | 100.00 |
| Deaf-blindness | 1 | 1 | 100.00 | 1 | 100.00 |
| Multiple disabilities | 82 | 65 | 79.27 | 58 | 70.73 |
| Autism | 742 | 702 | 94.61 | 691 | 93.13 |
| Traumatic brain injury | 2 | 2 | 100.00 | 2 | 100.00 |
| Not classified | 1 | 1 | 100.00 | 1 | 100.00 |
| Not economically disadvantaged | 385 | 349 | 90.65 | 340 | 88.31 |
| Economically disadvantaged | 1,016 | 952 | 93.70 | 926 | 91.14 |
| In US schools less than 12 months | 74 | 67 | 90.54 | 65 | 87.84 |
| In US schools 12 months or more | 1,299 | 1,211 | 93.23 | 1,178 | 90.69 |
| Duration unknown | 28 | 23 | 82.14 | 23 | 82.14 |
| Migrant education | 11 | 11 | 100.00 | 11 | 100.00 |
| Not migrant education | 1,390 | 1,290 | 92.81 | 1,255 | 90.29 |

Table 7.C. **Demographic Summary for Students: Grade Two**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 1,346 | 1,238 | 91.98 | 1,219 | 90.56 |
| Male | 939 | 870 | 92.65 | 856 | 91.16 |
| Female | 407 | 368 | 90.42 | 363 | 89.19 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 2 | 2 | 100.00 | 2 | 100.00 |
| Asian | 215 | 194 | 90.23 | 191 | 88.84 |
| Native Hawaiian or Other Pacific Islander | 8 | 8 | 100.00 | 8 | 100.00 |
| Filipino | 22 | 21 | 95.45 | 21 | 95.45 |
| Hispanic or Latino | 1,020 | 945 | 92.65 | 931 | 91.27 |
| Black or African American | 11 | 9 | 81.82 | 9 | 81.82 |
| White | 53 | 46 | 86.79 | 44 | 83.02 |
| Two or more races | 15 | 13 | 86.67 | 13 | 86.67 |
| Intellectual disability | 444 | 418 | 94.14 | 414 | 93.24 |
| Hearing impairment | 13 | 12 | 92.31 | 11 | 84.62 |
| Speech or language impairment | 41 | 39 | 95.12 | 39 | 95.12 |
| Visual impairment | 3 | 3 | 100.00 | 2 | 66.67 |
| Emotional impairment | 0 | 0 | N/A | 0 | N/A |
| Orthopedic impairment | 37 | 28 | 75.68 | 28 | 75.68 |
| Other health impairment | 67 | 59 | 88.06 | 57 | 85.07 |
| Specific learning disability | 5 | 5 | 100.00 | 5 | 100.00 |
| Deaf-blindness | 1 | 1 | 100.00 | 1 | 100.00 |
| Multiple disabilities | 91 | 69 | 75.82 | 67 | 73.63 |
| Autism | 641 | 601 | 93.76 | 593 | 92.51 |
| Traumatic brain injury | 2 | 2 | 100.00 | 2 | 100.00 |
| Not classified | 1 | 1 | 100.00 | 0 | 0.00 |
| Not economically disadvantaged | 365 | 326 | 89.32 | 321 | 87.95 |
| Economically disadvantaged | 981 | 912 | 92.97 | 898 | 91.54 |
| In US schools less than 12 months | 16 | 15 | 93.75 | 15 | 93.75 |
| In US schools 12 months or more | 1,319 | 1,212 | 91.89 | 1,193 | 90.45 |
| Duration unknown | 11 | 11 | 100.00 | 11 | 100.00 |
| Migrant education | 8 | 8 | 100.00 | 8 | 100.00 |
| Not migrant education | 1,338 | 1,230 | 91.93 | 1,211 | 90.51 |

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 4,799 | 4,300 | 89.60 | 4,231 | 88.16 |
| Male | 3,260 | 2,947 | 90.40 | 2,909 | 89.23 |
| Female | 1,539 | 1,353 | 87.91 | 1,322 | 85.90 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 9 | 8 | 88.89 | 8 | 88.89 |
| Asian | 708 | 623 | 87.99 | 611 | 86.30 |
| Native Hawaiian or Other Pacific Islander | 18 | 15 | 83.33 | 14 | 77.78 |
| Filipino | 102 | 92 | 90.20 | 91 | 89.22 |
| Hispanic or Latino | 3,654 | 3,297 | 90.23 | 3,248 | 88.89 |
| Black or African American | 40 | 33 | 82.50 | 31 | 77.50 |
| White | 194 | 170 | 87.63 | 168 | 86.60 |
| Two or more races | 74 | 62 | 83.78 | 60 | 81.08 |
| Intellectual disability | 1,845 | 1,685 | 91.33 | 1,663 | 90.14 |
| Hearing impairment | 31 | 24 | 77.42 | 23 | 74.19 |
| Speech or language impairment | 90 | 78 | 86.67 | 78 | 86.67 |
| Visual impairment | 12 | 8 | 66.67 | 5 | 41.67 |
| Emotional impairment | 12 | 8 | 66.67 | 8 | 66.67 |
| Orthopedic impairment | 152 | 112 | 73.68 | 108 | 71.05 |
| Other health impairment | 235 | 204 | 86.81 | 204 | 86.81 |
| Specific learning disability | 160 | 152 | 95.00 | 152 | 95.00 |
| Deaf-blindness | 0 | 0 | N/A | 0 | N/A |
| Multiple disabilities | 274 | 213 | 77.74 | 198 | 72.26 |
| Autism | 1,981 | 1,810 | 91.37 | 1,786 | 90.16 |
| Traumatic brain injury | 7 | 6 | 85.71 | 6 | 85.71 |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 1,239 | 1,065 | 85.96 | 1,046 | 84.42 |
| Economically disadvantaged | 3,560 | 3,235 | 90.87 | 3,185 | 89.47 |
| In US schools less than 12 months | 62 | 47 | 75.81 | 46 | 74.19 |
| In US schools 12 months or more | 4,702 | 4,220 | 89.75 | 4,153 | 88.32 |
| Duration unknown | 35 | 33 | 94.29 | 32 | 91.43 |
| Migrant education | 54 | 50 | 92.59 | 50 | 92.59 |
| Not migrant education | 4,745 | 4,250 | 89.57 | 4,181 | 88.11 |

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 4,315 | 3,676 | 85.19 | 3,624 | 83.99 |
| Male | 2,821 | 2,402 | 85.15 | 2,368 | 83.94 |
| Female | 1,494 | 1,274 | 85.27 | 1,256 | 84.07 |
| Nonbinary | 0 | 0 | N/A | 0 | N/A |
| American Indian or Alaska Native | 8 | 8 | 100.00 | 8 | 100.00 |
| Asian | 588 | 491 | 83.50 | 483 | 82.14 |
| Native Hawaiian or Other Pacific Islander | 14 | 10 | 71.43 | 10 | 71.43 |
| Filipino | 80 | 63 | 78.75 | 61 | 76.25 |
| Hispanic or Latino | 3,354 | 2,876 | 85.75 | 2,837 | 84.59 |
| Black or African American | 35 | 30 | 85.71 | 29 | 82.86 |
| White | 192 | 164 | 85.42 | 162 | 84.38 |
| Two or more races | 44 | 34 | 77.27 | 34 | 77.27 |
| Intellectual disability | 1,972 | 1,728 | 87.63 | 1,709 | 86.66 |
| Hearing impairment | 29 | 22 | 75.86 | 20 | 68.97 |
| Speech or language impairment | 45 | 43 | 95.56 | 42 | 93.33 |
| Visual impairment | 13 | 9 | 69.23 | 9 | 69.23 |
| Emotional impairment | 11 | 8 | 72.73 | 8 | 72.73 |
| Orthopedic impairment | 135 | 82 | 60.74 | 80 | 59.26 |
| Other health impairment | 206 | 174 | 84.47 | 173 | 83.98 |
| Specific learning disability | 254 | 229 | 90.16 | 227 | 89.37 |
| Deaf-blindness | 4 | 4 | 100.00 | 4 | 100.00 |
| Multiple disabilities | 285 | 212 | 74.39 | 197 | 69.12 |
| Autism | 1,350 | 1,158 | 85.78 | 1,148 | 85.04 |
| Traumatic brain injury | 11 | 7 | 63.64 | 7 | 63.64 |
| Not classified | 0 | 0 | N/A | 0 | N/A |
| Not economically disadvantaged | 1,039 | 845 | 81.33 | 829 | 79.79 |
| Economically disadvantaged | 3,276 | 2,831 | 86.42 | 2,795 | 85.32 |
| In US schools less than 12 months | 39 | 28 | 71.79 | 28 | 71.79 |
| In US schools 12 months or more | 4,249 | 3,624 | 85.29 | 3,572 | 84.07 |
| Duration unknown | 27 | 24 | 88.89 | 24 | 88.89 |
| Migrant education | 67 | 59 | 88.06 | 59 | 88.06 |
| Not migrant education | 4,248 | 3,617 | 85.15 | 3,565 | 83.92 |

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

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Student Group | Number Registered | Number Tested | Percent Tested | Number Analyzed | Percent Analyzed |
| All | 6,405 | 5,121 | 79.95 | 5,032 | 78.56 |
| Male | 4,151 | 3,311 | 79.76 | 3,258 | 78.49 |
| Female | 2,252 | 1,809 | 80.33 | 1,773 | 78.73 |
| Nonbinary | 2 | 1 | 50.00 | 1 | 50.00 |
| American Indian or Alaska Native | 5 | 4 | 80.00 | 4 | 80.00 |
| Asian | 878 | 695 | 79.16 | 686 | 78.13 |
| Native Hawaiian or Other Pacific Islander | 28 | 23 | 82.14 | 23 | 82.14 |
| Filipino | 140 | 105 | 75.00 | 104 | 74.29 |
| Hispanic or Latino | 4,989 | 4,007 | 80.32 | 3,934 | 78.85 |
| Black or African American | 39 | 36 | 92.31 | 35 | 89.74 |
| White | 266 | 203 | 76.32 | 200 | 75.19 |
| Two or more races | 60 | 48 | 80.00 | 46 | 76.67 |
| Intellectual disability | 3,154 | 2,595 | 82.28 | 2,560 | 81.17 |
| Hearing impairment | 53 | 44 | 83.02 | 44 | 83.02 |
| Speech or language impairment | 34 | 30 | 88.24 | 30 | 88.24 |
| Visual impairment | 30 | 27 | 90.00 | 26 | 86.67 |
| Emotional impairment | 50 | 32 | 64.00 | 32 | 64.00 |
| Orthopedic impairment | 282 | 179 | 63.48 | 170 | 60.28 |
| Other health impairment | 209 | 162 | 77.51 | 161 | 77.03 |
| Specific learning disability | 274 | 217 | 79.20 | 216 | 78.83 |
| Deaf-blindness | 6 | 4 | 66.67 | 4 | 66.67 |
| Multiple disabilities | 473 | 344 | 72.73 | 319 | 67.44 |
| Autism | 1,796 | 1,454 | 80.96 | 1,439 | 80.12 |
| Traumatic brain injury | 38 | 30 | 78.95 | 29 | 76.32 |
| Not classified | 6 | 3 | 50.00 | 2 | 33.33 |
| Not economically disadvantaged | 1,791 | 1,370 | 76.49 | 1,340 | 74.82 |
| Economically disadvantaged | 4,614 | 3,751 | 81.30 | 3,692 | 80.02 |
| In US schools less than 12 months | 22 | 16 | 72.73 | 16 | 72.73 |
| In US schools 12 months or more | 6,338 | 5,070 | 79.99 | 4,983 | 78.62 |
| Duration unknown | 45 | 35 | 77.78 | 33 | 73.33 |
| Migrant education | 51 | 41 | 80.39 | 41 | 80.39 |
| Not migrant education | 6,354 | 5,080 | 79.95 | 4,991 | 78.55 |

## Psychometric Analyses

This chapter summarizes the item- and test-level statistics from the analyses conducted for the 2021–22 operational field test administration of the Alternate English Language Proficiency Assessments for California (ELPAC). It also describes the analysis procedures and results for the dimensionality study.

### Overview

This chapter describes the psychometric analyses conducted by ETS for the Alternate ELPAC, including classical item analyses, differential item functioning (DIF) analyses, item response theory (IRT) calibration, and response time analyses, as well as analyses to support reliability and validity evidence.

#### Summary of the Analyses

Each of these analyses of the 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 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](#_Appendix_8.A_(formerly). Summary statistics, including the mean, minimum, and maximum values are presented in table 8.3.
2. **DIF Analyses—**DIF analysis is described in section [*8.3 Differential Item Functioning Analyses*](#_Differential_Item_Functioning). Table 8.4 provides the student groups included in the analyses. Table 8.5 and table 8.6 provide the DIF categories for dichotomous and polytomous items, respectively. Table 8.B.1 through table 8.B.3, in [appendix 8.B](#_Appendix_7.B:_Differential), present the results of the DIF analysis for Alternate ELPAC operational field test items.
3. **IRT Analyses—**IRT models, and analyses, including calibrations and linking, are presented in section [*8.4 Item Response Theory Analyses*](#_Item_Response_Theory). Table 8.8 presents the slopes and intercepts that convert theta scores to reported scale scores for the operational field test items. Table 8.C.1 through table 8.C.6 in [appendix 8.C](#_Appendix_7.C:_Item) provide IRT results for individual items by grade level or grade span.
4. **Response Time Analyses—**Alternate ELPAC assessments are untimed, but test examiners need guidance on anticipated test duration as they schedule administrations. Response time analysis is described in section [*8.6 Response Time Analyses*](#_Response_Time_Analyses). Table 8.D.1 in [appendix 8.D](#_Appendix_7.D:_Response) provides summary statistics of response times for the Alternate ELPAC at the first, tenth, twenty-fifth, fiftieth, seventy-fifth, ninetieth, and ninety-ninth percentiles.
5. **Dimensionality Analyses—**Dimensionality analyses are described in section [*8.7 Dimensionality Study*](#_Dimensionality_Study)*.* [Appendix 8.E](#_Appendix_7.E:_Factor) presents the results of the dimensionality analyses.
6. **Reliability Analyses—**Reliability estimation for the Alternate ELPAC is illustrated in section [*8.8 Reliability Analyses*](#_Reliability__Analyses_1).
7. **Validity Evidence—**Validity evidence related to the Alternate ELPAC is discussed in section [*8.9 Validity Evidence*](#_Validity_Evidence).

#### Samples Used for the Analyses

In general, analyses included in the technical report are based on all students in the tested population with valid scores available at the time of analysis. The actual data sample used depends on the time that data source becomes available as well as when the sample size is adequate to meet the analysis timeline.

The classical item analyses and item-level DIF analyses were conducted twice: the preliminary item analysis was conducted using the data file available in mid-February 2022 (i.e., the preliminary data), and the final item analysis was conducted using the final data from June 2022. The classical item analyses and item level DIF analyses provided in this technical report ([appendix 8.A](#_Appendix_8.A:_Classical) and [appendix 8.B](#_Appendix_7.B:_Differential) respectively) are based on the final item analysis with the June data (i.e., final data). The IRT analyses ([appendix 8.C](#_Appendix_7.C:_Item)) and the dimensionality analyses ([appendix 8.E](#_Appendix_7.E:_Factor)) were based on the data file available in mid-February 2022 (i.e., analysis sample). All other analyses, such as the response time analyses, used the final version of the production data file for student reports, which became available in July 2022. 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 test early. A student may not reach the end of a test, or the item(s) was 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.

Table 8.1 presents the differences in student counts between the two data sources (i.e., the analysis sample and the final production data file). Final production data files were received later and contained a larger number of test completers than the analysis sample.

Table 8.1 Alternate ELPAC Analyses Data Sources

|  |  |  |
| --- | --- | --- |
| Grade Level or Grade Span | Analysis Sample | Final Data Total |
| Kindergarten | 727 | 1,246 |
| 1 | 833 | 1,243 |
| 2 | 773 | 1,134 |
| 3–5 | 2,896 | 4,167 |
| 6–8 | 2,530 | 3,566 |
| 9–12 | 3,502 | 4,925 |

#### Test-Taking Rates

The decision to assign a student to take the Alternate ELPAC is made by the student’s individualized education program (IEP) team using the information on the California Department of Education (CDE) Alternate Assessment IEP Team Guidance web page. This web page describes the Alternate ELPAC and its administration as well as criteria for test taking and the students who should be identified to take this test (CDE, 2021).

A student must meet all three of the following criteria to be identified to take the Alternate ELPAC:

1. **The student has a significant cognitive disability.** Review of the student’s school records indicates a disability or multiple disabilities that significantly impact intellectual functioning and adaptive behavior essential for a person to live independently and to function safely in daily life.
2. **The student is learning content derived from the California Common Core State Standards (CA CCSS) or the California Next Generation Science Standards (CA NGSS)** **or is acquiring proficiency as identified in the 2012 English Language Development Standards (ELD Standards).** Goals and instruction listed in the IEP for the student are linked to the grade-level CA CCSS, CA NGSS, or 2012 ELD Standards and address knowledge and skills that are appropriate and set high expectations for this student.
3. **The student needs extensive, direct individualized instruction and substantial supports to achieve measurable gains in the grade-level and age-appropriate curriculum, including the following:**

* Instruction and support that are not of a temporary or transient nature
* Substantially adapted materials and individualized methods of accessing information in alternative ways to acquire, maintain, generalize, demonstrate, and transfer skills across multiple settings

All students who are identified by an IEP team to take the 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 test and, therefore, have a valid score (CDE, 2020).

Table 8.2 shows the test-taking rates by grade level.

Table 8.2 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 | 1,431 | 1,336 | 1,297 | 93.4 |
| 1 | 1,401 | 1,301 | 1,266 | 92.9 |
| 2 | 1,346 | 1,238 | 1,219 | 92.0 |
| 3 | 1,664 | 1,498 | 1,475 | 90.0 |
| 4 | 1,589 | 1,409 | 1,388 | 88.7 |
| 5 | 1,546 | 1,393 | 1,368 | 90.1 |

Table 8.2 *(continuation)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level | Number of Registered Students | Number of Test Takers | Number of Students with Valid Score | Test Takers as a Percent of Registered Students |
| 6 | 1,464 | 1,271 | 1,252 | 86.8 |
| 7 | 1,497 | 1,276 | 1,258 | 85.2 |
| 8 | 1,354 | 1,129 | 1,114 | 83.4 |
| 9 | 1,078 | 919 | 902 | 85.3 |
| 10 | 1,085 | 859 | 849 | 79.2 |
| 11 | 1,165 | 910 | 896 | 78.1 |
| 12 | 3,077 | 2,433 | 2,385 | 79.1 |

### Classical Item Analyses

Classical item analyses are conducted to evaluate the performance of all test items with respect to item difficulty, item discrimination, and distractor analysis. The associated flagging rules of these statistics are used to identify items that are not performing as expected.

#### 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

*Ni* 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_7) *for a description of this equation.*

 (8.2)

where,

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

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

*Max (Xi)* 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 test.

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_37) *for a description of this equation.*

 (8.3)

where,

 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; and

is the variance of the criterion (for example, the students’ total score).

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-ability students tend to perform better on the item than lower-ability students. An item with a negative item-total correlation typically signifies a problem with the item, as that indicates that

* the higher-ability students on the overall test tend to respond incorrectly to the item if dichotomous, or are assigned a low score for the item if polytomous; or
* the lower-ability students on the overall test 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.

Items with a low percentage (i.e., less than 3 percent) of students obtaining any score point were flagged for review. Such items may pose problems during IRT calibration. They need to be carefully reviewed and may need to be excluded from the item calibration analyses.

#### 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.”

##### 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.

#### Distractor Analyses

Distractor analyses were conducted on selected-response (SR) items (i.e., items that were not constructed response). The statistics for each item included the proportion of students selecting each distractor (incorrect response), computed for the group of all students in the analysis sample, and were also computed separately for the highest-performing 20 percent of students. Items were flagged for review if more high-performing students chose any distractor rather than the key. Such a result indicated that the item may have multiple correct answers or have the wrong key (i.e., the item was miskeyed).

For SR items, the distractor-total correlation describes the relationship between selecting a distractor for a specific item and performance on the total test. The polyserial correlation was calculated for the distractors, like the item-total correlation previously described, except that the regressions were implemented on the distractors rather than the keys. Items with distractor-total correlations not significantly below zero were flagged for review, as these items may have multiple correct answers, be miskeyed, or have other content issues.

#### Summary of Classical Item Analysis Flagging Criteria

An item was flagged for review if the item analysis yielded any of the following results. One item could have multiple flags if the statistics met the flagging criteria:

* **Difficulty flags** indicated extreme values of the proportion-correct (for dichotomous items) or the proportion of the possible maximum points earned (for polytomous items):
* A-flag: A *p-*value below 0.50 for two-choice dichotomous single-select items, below 0.30 for three-choice dichotomous single-select items, or below 0.20 for all other items.
* H-flag: A *p*-value above 0.95 for dichotomous items or above 0.80 for polytomous items.
* A **discrimination flag** (R-flag) indicated that the item did not discriminate effectively between high- and low-ability students. Items with a polyserial correlation less than 0.20 were flagged.
* An omit flag (O-flag) indicated an omission rate above 10 percent for dichotomous multiple-choice, single-select items or above 15 percent for all other items.
* A **distractor flag** (P-flag) was used for an item with any distractors having a correlation with the criterion score that is either positive, zero, or negative but not significantly below zero
* A **miskey flag** (D-flag) was used for multiple-choice items when more of the high-ability examinee group—the top 20 percent of examinees on the total assessment—choose any distractor rather than the response keyed as correct.
* An **underrepresented score point flag** (L-flag) was used for any item that had less than 3 percent of the students at any score level.

ETS’ Psychometric Analysis & Research staff and Assessment and Learning Technology Research & Development staff carefully reviewed each of the flagged items during and at the end of the item analyses. All flagged items were also reviewed by California educators at the data review meeting and then summarized for the CDE with recommendations for subsequent analyses.

#### Classical Item Analyses Results

Table 8.3 presents the summary statistics of item difficulty and the item discrimination statistics for all operational field test 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 8.3 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 | 34 | 1,246 | 0.40 | 0.15 | 0.77 | 0.66 | 0.36 | 0.82 |
| 1 | 34 | 1,243 | 0.48 | 0.30 | 0.82 | 0.67 | 0.55 | 0.79 |
| 2 | 34 | 1,134 | 0.48 | 0.20 | 0.79 | 0.67 | 0.43 | 0.77 |
| 3–5 | 49 | 4,167 | 0.58 | 0.26 | 0.82 | 0.69 | 0.39 | 0.81 |
| 6–8 | 50 | 3,566 | 0.67 | 0.41 | 0.91 | 0.74 | 0.38 | 0.82 |
| 9–12 | 50 | 4,925 | 0.68 | 0.39 | 0.92 | 0.74 | 0.56 | 0.83 |

The item difficulty and item discrimination statistics and the associated flag using the criteria described in subsection [*8.2.6 Summary of Classical Item Analysis Flagging Criteria*](#_Summary_of_Classical_1) 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](#_Appendix_8.A:_Classical). Most of the items were flagged for high omit rate, especially for lower grade levels.

### Differential Item Functioning Analyses

DIF is used to evaluate the consistency of individual item performance for students in different demographic student groups who have the same level of domain performance. For example, DIF evaluates whether female and male students matched to have the same test score perform similarly on each item in the test.

In examining the DIF between groups, the reference group is often designated as the group that is assumed to have an advantage, while the focal group refers to the group anticipated to possibly be disadvantaged by the test.

DIF analyses were conducted for field test items that met the sample size requirements. The sample size requirements for the DIF analyses were 100 in the smaller of either the focal group or the reference group and 400 in the combined focal and reference groups. These sample size requirements are based on standard operating procedures with respect to DIF analyses at ETS.

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

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

DIF analyses were conducted on each test for designated comparison groups. Groups are defined on the basis of demographic variables, such as gender, race or ethnicity, and primary disabilities, if the number of students in the group meets the sample size requirements.

These comparison groups are specified in table 8.4.

Table 8.4 Student Groups for DIF Comparison

|  |  |  |
| --- | --- | --- |
| DIF Type | Reference Group | Focal Group |
| **Primary Disability** | Intellectual Disability | * Autism * Deaf-blindness * Emotional disturbance * Hearing impairment * Multiple disabilities * Orthopedic impairment * Other health impairment * Specific learning disability * Speech or language impairment * Traumatic brain injury * Visual impairment * Deafness |
| **Gender** | Male | * Female |
| **Race or Ethnicity** | Hispanic or Latino | * Non-Hispanic or non-Latino |

**Note:** DIF comparisons are not performed using the nonbinary group because sample sizes are insufficient to make meaningful comparisons.

#### Differential Item Functioning Procedure for Dichotomous Items

The Mantel-Haenszel (MH) DIF (MH-DIF) statistic was calculated for dichotomous items (Mantel & Haenszel, 1959; Holland & Thayer, 1985). For this method, students are classified to relevant student groups of interest (e.g., gender or ethnicity). Students at each total score level in the focal group (e.g., females) are compared with students at each total score level in the reference group (e.g., males). The common odds ratio—that is, the proportion of correct response over the proportion of incorrect response—is estimated across all levels of matched student ability using the formula in equation 8.4 (Dorans & Holland, 1993). The resulting estimate is interpreted as the relative probability of success on a particular item for members of two groups when matched on ability. *Refer to the* [*Alternative Text for Equation 8.4*](#_Alternative_Text_for_38) *for a description of this equation.*

Equation 8.4; a link to the long description for this equation is found in the preceding paragraph. (8.4)

where,

*m* indexes the score categories,

*Rrm* is the number of students in the reference group at score level *m* who answer the item correctly,

*Wfm* is the number of students in the focal group at score level *m* who answer the item incorrectly,

*Ntm* is the total number of students at score level *m*,

*Rfm* is the number of students in the focal group at score level *m* who answer the item correctly, and

*Wrm* is the number of students in the reference group at score level *m* who answer the item incorrectly.

To facilitate the interpretation of MH results, the common odds ratio is frequently transformed to the delta scale using equation 8.5 (Holland & Thayer, 1985). *Refer to the* [*Alternative Text for Equation 8.5*](#_Alternative_Text_for_39) *for a description of this equation.*

 (8.5)

Positive values indicate DIF in favor of the focal group (i.e., positive DIF items are differentially easier for the focal group), whereas negative values indicate DIF in favor of the reference group (i.e., negative DIF items are differentially easier for the reference group).

#### Differential Item Functioning Procedure for Polytomous Items

The standardization DIF (Dorans & Schmitt, 1993; Zwick, Thayer, & Mazzeo, 1997; Dorans, 2013) in conjunction with the Mantel chi-square statistic (Mantel, 1963; Mantel & Haenszel, 1959) is calculated for polytomous items. The standardized mean difference (SMD) compares the item means of the two groups after adjusting for differences in the distribution of students across all items and is calculated using equation 8.6. *Refer to the* [*Alternative Text for Equation 8.6*](#_Alternative_Text_for_40) *for a description of this equation.*

 (8.6)

where,

*X* isthe criterion score (total raw score),

*Y* isthe item score,

*M* is the number of score levels on *X*,

*Nfm* is the number of students in the focal group at score level *m*,

*Er* is the expected item score for the reference group,

*Ef* is the expected item score for the focal group, and

*Dm* is the difference in the distribution of students at score level *m*.

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

A positive SMDvalue means that, conditional on the criterion score, the focal group has a higher mean item score than the reference group (i.e., the item is differentially easier for the focal group). In contrast, a negative SMD value means that, conditional upon the criterion score, the focal group has a lower mean item score than the reference group (i.e., the item is differentially harder for the focal group).

#### Classification

Based on the DIF statistics and significance tests, items are classified into three categories and assigned values of A, B, or C (Holland & Wainer, 1993). Category A items contain negligible DIF, Category B items exhibit slight to moderate DIF, and Category C items possess moderate to large DIF values.

The flagging criteria for dichotomous items are presented in table 8.5; the flagging criteria for polytomous items are provided in table 8.6. The determination of all significant differences is based on *p*-value < 0.05.

Table 8.5 DIF Categories for Dichotomous Items

|  |  |
| --- | --- |
| DIF Category | Criteria |
| A (negligible) | * Absolute value of MH D-DIF is less than one or is not significantly different from zero. * Positive values are classified as “A+” and negative values as “A-.” |
| B (moderate) | * Absolute value of MH D-DIF is significantly different from zero but not from one and is at least one; *or* absolute value of MH D-DIF is significantly different from one but is less than 1.5. * Positive values are classified as “B+” and negative values as “B-.” |
| C (large) | * Absolute value of MH D-DIF is at least 1.5 and is significantly different from one. * Positive values are classified as “C+” and negative values as “C-.” |

Table 8.6 DIF Categories for Polytomous Items

|  |  |
| --- | --- |
| DIF Category | Criteria |
| A (negligible) | Mantel chi-square *p-*value≥ 0.05 or |SMD/SD| ≤ 0.17 |
| B (moderate) | Mantel chi-square *p-*value *<* 0.05 and 0.17 < |SMD/SD| ≤ 0.25 |
| C (large) | Mantel chi-square *p-*value *<* 0.05 and |SMD*/*SD| > 0.25 |

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

#### Differential Item Functioning Analysis Results Summary

Summarized DIF results for the operational field-test items are given in [appendix 8.B](#_Appendix_7.B:_Differential), table 8.B.1 through table 8.B.3, for each grade level and grade span. If the sample size requirement for conducting DIF analyses was not met, that item was categorized in “insufficient counts.”

No items, across grade level or grade span, were flagged for gender or ethnicity DIF. One item in grade span six through eight and three items in grade span nine through twelve were flagged for primary disability DIF. Items that show C-level DIF and are considered biased by the DIF review panel will be deactivated for future use. For the four items flagged for primary disability DIF, none of them were considered biased after DIF panel review.

Test developers are instructed to avoid selecting C-level items considered unbiased by the DIF review panel for future test forms unless their inclusion is deemed essential to meeting test-content specifications.

### 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 to calibrate all the operational field test items to establish the base IRT scale. All operational field test items that were not rejected by both the data review committees and the CDE were included in the calibration process.

#### 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 (PCM) (Masters, 1982), was applied to both dichotomous and polytomous items.

The mathematical form of the GPCM is presented in equation 8.7. *Refer to the* [*Alternative Text for Equation 8.7*](#_Alternative_Text_for_41) *for a description of this equation.*

 (8.7)

where,

*Pih(θj)* is the probability of student with proficiency *θj* obtaining score *h* on item *i*,

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

*ai* is the discrimination parameter and 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*, and

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

When *ni* = 1, equation 8.7 becomes an expression of the one-parameter logistic (1PL) model for dichotomous items.

#### Data Preparation

Prior to IRT calibration analyses, ETS psychometricians reviewed the results of the classical item analyses to decide whether any items were of poor quality and needed to be removed from calibration. The results also were reviewed by ETS content experts and the CDE. The decision whether to remove items from calibration was made in consultation with the CDE.

For IRT calibration, scored item response data was used to create the IRT analysis input data files for each grade level. The IRT analysis input data file was a full matrix containing item-level scores for students who answered at least one expressive item and one receptive item.

Similar to the classical item analyses, “omit” items were treated as incorrect and “not presented” items were treated as blank.

#### 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).

For all assessments, regardless of whether they are preequated or postequated, IRT calibration and linking were conducted to put the field test item parameters onto the base IRT scale.

The 2021–22 administration is the first operational field test administration for the Alternate ELPAC. The data from this administration was used to establish the baseline scale.

##### Calibration

After the 2021–22 Alternate ELPAC administration, all operational field test items within each test (grade level and content area) were calibrated using all available data.

FlexMIRT (Cai, 2017), a multilevel and multiple-group IRT software package for item analysis and test scoring, was used for item calibration analysis. This software can fit a variety of IRT models to both single-level and multilevel data that are dichotomous, polytomous, or both, and was chosen for its superior flexibility among IRT software programs.

The evaluation of the calibration results includes the following steps:

1. Reviewing the item parameter estimates to examine whether these estimates were reasonable
   1. At the form level, the summary statistics for the *b*-parameter estimates (location difficulty) and *d*-parameter estimates (step parameter) were examined, including the mean, standard deviation (SD), median, minimum, maximum, and goodness-of-fit.
   2. At the item level, statistics of individual items were examined, including item difficulty estimates, model-fit statistics, and the IRT-based item parameters.
2. Flagging items that did not perform as expected (All flagged items were discussed thoroughly with the CDE to decide whether those items should be removed from calibration or whether the scoring categories need to be collapsed.)

The calibration process was paralleled by two ETS psychometricians to ensure quality and accuracy of results. Specifically, two psychometricians independently created flexMIRT control files and ran the same input data files and then compared the calibration results. Any differences in the output were investigated. Refer to section [*9.6 Quality Control of Psychometric Processes*](#_Quality_Control_of_1) for more details of this procedure.

#### Parameter Estimates

The overall summary of the IRT *b*-parameter estimates for all items is shown in table 8.7. The summary statistics, such as the minimum, maximum, mean, and SD values are also presented.

The number of items in each of the *b*-parameter intervals is shown for each grade level or grade span. The means of *b*-parameter estimates ranged from -1.33 to 0.64. The mean item difficulty level decreased for higher grades. Overall, there were not many items with large *b-*parameters (i.e., very difficult items), but all *b*-parameters were within acceptable ranges.

Table 8.7 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 | 1 |
| -3.8 ≤ b < -3.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| -3.6 ≤ b < -3.4 | 0 | 0 | 0 | 0 | 2 | 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 | 2 | 1 |
| -2.6 ≤ b < -2.4 | 0 | 0 | 0 | 0 | 0 | 0 |
| -2.4 ≤ b < -2.2 | 0 | 0 | 0 | 0 | 2 | 1 |
| -2.2 ≤ b < -2.0 | 0 | 0 | 0 | 2 | 0 | 3 |
| -2.0 ≤ b < -1.8 | 0 | 1 | 0 | 2 | 4 | 4 |
| -1.8 ≤ b < -1.6 | 1 | 0 | 1 | 0 | 4 | 5 |
| -1.6 ≤ b < -1.4 | 0 | 1 | 0 | 2 | 2 | 5 |
| -1.4 ≤ b < -1.2 | 0 | 0 | 0 | 1 | 6 | 7 |
| -1.2 ≤ b < -1.0 | 0 | 0 | 2 | 3 | 7 | 4 |
| -1.0 ≤ b < -0.8 | 2 | 1 | 2 | 6 | 5 | 7 |
| -0.8 ≤ b < -0.6 | 0 | 5 | 1 | 5 | 3 | 1 |
| -0.6 ≤ b < -0.4 | 0 | 0 | 1 | 4 | 2 | 2 |
| -0.4 ≤ b < -0.2 | 2 | 3 | 3 | 7 | 4 | 1 |
| -0.2 ≤ b < 0.0 | 1 | 2 | 2 | 3 | 1 | 2 |
| 0.0 ≤ b < 0.2 | 3 | 3 | 7 | 2 | 3 | 3 |
| 0.2 ≤ b < 0.4 | 3 | 4 | 5 | 4 | 2 | 0 |
| 0.4 ≤ b < 0.6 | 0 | 7 | 0 | 4 | 0 | 0 |
| 0.6 ≤ b < 0.8 | 6 | 3 | 5 | 2 | 1 | 0 |
| 0.8 ≤ b < 1.0 | 4 | 2 | 3 | 1 | 0 | 1 |

Table 8.7 *(continuation one)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| IRT *b*-value | Kindergarten | Grade 1 | Grade 2 | Grade Span 3–5 | Grade Span 6–8 | Grade Span 9–12 |
| 1.0 ≤ b < 1.2 | 3 | 2 | 0 | 0 | 0 | 0 |
| 1.2 ≤ b < 1.4 | 5 | 0 | 0 | 0 | 0 | 0 |
| 1.4 ≤ b < 1.6 | 1 | 0 | 1 | 1 | 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 | -1.79 | -1.99 | -1.80 | -2.11 | -3.45 | -3.98 |
| Maximum | 2.51 | 1.18 | 1.76 | 1.43 | 0.63 | 0.90 |
| Mean | 0.64 | 0.06 | 0.11 | -0.44 | -1.14 | -1.33 |
| SD | 0.87 | 0.71 | 0.74 | 0.79 | 0.89 | 0.89 |
| **Number of Items** | **34** | **34** | **34** | **49** | **50** | **50** |

The detailed information on item difficulty parameter *b,* the standard error of *b*-parameters, as well as the step parameters *d* for the polytomous items for all operational field test items are presented in table 8.C.1 through table 8.C.6 in [appendix 8.C](#_Appendix_8.C:_Item).

The item–person maps by grade level or grade span are provided in table 8.C.7 through table 8.C.12 in [appendix 8.C](#_Appendix_7.C:_Item). The item–person maps plot the item difficulty distribution side by side with the ability distribution of students and provide information on the alignment of item difficulty with students’ abilities. The item–person map for grade span nine through twelve indicates the test can benefit from more difficult items to provide a better measure for the Level 3 threshold.

### 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 [*8.5.2 Transformation from Theta Scores to Scale Scores*](#_Transformation_from_Theta_1)*.*

The raw-to-scale-score conversion tables for each form used in each grade level or grade span are included in table 8.G.1 through table 8.G.11 in [appendix 8.G](#_Appendix_8.G:_Raw-to-Scale-Score).

#### 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 file that was described in subsection [*8.4.2 Data Preparation*](#_Data_Preparation_1)*,* through 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 8.8. *Refer to the* [*Alternative Text for Equation 8.8*](#_Alternative_Text_for_42) *for a description of this equation.*

 (8.8)

where,

*i* indexes dichotomous items,

*j* indexes polytomous items,

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

*Pi(θ)* is the probability of a correct response to item *i* at ability *θ* on the dichotomous item in equation 8.7,

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

*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 8.7, 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 8.9, 8.10, and 8.11). *Refer to the* [*Alternative Text for Equation 8.9*](#_Alternative_Text_for_43) *for a description of this equation.*

 (8.9)

where,

theta-hat represents student ability.

The slope and intercept 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, if the IRT calibration is used, the slope and intercept in equation 8.9 are derived using the threshold scores from standard setting approved by the SBE ( in equations 8.10 and 8.11) and the desired threshold scale scores (two-digit scale score) ( in equations 8.10 and 8.11) using the following two formulas:

*Refer to the* [*Alternative Text for Equation 8.10*](#_Alternative_Text_for_44) *for a description of this equation.*

[](#_Alternative_Text_for_10) (8.10)

*Refer to the* [*Alternative Text for Equation 8.11*](#_Alternative_Text_for_45) *for a description of this equation.*

 (8.11)

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, 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 and content area are shown in table 8.8. Also, refer to subsection [*7.1.3 Scale Scores for the Total Assessment*](#_Scale_Scores_for)for the special requirements for the Alternate ELPAC reporting scale.

Table 8.8 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 | 244 | 260 | 10.62 | 44.35 |
| 1 | -0.2819 | 1.1807 | 344 | 360 | 10.94 | 47.08 |
| 2 | -0.1351 | 1.2967 | 444 | 460 | 11.17 | 45.52 |
| 3ؘ–5 | -0.7329 | 1.0099 | 544 | 560 | 9.18 | 50.73 |

Table 8.8 *(continuation)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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 |
| 6–8 | -1.2416 | 0.5178 | 644 | 660 | 9.09 | 55.29 |
| 9–10 | -1.0927 | 0.9151 | 744 | 760 | 7.97 | 52.71 |
| 11–12 | -1.0927 | 0.9151 | 844 | 860 | 7.97 | 52.71 |

### Response Time Analyses

Response time analyses are conducted at the item level and the total test level. At the item level, timing information was collected by the delivery platform for each “page” (screen) that was presented to the students. Information about the time required to answer a single question is available only for items that appear on a page alone. The time required to answer all questions on a page is available when multiple items appear on a page. At the total test level, response times are calculated by summing the page durations for all items in the Alternate ELPAC.

Table 8.D.1 in [appendix 8.D](#_Appendix_7.D:_Response) provides summary statistics of response times for the Alternate ELPAC, at the first, tenth, twenty-fifth, fiftieth, seventy-fifth, ninetieth, and ninety-ninth percentiles. Total test response times calculated for the fiftieth and ninetieth percentiles provide local educational agency administrators with an indicator of how much time students required on average, as well as how much time might be needed for students who require more time.

Because the response time distribution is often very skewed, the mean response time could be impacted by the unusually long response time at the end of the distribution. The median response time (i.e., half of the students spent that amount of time or less to finish the assessment) is more meaningful than the mean response time. The median testing time ranged from 13.39 minutes for grade span eleven and twelve at performance Level 3 (scale score interval of 860–899) to 21.99 minutes for grade two at performance Level 2 (scale score interval of 444–459). The ninety-ninth percentile of testing time (i.e., 99 percent of students completed the test within this testing time) ranged from 46.76 minutes for grade span eleven and twelve at performance Level 3 (scale score interval of 860–899) to 113.75 minutes for grade one at performance Level 2 (scale score interval of 344–359).

### Dimensionality Study

The dimensionality study evaluated the underlying construct measured and investigated whether practical unidimensionality held for the Alternate ELPAC.

#### Overview

The way IRT models are applied to item performance to report Alternate ELPAC scores is dependent on the dimensionality of the data. If a single dimension, or factor, can describe the performance of all items making up an assessment, that assessment would be described as *unidimensional*. IRT models in such an assessment using a single factor, typically called ability, and is given the Greek symbol *θ* (theta)*.* However, when assessments consist of multiple distinct sections of content, separate unidimensional IRT models are often used for each of those sections. When content is less distinct, an evaluation of whether the performance of items reflects one or multiple dimensions is needed.

Adopting the term “practical unidimensionality” acknowledges that all models are imperfect, and true unidimensionality is an ideal concept that is hardly tenable in practice (Reise, Morizot, & Hays, 2007), and the assessment of practical unidimensionality is essentially an evaluation of whether the data is unidimensional enough to allow the use of unidimensional models. The results of this analysis informed how the assessment items were to be calibrated and how the scores were to be reported to best measure students’ English language proficiency (ELP) using the Alternate ELPAC. If the student response data from the assessment provided evidence of practical unidimensionality, all the items could be calibrated jointly, providing a total test score that could be reported on the basis of all items the student took. If the student response data from the assessment provided evidence of multidimensionality, the items from different dimensions could be calibrated separately, allowing the scores to be combined to provide a weighted composite score from these dimensions.

Confirmatory factor analysis (CFA) was used to evaluate whether the data fits a prespecified model. For the Alternate ELPAC, the hypothesized structure to be evaluated was the communication mode, receptive and expressive, which are defined as follows:

1. Receptive items are those that require a student to demonstrate their comprehension of a stimulus by selecting a response from two or three options; the student is not required to generate any language.
2. Expressive test items are those that require students to communicate to others their understandings and ideas related to the stimulus using their individually preferred expressive mode of communication.

If the Alternate ELPAC was practically unidimensional, the unidimensional IRT model would be used to calibrate all items. However, if there were evidence of multidimensionality for items that measure receptive and expressive modes, they would be calibrated separately and separate scale scores would be produced, resulting in a weighted average scale score for the overall test score.

#### Methods

##### Model Specification

In CFA, two different methods within the IRT framework were used to evaluate the test dimensionality. The first method used the correlated common factor model, with some items measuring receptive as one factor and the rest measuring expressive as a second factor. The correlation between these two factors was estimated and the factor loading matrix was evaluated. A second method used a bifactor model and compared its results against a single-factor (i.e., unidimensional) model*.* Multiple indices were used to evaluate the relative strength of the general factor in the bifactor against the strength of the specific factors in the bifactor model and against the factor in the single-factor model. Figure 8.1, figure 8.2, and figure 8.3 illustrate the three models discussed in this section. In all three figures, the same items are modeled, represented by N, although only items 1 through 4 and items N-3 through N are shown, to save space.

The path diagram for the correlated two-factor model is shown in figure 8.1. In this model, each item has two sources that contribute to the item variance: the factor on which the item loads and the item’s error term. The two correlated factors are labeled “Receptive” and “Expressive,” representing the interpretation of the two factors. The error terms are labeled E1 through EN, where the subscript is consistent with the number of items in the model, from 1 to N. The items themselves are represented by rectangles labeled Item 1 to Item N.

Each of the two factors has one-sided arrows pointing from it towards all of the items that load on it. There are also one-sided arrows pointing from each item’s error term towards the item it belongs to. For a given item, its error term accounts for all the variance in the observed scores that is not explainable by the factor on which it loads.

The double-sided curved arrow in figure 8.1 represents the correlation between the receptive factor and the expressive factor, which is freely estimated in this model. The higher this correlation, the more evidence the model provides that the test is unidimensional; the lower the correlation, the more evidence the model provides that the test has two dimensions.

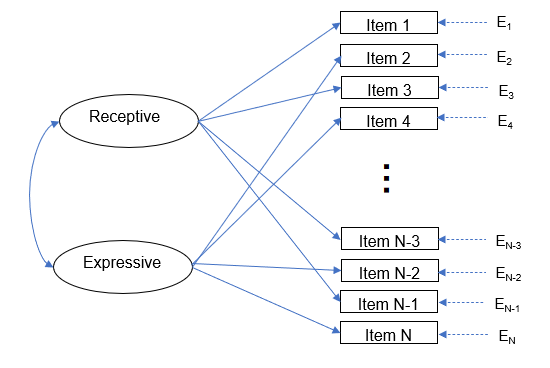


Figure 8.1 The path diagram for the correlated two-factor model

The second method used a bifactor model, which was evaluated relative to a single-factor model. The path diagrams for the bifactor model and the single-factor model are shown in figure 8.2 and figure 8.3. In figure 8.2, the specific factors are labeled “Receptive” and “Expressive” to represent the interpretation of the two factors. There is no double-sided curved arrow between these two ovals as there was in figure 8.1. The general factor is represented as an oval labeled “ELP” that represents the interpretation of that factor. The items and error terms are represented exactly as in figure 8.1, and all arrows between items and their specific factors and between items and their error terms are the same as in figure 8.1. Additionally, there are one-sided arrows pointing from the general factor to every item.

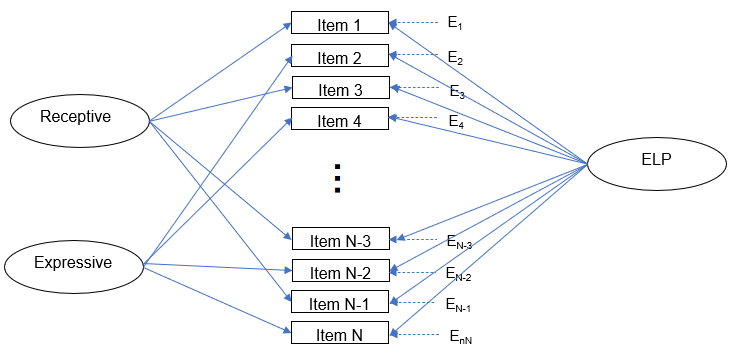


Figure 8.2 The path diagram for a bifactor model

In figure 8.3, the items and error terms are again represented as in figure 8.1 and figure 8.2. There is only one factor in this model, labeled with “ELP” to represent the interpretation of this factor, and there are one-sided arrows pointing from this factor leading to each of the items in the model. As with the two correlated-factors model shown in figure 8.1, each item has its own error term in addition to the one specific factor and one general factor on which it loads (figure 8.2) or in addition to the single factor on which all items load (figure 8.3).

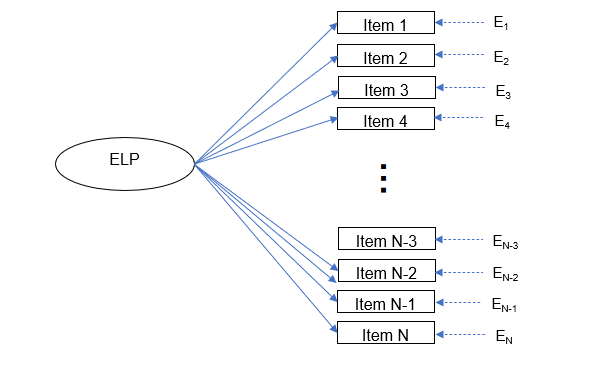


Figure 8.3 The path diagram for a one-factor model

In figure 8.2 for the bifactor model, each item has three sources that contribute to the item variance: the general factor (ELP), the content-specific factor (receptive or expressive) that is not explained by the general factor, and the error term. The correlations between the communication modes are uncorrelated in this model because their correlations have been reflected by the general factor.

In figure 8.3, the hypothesis is that the items assess only one latent variable, interpreted as general ELP, and that specific factors, interpreted as receptive and expressive ELP, cannot account for a meaningful amount of variance in the item responses beyond this. In this model, all items measure ELP, and each item has two sources that contribute to the item variance: the general factor (ELP) and the error term.

In the IRT framework, the probability of obtaining a response pattern on *K* items is denoted in equation 8.12. *Refer to the* [*Alternative Text for Equation 8.12*](#_Alternative_Text_for_20) *for a description of this equation.*

 (8.12)

where,

**** is the probability for a student with ability vector *θ* to get a score point *y* on item *k*.

This probability  is related to a linear function of the latent variables *θ* through a link function *g(π)*, typically a probit or logit function. The difference in the bifactor model and the correlated factor multidimensional item response theory (MIRT) model is in the link function. The link function for a bifactor model is defined in equation 8.13. *Refer to the* [*Alternative Text for Equation 8.13*](#_Alternative_Text_for_23) *for a description of this equation.*

 (8.13)

where,

*ck* is the intercept parameter for item *k*,

*akg* is the loading on the general factor, and

*akm* is the loading on the dimension *m*.

The latent variables *θg* and *θm* are specified as uncorrelated; the link function for the correlated factor MIRT model is defined in equation 8.14. *Refer to the* [*Alternative Text for Equation 8.14*](#_Alternative_Text_for_24) *for a description of this equation.*

 (8.14)

where the instances of *θm* are correlated.

The bifactor model and the correlated factor MIRT model provide different parameter estimates but should provide similar information on the test dimensionality. Strong unidimensionality will be reflected in the MIRT model with high correlations among all dimensions. While in the bifactor model, high correlations among the dimensions will be absorbed by the general factor, and the items will have insignificant loadings on the specific factors. The bifactor model has an added benefit of showing the proportion of variance that can be explained by the specific factor as a direct way to evaluate the test dimensionality.

All models used in this study were fitted with the commercial *FlexMIRT* 3.5 software (Cai, 2017).

##### Model Evaluation Criteria

A variety of evidence was considered when determining whether the Alternate ELPAC follows a practical unidimensional structure.

Item loadings on the general factor and the content-specific factors were examined from the bifactor model. High loadings on the general factor and low loadings on the content-specific factors for most of the items suggested that the unidimensional model may be sufficient for the data. It should be noted in interpreting the results of content-specific factors that these are not loadings on the named factor, but, rather, low loadings on an uninterpretable factor that is orthogonal to the general factor. That is, these content-specific factors are residuals after the general factor is removed.

Rodriguez, Reise, and Haviland (2016) proposed to use a group of statistical indices computed on the basis of factor loadings of the bifactor model to evaluate the strength of the general factor and hence indirectly assess practical unidimensionality. Specifically, the explained common variance (ECV) and the relative parameter bias were used to evaluate the strength of the general factor and to determine whether the test can be considered as practically unidimensional.

The ECV measures the extent to which data are essential unidimensional (Ten Berge & Socan, 2004). High ECV values suggest high similarity between factor loadings of the one-factor model and those on the general factor of the bifactor model, which also supports practical unidimensionality. Equation 8.15 shows how the ECV index can be computed. *Refer to the* [*Alternative Text for Equation 8.15*](#_Alternative_Text_for_25) *for a description of this equation.*

 (8.15)

An item parameter estimate could be biased when a test with multidimensional structure is forced to conform to a unidimensional model. That is, a unidimensional model is fit to data with a multidimensional structure. Relative parameter bias (RPB) reflects the amount of bias in the item loadings when the test with a bifactor structure is fit using a unidimensional model and is defined in equation 8.16. *Refer to the* [*Alternative Text for Equation 8.16*](#_Alternative_Text_for_26) *for a description of this equation.*

Equation 8.16; a link to the long description for this equation is found in the preceding paragraph. (8.16)

where,

gamma sub g hat is the item loading on the general factor in the bifactor model, and

gamma sub u hat is the item loading in the unidimensional model.

The one-factor model was fit, in part, to provide the information needed by RPB.

The mean and variance of theta in the unidimensional model were constrained to be (0,1), and the same constraints were applied to the factors in the bifactor model. This ensured that the bifactor general factor and the single dimension from the unidimensional analysis were on the same metric.

##### Analysis Sample

The dimensionality analysis was conducted after preliminary item analysis (PIA) and DIF were completed by using data received on February 16, 2022. All valid responses for the Alternate ELPAC were used for this analysis. It should be noted that the data inclusion rules described in [subsection 8.1.2](#_Samples_for_the) were applied to the dimensionality study.

The dimensionality analysis was conducted separately for each grade level (kindergarten, grade one, and grade two) or grade span (grade spans three through five, six through eight, and nine through twelve). For kindergarten, grade one, and grade two, there was one form with 34 items used for each grade level. For grade spans three through five, six through eight, and nine through twelve, each grade span had two forms, each with 34 items, including 18 common items between the two forms. There were 50 unique items across two forms, and all 50 items were included in the analyses.

Table 8.9 provides the Alternate ELPAC operational field test form structures.

Table 8.9 Item Distribution of the Alternate ELPAC Operational Field Test by Communication Mode

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test | Expressive—Number of Items | Expressive—Possible Max Score | Receptive—Number of Items | Receptive—Possible Max Score | Total Number of Items | Total Possible Max Score |
| Kindergarten | 15 | 23 | 19 | 19 | 34 | 42 |
| 1 | 15 | 25 | 19 | 19 | 34 | 44 |
| 2 | 15 | 24 | 19 | 19 | 34 | 43 |
| 3–5 Form 1 | 15 | 24 | 19 | 19 | 34 | 43 |
| 3–5 Form 2 | 15 | 24 | 19 | 19 | 34 | 43 |
| 6–8 Form 1 | 15 | 23 | 19 | 19 | 34 | 42 |
| 6–8 Form 2 | 15 | 23 | 19 | 19 | 34 | 42 |
| 9–12 Form 1 | 15 | 26 | 19 | 19 | 34 | 45 |
| 9–12 Form 2 | 15 | 26 | 19 | 19 | 34 | 45 |

#### Results

One discernible pattern in the datasets for the Alternate ELPAC is that factor loadings on the general factor of the bifactor models are greater than those on the secondary factors after the general factor has been accounted for. This indicates the existence of a strong general factor in the Alternate ELPAC operational field test data.

Table 8.E.1 through Table 8.E.6 in [appendix 8.E](#_Appendix_8.E:_Factor) present the standardized factor loadings of both the bifactor and the one-factor models fit to the Alternate ELPAC six datasets. In these tables, “M1” refers to the one-factor model and “M2” refers to the bifactor model. As the general factor is the main contributor to the total score variance, all items have a relatively high loading on the general factor, while most items have a very small loading on secondary factors. A few items have high loadings on both the general and the secondary factor (e.g., VR154753 and VR133983 in table 8.E.2). Possible causes of the high loadings on the secondary factor are that these items have low discriminating power or that they measure something different than all other items in general.

Figure 8.E.1 through figure 8.E.6 plot the factor loadings from the single-factor model and the general factor from the bifactor model. In most of the scatterplots, the points cluster around the regression line closely, indicating that most of the differences between the two sets of factor loading values are negligibly small. In the scatterplot for kindergarten (Figure 8.E.1), the points cluster around the regression line with some distance between the two sets of factor loadings, and some points are appreciably away from the regression line, suggesting that, for some of the items in those tests, these two sets of factor loadings are not as similar as those observed in the other scatterplots.

To further evaluate whether those tests are practically unidimensional, the bifactor model–based statistical indices ECV and RPB discussed earlier are presented in table 8.10.

Table 8.10 Dimensionality Evaluation

|  |  |  |
| --- | --- | --- |
| Test | ECV | RPB |
| Kindergarten | 0.87 | 0.07 |
| Grade 1 | 0.87 | 0.04 |
| Grade 2 | 0.86 | 0.04 |
| Grades 3–5 | 0.90 | 0.04 |
| Grades 6–8 | 0.92 | 0.01 |
| Grades 9–12 | 0.91 | 0.03 |

The ECV index shows the percent of common variance explained by the general factor. It is a direct way to measure the strength of the general factor in comparison to the secondary receptive and expressive communication mode factors after the general factor is accounted for. The lowest ECV is 0.86 for grade two, which means that 86 percent of the common variance was explained by the general factor and 14 percent by the secondary factors.

The largest absolute RPB is 0.07 for kindergarten, indicating that the data has a multidimensional structure but, when fitted with a unidimensional model, the average bias in the factor loadings is small and negligible. Muthén, Kaplan, and Hollis (1987) suggested that parameter bias less than 10–15 percent is acceptable and poses no serious concern.

Table 8.11 shows the correlations of the expressive and receptive factors when the data was fitted with the correlated two-factor MIRT model. The high correlations between two communication mode scores support the unidimensionality of the test.

Table 8.11 Correlations between the Communication Mode Scores

|  |  |
| --- | --- |
| Grade Level or Grade Span | Correlation between Receptive and Expressive Scores |
| Kindergarten | 0.90 |
| Grade 1 | 0.90 |
| Grade 2 | 0.87 |
| Grades 3–5 | 0.94 |
| Grades 6–8 | 0.97 |
| Grades 9–12 | 0.95 |

Overall, the findings suggest that the Alternate ELPAC is practically unidimensional, which supports overall scale score reporting, interpretation, use, and reporting and interpretation of measurement error and reliability coefficient based on the unidimensional IRT model. Further, as the assumption of practical unidimensionality in the Alternate ELPAC across all grade levels and grade spans is tenable, the use of unidimensional IRT models as the psychometric machinery for the test is psychometrically defensible.

### Reliability Analyses

The reliability for a particular group of students’ test scores estimates the extent to which the scores would remain consistent if those same students were retested with a parallel version of the same test. 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 test.

#### Reliability Measures

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

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

The Alternate ELPAC test reliabilities were evaluated for each domain, the composite scores, or for item types as dichotomous items and polytomous items, using the coefficient alpha (Cronbach, 1951) index of internal consistency, which is calculated as presented in equation 8.17.

*Refer to the* [*Alternative Text for Equation 8.17*](#_Appendix_7.A:_Classical) *for a description of this equation.*

 (8.17)

where,

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

is the estimated variance of item *i*, and

is the estimated total test variance.

The reliability of the overall score was estimated by substituting sample estimates into the following definitional formula for composite reliability (Feldt & Brennan, 1989), equation 8.18. *Refer to the* [*Alternative Text for Equation 8.18*](#_Alternative_Text_for_46) *for a description of this equation.*

 (8.18)

where,

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

 is the variance of scores on the *j*th component,

 is the reliability of scores on the *j*th component, and

 is the variance of the composite score.

#### 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 test. If it were possible to compute the error of measurement for each student’s score in a large group of students, these errors of measurement would have a mean of zero. These SEMs would be 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. In a large group of students, approximately two-thirds of the students will earn scores within one SEM of the scores they would earn on a perfectly reliable test.

The SEM is the square root of the error variance in the scores (i.e., the SD of the distribution of the differences between students’ observed scores and their true scores). The SEM is calculated using equation 8.19. *Refer to the* [*Alternative Text for Equation 8.19*](#_Alternative_Text_for_47) *for a description of this equation.*

 (8.19)

where,

rho sub theta theta prime is the reliability estimated in equation 8.17,

*Sθ* is the SD of the total test *θ* score, and

*a* is the slope of the scaling transformation of *θ* to 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 test. Across those administrations, approximately 95 percent of the time the student’s observed score -1.96 SEMs to the student’s observed score +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 test equals 250 points and the SEM equals 5, one can be 95 percent confident that the student’s true score lies between 240 and 260 points (250  10).

Table 8.12 gives the mean and SD of the scale scores, the reliability, and the SEM for each test form. These results indicate that the reliability estimates for all tests are moderately high.

Table 8.12 Test Reliability of the Total Scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | N Students | Mean | SD | Reliability | SEM |
| Kindergarten | 1,166 | 246 | 14.5 | 0.87 | 5.28 |
| Grade 1 | 1,163 | 348 | 13.9 | 0.86 | 5.17 |
| Grade 2 | 1,131 | 446 | 15.5 | 0.87 | 5.50 |
| Grades 3–5 | 3,936 | 551 | 13.0 | 0.88 | 4.59 |
| Grades 6–8 | 3,070 | 653 | 14.4 | 0.86 | 5.30 |
| Grades 9–10 | 1,514 | 750 | 12.5 | 0.87 | 4.51 |
| Grades 11–12 | 2,761 | 850 | 13.0 | 0.87 | 4.62 |

#### Student Group Reliabilities

The reliabilities were examined for various student groups that tested. The student groups included in these analyses were defined by their gender, ethnicity, primary disabilities, economic status, length of attendance in US schools, and migrant education status. SEM information for the total test scores by test form are reported for each student group.

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.F.1 through table 8.F.7 in [appendix 8.F](#_Appendix_8.F:_Reliability) present the overall test reliabilities for the various student groups. Most student groups in all seven grade levels or grade spans have reliability greater than 0.85. Note that students associated with a specific learning disability group have low reliability in a few grade levels or grade spans, with ranges from 0.29 in grade span six through eight to 0.47 in grade span nine through ten. The low reliability for this student group may likely be due to a lack of variation in student performance.

#### 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 (TIF) based on the items taken by each student. The CSEM for a student with proficiency 𝜃𝑗 is calculated using equation 8.20. *Refer to the* [*Alternative Text for Equation 8.20*](#_Alternative_Text_for_48) *for a description of this equation.*

 (8.20)

where,

*I(θj)* is the test information for student *j* and is calculated using equation 8.21. *Refer to the* [*Alternative Text for Equation 8.21*](#_Alternative_Text_for_49) *for a description of this equation.*

 (8.21)

where,

*Ii(θj)* is the item information of item *i* for student *j*.

Item information is calculated as presented in equation 8.22. *Refer to the* [*Alternative Text for Equation 8.22*](#_Alternative_Text_for_50) *for a description of this equation.*

 (8.22)

where,

*si(θj)* is the expected item score for item *i* on a theta score *θj* calculated as presented in equation 8.23 *(Refer to the* [*Alternative Text for Equation 8.23*](#_Alternative_Text_for_51) *for a description of this equation.*)

 (8.23)

and equation 8.24. *Refer to the* [*Alternative Text for Equation 8.24*](#_Alternative_Text_for_52) *for a description of this equation.*

 (8.24)

where,

*pih(θj)* is the probability of an examinee with *θj* getting score *h* on item *i*, the computation of which is shown in equation 8.7; and

*ni* is the maximum number of score points for item *i*.

CSEMs for scale scores are computed by transforming CSEMs of theta scores onto the reporting scale. Refer to subsection [*8.5.2* *Transformation from Theta Scores to Scale Scores*](#_Transformation_from_Theta_1)for scaling factors of transformation. A student’s CSEM for scale scores under the IRT framework is equal to the reciprocal of the square root of the TIF multiplied by the scaling factor *a*, as presented in equation 8.25. *Refer to the* [*Alternative Text for Equation 8.25*](#_Alternative_Text_for_54) *for a description of this equation.*

 (8.25)

where,

;

CSEM(*SS*) is the conditional SEM on scale score scale;

 is the TIF at ability level  as shown in equations 8.22, 8.23, and 8.24; and

*a* is the scaling factor (the slope) needed to transform theta to the scale score metric.

Table 8.13 presents the scale score CSEMs at the lowest score required for a student to be classified in the score reporting range two and score reporting range three for Alternate ELPAC tests.

Table 8.13 Scale Score CSEMs at the Reporting Score Range Thresholds

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Level 2‌―Reporting Score Range Threshold | Level 2‌—CSEM | Level 3‌―Reporting Score Range Threshold | Level 3‌—CSEM |
| Kindergarten | 244 | 4 | 260 | 4 |
| Grade 1 | 344 | 4 | 360 | 5 |
| Grade 2 | 444 | 4 | 460 | 5 |
| Grade span 3–5 Form 1 | 544 | 3 | 560 | 4 |
| Grade span 3–5 Form 2 | 544 | 3 | 560 | 4 |
| Grade span 6–8 Form 1 | 644 | 3 | 660 | 5 |
| Grade span 6–8 Form 2 | 644 | 4 | 660 | 4 |
| Grade span 9–10 Form 1 | 744 | 3 | 760 | 5 |
| Grade span 9–10 Form 2 | 744 | 3 | 760 | 5 |
| Grade span 11–12 Form 1 | 844 | 3 | 860 | 5 |
| Grade span 11–12 Form 2 | 844 | 3 | 860 | 5 |

The CSEMs for all score points are included in the raw-to-scale-score conversion tables for each form used in each grade level or grade span in table 8.G.1 through table 8.G.11 in [appendix 8.G](#_Appendix_7.G_=).

#### 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 test (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 test?

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 8.14 and table 8.15 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 8.14 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 8.15 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 proficiency level for the Alternate ELPAC are presented in table 8.F.8 in [appendix 8.F](#_Appendix_7.F_=) for all grade levels and grade spans.

At each threshold, the classification at adjacent performance levels appeared to be acceptably reliable and consistent. The classification accuracy ranged from 0.87 to 0.90 for the cut between Level 1 and Level 2, and from 0.88 to 0.91 for the cut between Level 2 and Level 3. The classification consistency ranged from 0.83 to 0.87 for the cut between Level 1 and Level 2, and 0.85 to 0.88 for the cut between Level 2 and Level 3. Overall, the accuracy ranged from 0.77 to 0.79 and classification consistency ranged from 0.70 to 0.72.

#### Interrater Agreement

The interrater reliability analyses are performed on approximately 10 percent of the overall testing population, randomly selected from the total population; those students’ responses are scored by two raters. The two sets of ratings are used to compute statistics describing the consistency (or reliability) of the ratings. This interrater consistency is described in three ways:

1. Percentage agreement between two raters
2. Cohen’s Kappa
3. Quadratic-weighted kappa (QWK) coefficient

In some scoring rubrics, zero is a valid score for the responses but is not provided by a rater. Instead, a score of zero is assigned when the student attempted the writing task but did not provide a response. Responses with zero scores should not be included in the calculation of the agreement statistics for these items.

##### Percentage Agreement

Percentage agreement between two raters is frequently defined as the percentage of exact score agreement and adjacent score agreement. Exact score agreement means two raters give exact same scores. Adjacent score agreement means agreement between scores that differ by just one point. The percentage of exact score agreement is a stringent criterion, which tends to decrease with an increasing number of item score points. The fewer the item score points, the fewer degrees of freedom on which two raters can vary, and the higher the percentage of agreement.

##### Kappa

Interrater reliability or consistency is an indicator of homogeneity and is most frequently measured using Cohen’s Kappa statistic (1960), which takes chance agreement. For a human-scored item with *m+1* categories, one can construct an (*m+1)* × *(m+1)* rating table with scores provided by two raters, *X* and *Y*, as shown in table 8.16. Let n sub s t denote the number of responses for which rater *X’s* score = *s* and rater *Y’s* score = *t,* n sub s plus is the number of responses for which rater *X’s* score = *s*, n sub plus t is the number of responses for which rater *Y’s* score = *t*, and n sub plus plus is the number of all responses.

Table 8.16 Frequencies of Ratings

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Rating | Y = 0 | Y = 1 | Y = 2 | …\*\* | Y = m\* |
| X = 0 | n00 | n01 | n02 | … | n0m |
| X = 1 | N10 | n11 | n12 | … | n1m |
| X = 2 | n20 | n21 | n22 | … | n2m |
| … | … | … | … | … | … |
| X = m | nm0 | nm1 | nm2 | … | nmm |

\* m is the number of score categories of an item.

\*\* An ellipsis (…) signifies that there might be more rows (or columns) in the table.

*Refer to the* [*Alternative Text for Equation 8.26*](#_Alternative_Text_for_53) *for a description of this equation.* The kappa statistic is defined as

 (8.26)

*Refer to the* [*Alternative Text for Equation 8.27*](#_Alternative_Text_for_55) *for a description of this equation.*

 (8.27)

*Refer to the* [*Alternative Text for Equation 8.28*](#_Alternative_Text_for_56) *for a description of this equation.*

 (8.28)

where,

p sub obs is the observed agreement, and

p sub exp is the expected agreement between *X* and *Y*.

When p sub obs and p sub exp agree only at the chance level, the value of kappa is 0. When the two measurements agree perfectly, the value of kappa is 1.0.

##### Quadratic-Weighted Kappa

QWK is used because kappa does not take into account the degree of disagreement between raters. It is a generalization of the simple kappa coefficient using weights to quantify the relative difference between categories. The range of the QWK is from 0.0 to 1.0, with perfect agreement being equal to 1.0.

For a human-scored item with *m+1* categories, one can construct an (*m+1)* × *(m+1)* rating table with scores provided by two raters, *X* and *Y,* as described in table 8.16. The weighted kappa coefficient is defined as presented in equation 8.29. *Refer to the* [*Alternative Text for Equation 8.29*](#_Alternative_Text_for_57) *for a description of this equation.*

 (8.29)

For QWK, the weights are calculated using equation 8.30. *Refer to the* [*Alternative Text for Equation 8.30*](#_Alternative_Text_for_58) *for a description of this equation.*

 (8.30)

##### Results of Interrater Reliability

The interrater reliability statistics for each item in each grade level or grade span are shown in table 8.H.1 through table 8.H.6 in [appendix 8.H](#_Appendix_8.H:_Interrater). All rubric-scored items are two-point items. These tables show that 37 out of 67 items have percent of exact agreement at least 90 percent. All other items have percent exact agreement no lower than 85 percent. The QWK ranged from 0.71 to 0.95 for all items across all grade levels or grade spans. This indicates the scoring has moderate to high interrater reliabilities.

### 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). It constitutes the central notion underlying the development, administration, and scoring of a test and the uses and interpretations of test scores.

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 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 test, 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 Alternate ELPAC, followed by a description and discussion of different kinds of validity evidence that have been gathered.

#### Design of the Alternate ELPAC

##### Purpose

The Alternate ELPAC is designed to assess students with the most significant cognitive disabilities whose individualized education program (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. 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.

##### The Constructs to Be Measured

The 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. 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 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 California State Board of Education–approved ELD Connectors and test blueprints (CDE, 2019).

##### Interpretations and Uses of the Scores

Overall student performance is expressed as a scale score that is generated for the 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.

An LEA may use Alternate ELPAC results to help make decisions about student placement in programs that support the student’s ongoing development toward English language proficiency, student exit from EL programs, and student growth in proficiency while in EL programs. The 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 Alternate ELPAC results.

##### Intended Test Population

The Alternate ELPAC is intended only for EL students, and potential EL students, in kindergarten through grade twelve (up to age twenty-one) who have been identified as having the most significant cognitive disabilities and who have been found eligible for alternate assessments by their IEP team.

#### 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 tests 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 2021–22 Alternate ELPAC administration and includes a description 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 Alternate ELPAC was conducting cognitive labs to understand how students interact with the test and to ensure that the test was 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.

##### Teacher Ratings of Student Performance

[Chapter 10](#_In-Test_Survey_1) describes the in-test survey that accompanied the Alternate ELPAC. This survey provided test administrators with an opportunity to provide ratings of students’ English language proficiency based on their in-class performance. Teachers’ ratings of students as EL (performance levels 1 and 2) or as fluent English proficient (performance level 3) across grades was 73.3 percent. The percent agreement differed by grade level or grade span and ranged from 68.3 to 80.5 (ETS, 2023). The in-survey test will continue to be administered and evidence will continue to be collected in this area.

##### Testing Time Analysis

Testing time for each administration can be evaluated for consistency by examining the expected response processes for the items presented to students. The length of time it took students to complete a test was collected and analyzed to build a profile describing what a typical testing event looked like for expressive and receptive items separately at each grade level. In addition, variability in testing time was investigated to determine whether a student’s testing time should be viewed as unusual or irregular. It should be noted that the Alternate ELPAC assessments are untimed tests.

Students with no item responses and students who did not answer at least one expressive item and one receptive item were removed from these analyses. The remaining testing population is partitioned into three score intervals based on the performance level thresholds. Descriptive statistics of the time required to complete the total test are computed for each of the three groups. Refer to section [*8.6 Response Time Analyses*](#_Response_Time_Analyses) for more information and results.

Some cases of extremely long testing time may be attributed to students with special needs taking longer to complete the tests or the test not being closed down properly. With that being said, the results should be interpreted with caution. The medians (fiftieth percentile) are more meaningful in the interpretation of the time comparisons because medians are less impacted by extreme values than means.

#### Relationship to Other Variables

Evidence based on *relations to other variables* can be evaluated using the correlation evidence between the Alternate ELPAC and the California Alternate Assessments (CAAs) for English language arts/literacy (ELA), mathematics, and science, as they are provided for students with the most significant cognitive disabilities and the majority of students who take the Alternate ELPAC take the CAAs. This type of evidence is essential for supporting the validity of certain inferences based on scores from the Alternate ELPAC.

##### Relationship Between Alternate ELPAC and California Alternate Assessment for English Language Arts/Literacy Test Scores

Table 8.17 presents the total number of students who took the Alternate ELPAC and the matched and unmatched percentages. “Matched Percentage” represents the percentage of students who also took the CAA for ELA. *Not Matched Percentage* includes the percentage of students who took the Alternate ELPAC but did not take the CAA for ELA. Kindergarten through grade two and grades nine, ten, and twelve are not included in this table because the CAA for ELA is not administered in these grade levels.

Table 8.17 also presents the correlations of Alternate ELPAC and CAA for ELA scores. Correlations between .67 and .73 are observed; these values suggest that student scores are similarly ordered for the two assessments, which is consistent with their being measures of related skills and supports the assertion that the Alternate ELPAC is an appropriate test for students with the most significant cognitive disabilities.

Table 8.17 Correlation of Alternate ELPAC and CAA for ELA Scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level | Number of Students Taking Alt ELPAC | Matched Percentage | Not Matched Percentage | Number of Students Taking Both | Correlation |
| 3 | 1,475 | 92.61 | 7.39 | 1,366 | 0.70 |
| 4 | 1,388 | 92.44 | 7.56 | 1,283 | 0.73 |
| 5 | 1,368 | 92.18 | 7.82 | 1,261 | 0.72 |
| 6 | 1,252 | 92.41 | 7.59 | 1,157 | 0.71 |
| 7 | 1,258 | 93.64 | 6.36 | 1,178 | 0.71 |
| 8 | 1,114 | 92.82 | 7.18 | 1,034 | 0.69 |
| 11 | 896 | 88.50 | 11.50 | 793 | 0.67 |

##### Relationship Between Alternate ELPAC and California Alternate Assessment for Mathematics Test Scores

Table 8.18 presents the total number of students who took the Alternate ELPAC and the matched and unmatched percentages. “Matched Percentage” represents the percentage of students who also took the CAA for Mathematics. *Not Matched Percentage* includes the percentage of students who took the Alternate ELPAC but did not take the CAA for Mathematics. Kindergarten through grade two and grades nine, ten, and twelve are not included in this table because the CAA for Mathematics is not administered in these grade levels.

Table 8.18 also presents the correlations of Alternate ELPAC and CAA for Mathematics scores. Correlations between .59 and .68 are observed; these values suggest that the Alternate ELPAC is an appropriate assessment for students with the most significant cognitive disabilities.

Table 8.18 Correlation of Alternate ELPAC and CAA for Mathematics Scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level | Number of Students Taking Alt ELPAC | Matched Percentage | Not Matched Percentage | Number of Students Taking Both | Correlation |
| 3 | 1,475 | 92.34 | 7.66 | 1,362 | 0.64 |
| 4 | 1,388 | 92.58 | 7.42 | 1,285 | 0.67 |
| 5 | 1,368 | 92.25 | 7.75 | 1,262 | 0.63 |
| 6 | 1,252 | 92.25 | 7.75 | 1,155 | 0.59 |
| 7 | 1,258 | 93.32 | 6.68 | 1,174 | 0.65 |
| 8 | 1,114 | 92.37 | 7.63 | 1,029 | 0.68 |
| 11 | 896 | 88.06 | 11.94 | 789 | 0.66 |

##### Relationship Between Alternate ELPAC and California Alternate Assessment for Science Test Scores

Table 8.19 presents the total number of students who took the Alternate ELPAC and the matched and unmatched percentages. “Matched Percentage” represents the percentage of students who also took the CAA for Science. *Not Matched Percentage* includes the percentage of students who took the Alternate ELPAC but did not take the CAA for Science. Grade five, grade eight, and grades ten through twelve are shown in this table because these are the grade levels at which the CAA for Science is administered. The higher proportion of matched students at grades five and eight is because students only take the CAA for Science once during high school, so the potential high school sample is diluted across three different grade levels.

Table 8.19 presents the correlations of Alternate ELPAC and CAA for Science scores. Correlations between .58 and .77 are observed; these values suggest that the Alternate ELPAC is an appropriate assessment for students with the most significant cognitive disabilities.

Table 8.19 Correlation of Alternate ELPAC and CAA for Science Scores

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Grade Level | Number of Students Taking Alt ELPAC | Matched Percentage | Not Matched Percentage | Number of Students Taking Both | Correlation |
| 5 | 1,368 | 89.33 | 10.67 | 1,222 | 0.77 |
| 8 | 1,114 | 86.80 | 13.20 | 967 | 0.75 |
| 10 | 849 | 3.06 | 96.94 | 26 | 0.58 |
| 11 | 896 | 49.55 | 50.45 | 444 | 0.71 |
| 12 | 2,385 | 24.11 | 75.89 | 575 | 0.71 |

#### Internal Structure

Internal structure evidence evaluates the strength or salience of the major dimensions underlying an assessment using dimensionality evaluation, which includes DIF analyses. These analyses were conducted using the 2021–22 operational field test data.

##### Test Dimensionality

Results of the test dimensionality study are summarized in [section 8.7](#_Dimensionality_Study) and [appendix 8.E](#_Appendix_7.E:_Factor) of this report.

Evidence collected from the 2021–22 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.

##### Differential Item Functioning

As described in section [*8.3 Differential Item Functioning Analysis*](#_Differential_Item_Functioning)*,* DIF analyses were conducted to assess differences in the item performance of groups of students who differ in their demographic characteristics. No items were identified as having significant levels of DIF for the gender and race or ethnicity groups. There were three items identified as having C-level DIF for the primary disability types, but all of these items were confirmed to be unbiased by a DIF review panel (refer to subsection [*8.3.4 Differential Item Functional Analysis Results Summary*](#_Differential_Item_Functioning_1) for additional information).

##### Overall Reliability Estimates

The results of reliability and standard error of measurements on the scale score for each test are presented in subsection [*8.8.2 Standard Error of Measurement*](#_Standard_Error_of_2). Results indicate that the scores for the Alternate ELPAC tests are reliable.

##### Student Group Reliability Estimates

The reliabilities are also examined for various student groups. The student groups considered are based on gender, ethnicity, economic status, primary disability, migrant status, and English language fluency. These results are presented in subsection [*8.8.3 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 [*8.8.5 Decision Classification Analyses*](#_Decision_Classification_Analyses).

#### English-Only Student Sample Validation Analysis

The purpose of the Alternate ELPAC operational field test administration was to provide a standardized measure of ELP for EL students, and potential EL students, with the most significant cognitive disabilities.

A challenge in developing the Alternate ELPAC was to identify items that clearly evaluated ELP for this group of students. If these students did not perform well on the items, it is important to try to understand whether this low performance was due to their EL status or if it was due to construct-irrelevant factors such as content appropriateness or a student’s method of communication.

##### Overview

One way to help determine whether this lack of performance was due to EL status was to compare the performance of EL test-takers to that of English only (EO) students with the most significant cognitive disabilities. The Alternate ELPAC is a new assessment, and the extent to which its items measure ELP for these EL and potential EL students is unknown. Therefore, a study including results from EO students was planned to identify items that are difficult for EL students but not difficult for EO students when those students are matched in terms of primary disability type. Such items would be assumed to measure English language skill rather than some other factor.

It was expected that the results of this study would provide useful information for future item development for EL students and potential EL students with the most significant cognitive disabilities.

##### Analysis

Despite the efforts made to recruit EO students in the study, there were very few participants in the EO study in the 2021–22 operational field test administration. Table 8.20 presents the number of targeted sample size for this study and the number of students who participated by grade level and grade span.

Table 8.20 Target Sample Size and the Number of Participants

|  |  |  |
| --- | --- | --- |
| Grade Level and Grade Span | Target Number of EO Students | Number of EO Participants |
| Kindergarten | 250 | 25 |
| Grade 1 | 250 | 20 |
| Grade 2 | 250 | 20 |
| Grade span 3–5 | 500 | 81 |
| Grade span 6–8 | 500 | 74 |
| Grade span 9–12 | 500 | 77 |
| **Total:** | **2,250** | **297** |

Further details on the sampling and recruitment plan can be found in the EO recruitment memorandum (ETS, 2020).

The planned analyses were not conducted because of the insufficient number of EO students tested in the spring 2022 operational field test administration.

### 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.

Cai, L. (2017). FlexMIRT®: *Flexible multilevel, multidimensional item analysis and test scoring* (Version 3.51) [computer software]. Chapel Hill, NC: Vector Psychometric Group.

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. (2021). *Alternate assessment IEP team guidance.* California Department of Education website.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20 (1): 37–46.

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. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

Dorans, N. J., & Schmitt, A. P. (1993). Constructed response and differential item functioning: A pragmatic approach. In R. E. Bennett & W. C. Ward (Eds.), *Construction versus choice in cognitive measurement* (pp. 135–65). Hillsdale, NH: Lawrence Erlbaum Associates, Inc.

Embretson (Whitley), S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin,* *93*, 179–97.

ETS. (2020). *EO recruitment memorandum.* [Unpublished memorandum]. Princeton, NJ: ETS.

ETS. (2023). *Summative Alternate English Language Proficiency Assessments for California threshold score validation study report* [Unpublished report.] Princeton, NJ: ETS.

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.

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.

Holland, P. W., & Thayer, D. T. (1985). *An alternative definition of the ETS delta scale of item difficulty* (Research Report 85–43). Princeton, NJ: Educational Testing Service.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.

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.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, *58*, 690–700.

Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analyses of data from retrospective studies of disease. *Journal of the National Cancer Institute,* *22*, 719–48.

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.

Muthén, B., Kaplan, D. & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika,* *52,* 431–62.

Olsson, U., Drasgow, F., & Dorans, N. J. (1982). The polyserial correlation coefficient. *Psychometrika, 47*, 337–47.

Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, *16*(1), 19‍–‍31.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–50.

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.

Ten Berge, J. M., & Socan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*(4), 613–25.

Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing differential item functioning in polytomous items. *Applied Measurement in Education, 10*(4), 321–44.

### Accessibility Information

#### Alternative Text for Equation 8.1

P value sub dich equals the fraction with the numerator the sum of X sub ij and the denominator N sub I end fraction.

#### Alternative Text for Equation 8.2

P value sub poly equals the fraction with the numerator X sub ij and the denominator N sub i times Max of X sub I end fraction.

#### Alternative Text for Equation 8.3

r sub polyreg equals the fraction beta sub hat times S tot divided by the square root of Beta sub hat squared times s sub tot squared plus 1.

#### Alternative Text for Equation 8.4

Alpha sub MH equals the numerator open parenthesis the sum sub m of R sub rm times W sub fm divided by N sub tm close parenthesis divided by the denominator open parenthesis the sum sub m of R sub fm times W sub rm divided by N sub tm closed parenthesis.

#### Alternative Text for Equation 8.5

MH D-DIF equals negative 2.35 times the natural logarithm open bracket alpha sub MH close bracket.

#### Alternative Text for Equation 8.6

SMD equals the fraction with numerator the sum from m equals 1 to M of N sub fm times E sub f of Y from X equals m and denominator the sum from m equals 1 to M of N sub fm end fraction minus the fraction with numerator the sum from m equals 1 to M of N sub fm times E sub r of Y from X equals m and denominator the sum from m equals 1 to M of N sub fm end fraction equals the fraction with the numerator the sum from m equals 1 to M of D sub fm and the denominator m equals1 to M of N sub fm end fraction.

#### Alternative Text for Equation 8.7

P sub ih of theta sub j equals the numerator exp open parenthesis the sum from v equals 1 to h of Da sub i open parenthesis theta 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 n sub I exp open parenthesis the sum from v equals 1 to c of Da 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 1, 2, …, n sub i.

P sub ih of theta sub j equals 1 divided by the denominator open parenthesis 1 plus the sum from c equals 1 to n sub I exp open parenthesis the sum from v equals 1 to c of Da 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.

#### Alternative Text for Equation 8.8

Epsilon of theta equals the sum from i equals 1 to ndich of P sub i of theta plus the sum from j equals 1 to npoly over each sum of x equals 1 to m of s sub xj times P sub xj of theta.

#### Alternative Text for Equation 8.9

Scale score equals intercept plus slope times theta-hat.

#### Alternative Text for Equation 8.10

Slope 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.

#### Alternative Text for Equation 8.11

Intercept 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.

#### Alternative Text for Equation 8.12

The probability for a student of ability vector theta to have a response vector y equals to the product from item 1 to item K of the probability of the student obtaining response yi for item i.

#### Alternative Text for Equation 8.13

The link function g of pi equals a sub kg times theta sub g plus a sub km times theta sub m plus c sub k.

#### Alternative Text for Equation 8.14

The link function g of pi equals a sub km times theta sub m plus c sub k.

#### Alternative Text for Equation 8.15

ECV equals the fraction with the numerator sum of lambda sub gen squared and the denominator sum of lambda sub gen squared plus sum over k sum lambda sub grpk squared.

#### Alternative Text for Equation 8.16

RPB equals the sum of the left bracket gamma hat sub g minus gamma hat sub u right bracket divided by n.

#### Alternative Text for Equation 8.17

Alpha hat equals the numerator k divided by the denominator k minus 1 multiplied by open bracket 1 minus the numerator summation from i equals 1 to k multiplied by sigma-hat squared sub i divided by the denominator sigma-hat squared sub x close bracket.

#### Alternative Text for Equation 8.18

Alpha hat sub c equals 1 minus fraction with numerator sum of j of w squared sub j times Sigma squared hat sub j times open parenthesis 1 minus alpha hat sub j close parenthesis and denominator Sigma squared hat sub c.

#### Alternative Text for Equation 8.19

SEM sub scaled equals a times s sub theta times the square root of 1 minus rho sub theta prime.

#### Alternative Text for Equation 8.20

SEM of Theta sub j equals 1 divided by the square root of I of theta sub j.

#### Alternative Text for Equation 8.21

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

#### Alternative Text for Equation 8.22

I sub i of theta sub j equals open bracket s sub i2 of theta sub j minus s sub i squared of theta sub j.

#### Alternative Text for Equation 8.23

S sub i of Theta sub j equals the sum from h equals 0 to n sub i of h times p sub ih of theta sub j.

#### Alternative Text for Equation 8.24

S sub i2 of Theta sub j equals the sum from h equals 0 to n sub i of h squared times p sub ih of theta sub j.

#### Alternative Text for Equation 8.25

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

#### Alternative Text for Equation 8.26

kappa equals the fraction with the numerator p sub obs minus p sub exp the denominator 1 minus p sub exp.

#### Alternative Text for Equation 8.27

P sub obs equals 1 divided by n times the sum from s equals 0 to m n sub ss.

#### Alternative Text for Equation 8.28

P sub exp equals 1 divided by n square times the sum from s equals 0 to m n sub s plus times n sub plus s.

#### Alternative Text for Equation 8.29

K sub ij equals open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub ij divided by n sub plus plus close parenthesis minus open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub I plus times n sub plus j divided by n squared sub plus plus close parenthesis divided open parenthesis 1 minus open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub i plus times n sub plus j divided by n squared sub plus plus close parenthesis close parenthesis, K sub ij equals open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub ij divided by n sub plus plus close parenthesis minus open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub i plus times n sub plus j divided by n squared sub plus plus close parenthesis divided open parenthesis 1 minus open parenthesis the sum from i equals zero to m the sum from j equals zero to m of w sub ij times n sub i plus times n sub plus j divided by n squared sub plus plus close parenthesis close parenthesis.

#### Alternative Text for Equation 8.30

W sub ij equals 1 minus open parenthesis I minus j close parenthesis squared divided by m squared.

### Appendix 8.A: Classical Item Analyses Results

Notes: The following abbreviations are used in table 8.A.1 through table 8.A.6:

Item Types:

* D = dichotomous item,
* P = polytomous item

Statistics:

* AIS = average item score

Flags:

* A = difficult item,
* H = easy item,
* Rpt = low item-total correlation,
* P = positive distractor-total correlation,
* O = high omit rate, and
* D = more high-scoring students chose a distractor.

Table 8.A. 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 | Flag |
| VR052871 | D | 1,246 | 0.27 | 0.57 | 25 | 48 | 27 | N/A | A O |
| VR053826 | D | 1,246 | 0.77 | 0.68 | 12 | 11 | 77 | N/A | O |
| VR131310 | D | 1,246 | 0.15 | 0.36 | 23 | 62 | 15 | N/A | A D O |
| VR131342 | D | 1,246 | 0.41 | 0.65 | 24 | 35 | 41 | N/A | O |
| VR131450 | D | 1,246 | 0.41 | 0.63 | 25 | 34 | 41 | N/A | A O |
| VR131594 | D | 1,246 | 0.23 | 0.44 | 25 | 52 | 23 | N/A | A O |
| VR131679 | D | 1,246 | 0.29 | 0.56 | 24 | 47 | 29 | N/A | A O |
| VR131683 | D | 1,246 | 0.37 | 0.60 | 23 | 39 | 37 | N/A | O |
| VR131687 | D | 1,246 | 0.29 | 0.57 | 26 | 45 | 29 | N/A | A O |
| VR137085 | D | 1,246 | 0.67 | 0.66 | 16 | 17 | 67 | N/A | O |
| VR138924 | D | 1,246 | 0.50 | 0.59 | 19 | 31 | 50 | N/A | A O |
| VR138950 | D | 1,246 | 0.20 | 0.50 | 30 | 50 | 20 | N/A | A O |
| VR138982 | D | 1,246 | 0.55 | 0.70 | 20 | 24 | 55 | N/A | O |
| VR139042 | D | 1,246 | 0.39 | 0.66 | 32 | 29 | 39 | N/A | O |
| VR139666 | D | 1,246 | 0.44 | 0.69 | 25 | 31 | 44 | N/A | O |
| VR139673 | D | 1,246 | 0.45 | 0.65 | 24 | 31 | 45 | N/A | O |

Table 8.A.1 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR139729 | D | 1,245 | 0.32 | 0.52 | 27 | 41 | 32 | N/A | A O |
| VR154406 | D | 1,246 | 0.50 | 0.79 | 30 | 20 | 50 | N/A | O |
| VR154449 | D | 1,246 | 0.36 | 0.53 | 24 | 40 | 36 | N/A | O |
| VR170322 | D | 1,246 | 0.37 | 0.68 | 25 | 38 | 37 | N/A | O |
| VR193093 | D | 1,246 | 0.64 | 0.67 | 14 | 21 | 64 | N/A | O |
| VR221674 | D | 1,246 | 0.35 | 0.57 | 27 | 39 | 35 | N/A | O |
| VR244385 | D | 1,246 | 0.32 | 0.57 | 27 | 41 | 32 | N/A | O |
| VR131419 | P | 1,246 | 0.59 | 0.77 | 17 | 46 | 14 | 22 | O |
| VR131711 | P | 1,246 | 0.61 | 0.79 | 17 | 45 | 15 | 23 | O |
| VR139022 | P | 1,246 | 0.76 | 0.77 | 13 | 42 | 13 | 32 | [no flag] |
| VR139973 | P | 1,246 | 0.98 | 0.79 | 9 | 36 | 12 | 43 | [no flag] |
| VR139975 | P | 1,246 | 0.55 | 0.74 | 18 | 47 | 15 | 20 | O |
| VR154458 | P | 1,246 | 0.57 | 0.79 | 19 | 44 | 17 | 20 | O |
| VR154465 | P | 1,246 | 1.15 | 0.82 | 16 | 20 | 13 | 51 | O |
| VR193113 | P | 1,246 | 1.02 | 0.74 | 8 | 32 | 19 | 41 | [no flag] |
| VR215978 | P | 1,246 | 0.71 | 0.79 | 18 | 39 | 16 | 28 | O |
| VR216450 | P | 1,246 | 1.15 | 0.75 | 6 | 30 | 12 | 52 | [no flag] |
| VR223164 | P | 1,246 | 0.48 | 0.73 | 19 | 49 | 16 | 16 | O |

Table 8.A. 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 | Flag |
| VR053944 | D | 1,243 | 0.82 | 0.69 | 8 | 10 | 82 | N/A | [no flag] |
| VR130319 | D | 1,243 | 0.57 | 0.69 | 15 | 28 | 57 | N/A | O |
| VR130326 | D | 1,243 | 0.41 | 0.60 | 16 | 43 | 41 | N/A | O |
| VR130331 | D | 1,243 | 0.49 | 0.58 | 17 | 34 | 49 | N/A | O |
| VR130345 | D | 1,243 | 0.49 | 0.65 | 17 | 35 | 49 | N/A | A O |
| VR130374 | D | 1,243 | 0.45 | 0.58 | 15 | 40 | 45 | N/A | O |
| VR130402 | D | 1,243 | 0.53 | 0.59 | 15 | 32 | 53 | N/A | O |
| VR130421 | D | 1,243 | 0.39 | 0.60 | 16 | 45 | 39 | N/A | O |
| VR133917 | D | 1,243 | 0.50 | 0.72 | 17 | 32 | 50 | N/A | O |
| VR133975 | D | 1,243 | 0.30 | 0.61 | 18 | 52 | 30 | N/A | A O |
| VR134007 | D | 1,243 | 0.61 | 0.74 | 16 | 22 | 61 | N/A | O |
| VR137615 | D | 1,243 | 0.63 | 0.66 | 14 | 23 | 63 | N/A | O |
| VR138495 | D | 1,243 | 0.55 | 0.56 | 16 | 30 | 55 | N/A | O |
| VR138505 | D | 1,243 | 0.62 | 0.68 | 14 | 24 | 62 | N/A | O |
| VR138567 | D | 1,243 | 0.66 | 0.69 | 14 | 19 | 66 | N/A | O |
| VR138601 | D | 1,243 | 0.41 | 0.57 | 15 | 44 | 41 | N/A | A O |
| VR138628 | D | 1,243 | 0.50 | 0.67 | 25 | 26 | 50 | N/A | O |
| VR150660 | D | 1,243 | 0.45 | 0.65 | 19 | 36 | 45 | N/A | O |
| VR150666 | D | 1,243 | 0.54 | 0.69 | 17 | 29 | 54 | N/A | O |
| VR150685 | D | 1,243 | 0.32 | 0.55 | 18 | 50 | 32 | N/A | O |
| VR150709 | D | 1,243 | 0.36 | 0.66 | 26 | 38 | 36 | N/A | O |
| VR154742 | D | 1,243 | 0.41 | 0.60 | 17 | 42 | 41 | N/A | O |
| VR154751 | D | 1,243 | 0.43 | 0.68 | 27 | 31 | 43 | N/A | O |
| VR193603 | D | 1,243 | 0.75 | 0.67 | 9 | 16 | 75 | N/A | [no flag] |

Table 8.A.2 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR053970 | P | 1,243 | 1.31 | 0.79 | 5 | 24 | 12 | 59 | [no flag] |
| VR130343 | P | 1,243 | 0.64 | 0.75 | 11 | 47 | 19 | 22 | [no flag] |
| VR130428 | P | 1,243 | 0.70 | 0.74 | 12 | 44 | 17 | 26 | [no flag] |
| VR133983 | P | 1,243 | 0.69 | 0.72 | 13 | 44 | 17 | 26 | [no flag] |
| VR134003 | P | 1,243 | 0.81 | 0.77 | 12 | 39 | 16 | 32 | [no flag] |
| VR137618 | P | 1,243 | 0.85 | 0.68 | 8 | 44 | 13 | 36 | [no flag] |
| VR150707 | P | 1,243 | 0.75 | 0.77 | 14 | 41 | 15 | 30 | [no flag] |
| VR154753 | P | 1,243 | 0.63 | 0.70 | 13 | 47 | 18 | 22 | [no flag] |
| VR154755 | P | 1,243 | 0.91 | 0.70 | 12 | 25 | 35 | 28 | [no flag] |
| VR193651 | P | 1,243 | 1.33 | 0.79 | 5 | 23 | 10 | 61 | [no flag] |

Table 8.A. 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 | Flag |
| VR130125 | D | 1,134 | 0.44 | 0.67 | 15 | 42 | 44 | N/A | O |
| VR130141 | D | 1,134 | 0.47 | 0.65 | 14 | 38 | 47 | N/A | O |
| VR130233 | D | 1,134 | 0.52 | 0.63 | 14 | 33 | 52 | N/A | O |
| VR130277 | D | 1,134 | 0.53 | 0.74 | 19 | 28 | 53 | N/A | O |
| VR130352 | D | 1,134 | 0.44 | 0.65 | 14 | 42 | 44 | N/A | O |
| VR130367 | D | 1,134 | 0.48 | 0.73 | 14 | 38 | 48 | N/A | O |
| VR130401 | D | 1,134 | 0.36 | 0.67 | 13 | 50 | 36 | N/A | O |
| VR130438 | D | 1,134 | 0.47 | 0.69 | 14 | 39 | 47 | N/A | O |
| VR134637 | D | 1,134 | 0.50 | 0.62 | 16 | 34 | 50 | N/A | O |
| VR134649 | D | 1,134 | 0.51 | 0.74 | 21 | 28 | 51 | N/A | O |
| VR134668 | D | 1,134 | 0.36 | 0.69 | 23 | 41 | 36 | N/A | O |
| VR140204 | D | 1,134 | 0.65 | 0.64 | 10 | 25 | 65 | N/A | [no flag] |
| VR140495 | D | 1,134 | 0.57 | 0.71 | 12 | 31 | 57 | N/A | O |
| VR140498 | D | 1,134 | 0.43 | 0.67 | 18 | 39 | 43 | N/A | O |
| VR140501 | D | 1,134 | 0.70 | 0.73 | 10 | 20 | 70 | N/A | O |
| VR140520 | D | 1,134 | 0.58 | 0.68 | 17 | 25 | 58 | N/A | O |
| VR140523 | D | 1,134 | 0.35 | 0.60 | 20 | 45 | 35 | N/A | O |
| VR151565 | D | 1,134 | 0.38 | 0.56 | 19 | 43 | 38 | N/A | O |
| VR151573 | D | 1,134 | 0.42 | 0.69 | 16 | 41 | 42 | N/A | O |
| VR151587 | D | 1,134 | 0.25 | 0.43 | 23 | 52 | 25 | N/A | A D O |
| VR151643 | D | 1,134 | 0.20 | 0.44 | 28 | 52 | 20 | N/A | A O |
| VR155513 | D | 1,134 | 0.49 | 0.72 | 14 | 37 | 49 | N/A | O |
| VR155674 | D | 1,134 | 0.35 | 0.64 | 16 | 49 | 35 | N/A | O |
| VR193816 | D | 1,134 | 0.79 | 0.73 | 7 | 14 | 79 | N/A | [no flag] |
| VR193873 | D | 1,134 | 0.69 | 0.64 | 6 | 25 | 69 | N/A | [no flag] |
| VR130298 | P | 1,134 | 0.74 | 0.71 | 12 | 40 | 22 | 26 | [no flag] |
| VR134664 | P | 1,134 | 0.61 | 0.73 | 13 | 46 | 19 | 21 | [no flag] |
| VR134677 | P | 1,134 | 0.87 | 0.77 | 13 | 36 | 16 | 36 | [no flag] |
| VR140209 | P | 1,134 | 1.30 | 0.76 | 6 | 24 | 9 | 60 | [no flag] |

Table 8.A.3 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR151624 | P | 1,134 | 0.69 | 0.74 | 15 | 42 | 16 | 26 | [no flag] |
| VR155670 | P | 1,134 | 0.92 | 0.74 | 13 | 22 | 39 | 27 | [no flag] |
| VR193828 | P | 1,134 | 1.11 | 0.66 | 6 | 29 | 19 | 46 | [no flag] |
| VR193885 | P | 1,134 | 1.39 | 0.77 | 6 | 20 | 10 | 64 | [no flag] |
| VR223063 | P | 1,134 | 0.97 | 0.73 | 12 | 21 | 36 | 31 | [no flag] |

Table 8.A. 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 | Flag |
| VR053180 | D | 2,054 | 0.71 | 0.68 | 7 | 22 | 71 | N/A | [no flag] |
| VR053988 | D | 2,113 | 0.82 | 0.68 | 5 | 13 | 82 | N/A | [no flag] |
| VR054453 | D | 2,054 | 0.61 | 0.74 | 10 | 29 | 61 | N/A | [no flag] |
| VR054493 | D | 2,054 | 0.56 | 0.74 | 10 | 34 | 56 | N/A | [no flag] |
| VR060298 | D | 2,113 | 0.58 | 0.65 | 9 | 34 | 58 | N/A | [no flag] |
| VR130469 | D | 2,113 | 0.65 | 0.75 | 8 | 27 | 65 | N/A | [no flag] |
| VR130481 | D | 2,113 | 0.64 | 0.70 | 9 | 27 | 64 | N/A | [no flag] |
| VR130500 | D | 2,113 | 0.39 | 0.70 | 12 | 49 | 39 | N/A | O |
| VR130526 | D | 2,113 | 0.51 | 0.70 | 12 | 37 | 51 | N/A | O |
| VR131591 | D | 2,113 | 0.55 | 0.69 | 8 | 36 | 55 | N/A | [no flag] |
| VR131622 | D | 2,113 | 0.56 | 0.60 | 8 | 36 | 56 | N/A | [no flag] |
| VR131627 | D | 2,113 | 0.42 | 0.54 | 12 | 46 | 42 | N/A | O |
| VR131628 | D | 2,113 | 0.66 | 0.75 | 12 | 22 | 66 | N/A | O |
| VR140200 | D | 2,113 | 0.66 | 0.80 | 13 | 21 | 66 | N/A | O |
| VR140214 | D | 2,113 | 0.56 | 0.76 | 9 | 35 | 56 | N/A | [no flag] |
| VR140221 | D | 2,113 | 0.59 | 0.77 | 12 | 29 | 59 | N/A | O |
| VR144415 | D | 2,054 | 0.82 | 0.72 | 6 | 12 | 82 | N/A | [no flag] |
| VR144676 | D | 2,113 | 0.75 | 0.66 | 6 | 19 | 75 | N/A | [no flag] |
| VR145701 | D | 2,113 | 0.80 | 0.73 | 7 | 13 | 80 | N/A | [no flag] |
| VR145817 | D | 2,113 | 0.64 | 0.67 | 7 | 29 | 64 | N/A | [no flag] |
| VR145916 | D | 2,113 | 0.65 | 0.63 | 7 | 27 | 65 | N/A | [no flag] |
| VR146024 | D | 2,113 | 0.69 | 0.69 | 10 | 21 | 69 | N/A | O |
| VR146075 | D | 2,113 | 0.45 | 0.65 | 11 | 44 | 45 | N/A | O |
| VR150967 | D | 2,113 | 0.50 | 0.58 | 10 | 39 | 50 | N/A | O |
| VR150970 | D | 2,113 | 0.57 | 0.73 | 12 | 32 | 57 | N/A | O |
| VR151004 | D | 2,113 | 0.40 | 0.51 | 15 | 45 | 40 | N/A | O |
| VR151042 | D | 2,054 | 0.45 | 0.43 | 10 | 45 | 45 | N/A | O |
| VR151060 | D | 2,054 | 0.38 | 0.66 | 11 | 50 | 38 | N/A | O |
| VR151083 | D | 2,054 | 0.26 | 0.39 | 14 | 60 | 26 | N/A | A O |
| VR151087 | D | 2,054 | 0.46 | 0.44 | 13 | 41 | 46 | N/A | O |
| VR155150 | D | 2,054 | 0.53 | 0.75 | 12 | 35 | 53 | N/A | O |

Table 8.A.4 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR155166 | D | 2,054 | 0.61 | 0.68 | 11 | 27 | 61 | N/A | O |
| VR168670 | D | 2,054 | 0.49 | 0.69 | 10 | 41 | 49 | N/A | O |
| VR168684 | D | 2,054 | 0.81 | 0.73 | 6 | 12 | 81 | N/A | [no flag] |
| VR194284 | D | 2,113 | 0.77 | 0.73 | 5 | 18 | 77 | N/A | [no flag] |
| VR223852 | D | 2,113 | 0.62 | 0.75 | 13 | 26 | 62 | N/A | O |
| VR053990 | P | 2,113 | 1.47 | 0.71 | 4 | 17 | 11 | 68 | [no flag] |
| VR140236 | P | 2,113 | 0.93 | 0.78 | 9 | 36 | 18 | 38 | [no flag] |
| VR140266 | P | 2,113 | 1.03 | 0.76 | 9 | 32 | 17 | 43 | [no flag] |
| VR144428 | P | 2,054 | 1.46 | 0.79 | 4 | 19 | 8 | 69 | [no flag] |
| VR144686 | P | 2,113 | 1.40 | 0.78 | 4 | 22 | 7 | 66 | [no flag] |
| VR150995 | P | 2,113 | 1.18 | 0.80 | 8 | 28 | 11 | 54 | [no flag] |
| VR151097 | P | 2,054 | 0.83 | 0.70 | 8 | 40 | 20 | 31 | [no flag] |
| VR155154 | P | 2,054 | 1.28 | 0.73 | 7 | 12 | 35 | 46 | [no flag] |
| VR155163 | P | 2,054 | 1.18 | 0.75 | 7 | 18 | 33 | 43 | [no flag] |
| VR194301 | P | 2,113 | 1.26 | 0.71 | 4 | 26 | 14 | 56 | [no flag] |
| VR210937 | P | 2,113 | 0.70 | 0.68 | 8 | 48 | 18 | 26 | [no flag] |
| VR222572 | P | 2,113 | 1.24 | 0.76 | 8 | 17 | 27 | 49 | [no flag] |
| VR222573 | P | 2,113 | 0.93 | 0.81 | 8 | 38 | 14 | 39 | [no flag] |

Table 8.A. 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 | Flag |
| VR054031 | D | 1,747 | 0.90 | 0.73 | 4 | 5 | 90 | N/A | [no flag] |
| VR131713 | D | 1,819 | 0.54 | 0.75 | 7 | 39 | 54 | N/A | [no flag] |
| VR131809 | D | 1,819 | 0.76 | 0.76 | 7 | 17 | 76 | N/A | [no flag] |
| VR132153 | D | 1,819 | 0.64 | 0.61 | 7 | 29 | 64 | N/A | [no flag] |
| VR132159 | D | 1,819 | 0.65 | 0.78 | 9 | 26 | 65 | N/A | [no flag] |
| VR132196 | D | 1,819 | 0.70 | 0.78 | 7 | 23 | 70 | N/A | [no flag] |
| VR132209 | D | 1,819 | 0.75 | 0.77 | 7 | 18 | 75 | N/A | [no flag] |
| VR132300 | D | 1,819 | 0.46 | 0.38 | 9 | 45 | 46 | N/A | [no flag] |
| VR132305 | D | 1,819 | 0.67 | 0.79 | 8 | 25 | 67 | N/A | [no flag] |
| VR133740 | D | 1,819 | 0.77 | 0.75 | 6 | 17 | 77 | N/A | [no flag] |
| VR133759 | D | 1,819 | 0.71 | 0.69 | 7 | 22 | 71 | N/A | [no flag] |
| VR133797 | D | 1,819 | 0.74 | 0.76 | 9 | 17 | 74 | N/A | [no flag] |
| VR135915 | D | 1,819 | 0.69 | 0.82 | 9 | 23 | 69 | N/A | [no flag] |
| VR146736 | D | 1,819 | 0.85 | 0.78 | 5 | 9 | 85 | N/A | [no flag] |
| VR147461 | D | 1,819 | 0.82 | 0.72 | 6 | 12 | 82 | N/A | [no flag] |
| VR148419 | D | 1,819 | 0.67 | 0.69 | 9 | 24 | 67 | N/A | [no flag] |
| VR148512 | D | 1,819 | 0.41 | 0.51 | 9 | 50 | 41 | N/A | [no flag] |
| VR148515 | D | 1,819 | 0.65 | 0.76 | 8 | 27 | 65 | N/A | [no flag] |
| VR148557 | D | 1,819 | 0.48 | 0.55 | 10 | 41 | 48 | N/A | O |
| VR148842 | D | 1,747 | 0.75 | 0.82 | 10 | 15 | 75 | N/A | O |
| VR148853 | D | 1,747 | 0.63 | 0.77 | 10 | 27 | 63 | N/A | O |
| VR148858 | D | 1,747 | 0.54 | 0.78 | 11 | 36 | 54 | N/A | O |
| VR148916 | D | 1,747 | 0.46 | 0.60 | 13 | 41 | 46 | N/A | O |
| VR150141 | D | 1,819 | 0.78 | 0.78 | 8 | 14 | 78 | N/A | [no flag] |
| VR150176 | D | 1,819 | 0.58 | 0.75 | 8 | 34 | 58 | N/A | [no flag] |
| VR150177 | D | 1,819 | 0.86 | 0.76 | 6 | 8 | 86 | N/A | [no flag] |
| VR150178 | D | 1,819 | 0.70 | 0.78 | 7 | 22 | 70 | N/A | [no flag] |
| VR150180 | D | 1,819 | 0.62 | 0.78 | 8 | 30 | 62 | N/A | [no flag] |
| VR150186 | D | 1,747 | 0.77 | 0.78 | 9 | 14 | 77 | N/A | [no flag] |

Table 8.A.5 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR150187 | D | 1,747 | 0.65 | 0.76 | 9 | 25 | 65 | N/A | [no flag] |
| VR150188 | D | 1,747 | 0.73 | 0.72 | 9 | 18 | 73 | N/A | [no flag] |
| VR150189 | D | 1,747 | 0.66 | 0.80 | 9 | 25 | 66 | N/A | [no flag] |
| VR150190 | D | 1,747 | 0.69 | 0.67 | 9 | 22 | 69 | N/A | [no flag] |
| VR166526 | D | 1,747 | 0.69 | 0.77 | 9 | 22 | 69 | N/A | [no flag] |
| VR166576 | D | 1,747 | 0.69 | 0.78 | 10 | 21 | 69 | N/A | O |
| VR166709 | D | 1,819 | 0.57 | 0.63 | 7 | 37 | 57 | N/A | [no flag] |
| VR167935 | D | 1,819 | 0.67 | 0.69 | 9 | 24 | 67 | N/A | [no flag] |
| VR196656 | D | 1,819 | 0.91 | 0.73 | 4 | 4 | 91 | N/A | [no flag] |
| VR133773 | P | 1,819 | 1.27 | 0.77 | 5 | 23 | 17 | 55 | [no flag] |
| VR133811 | P | 1,819 | 1.18 | 0.79 | 6 | 23 | 23 | 47 | [no flag] |
| VR146758 | P | 1,819 | 1.50 | 0.77 | 4 | 18 | 7 | 72 | [no flag] |
| VR147469 | P | 1,819 | 1.63 | 0.78 | 4 | 12 | 7 | 78 | H |
| VR148552 | P | 1,819 | 1.00 | 0.79 | 7 | 32 | 22 | 39 | [no flag] |
| VR148864 | P | 1,747 | 1.23 | 0.78 | 7 | 25 | 13 | 55 | [no flag] |
| VR166594 | P | 1,747 | 1.14 | 0.78 | 7 | 27 | 17 | 48 | [no flag] |
| VR166612 | P | 1,747 | 1.35 | 0.76 | 6 | 12 | 27 | 54 | [no flag] |
| VR167959 | P | 1,819 | 0.90 | 0.76 | 7 | 36 | 25 | 33 | [no flag] |
| VR167974 | P | 1,819 | 1.30 | 0.77 | 6 | 13 | 32 | 49 | [no flag] |
| VR196675 | P | 1,819 | 1.64 | 0.78 | 3 | 12 | 6 | 79 | H |
| VR213047 | P | 1,747 | 1.69 | 0.76 | 3 | 9 | 6 | 81 | H |

Table 8.A. 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 | Flag |
| VR056029 | D | 2,578 | 0.68 | 0.75 | 9 | 23 | 68 | N/A | [no flag] |
| VR056049 | D | 2,578 | 0.46 | 0.71 | 11 | 43 | 46 | N/A | O |
| VR132587 | D | 2,578 | 0.79 | 0.78 | 7 | 14 | 79 | N/A | [no flag] |
| VR132598 | D | 2,578 | 0.67 | 0.79 | 8 | 25 | 67 | N/A | [no flag] |
| VR132612 | D | 2,578 | 0.68 | 0.66 | 7 | 25 | 68 | N/A | [no flag] |
| VR132681 | D | 2,578 | 0.63 | 0.72 | 8 | 29 | 63 | N/A | [no flag] |
| VR132695 | D | 2,578 | 0.70 | 0.77 | 9 | 21 | 70 | N/A | [no flag] |
| VR132848 | D | 2,578 | 0.72 | 0.76 | 10 | 18 | 72 | N/A | [no flag] |
| VR133507 | D | 2,578 | 0.60 | 0.73 | 10 | 30 | 60 | N/A | [no flag] |
| VR133887 | D | 2,578 | 0.72 | 0.75 | 8 | 20 | 72 | N/A | [no flag] |
| VR133915 | D | 2,578 | 0.69 | 0.76 | 8 | 23 | 69 | N/A | [no flag] |
| VR133981 | D | 2,578 | 0.65 | 0.76 | 9 | 26 | 65 | N/A | [no flag] |
| VR134023 | D | 2,578 | 0.75 | 0.77 | 9 | 16 | 75 | N/A | [no flag] |
| VR144835 | D | 2,578 | 0.87 | 0.80 | 6 | 7 | 87 | N/A | [no flag] |
| VR145306 | D | 2,578 | 0.83 | 0.72 | 7 | 10 | 83 | N/A | [no flag] |
| VR147932 | D | 2,347 | 0.72 | 0.79 | 6 | 22 | 72 | N/A | [no flag] |
| VR148029 | D | 2,347 | 0.70 | 0.78 | 8 | 22 | 70 | N/A | [no flag] |
| VR148031 | D | 2,347 | 0.78 | 0.68 | 6 | 16 | 78 | N/A | [no flag] |
| VR148050 | D | 2,347 | 0.65 | 0.77 | 7 | 27 | 65 | N/A | [no flag] |
| VR148051 | D | 2,347 | 0.71 | 0.77 | 8 | 21 | 71 | N/A | [no flag] |
| VR150116 | D | 2,578 | 0.78 | 0.81 | 7 | 15 | 78 | N/A | [no flag] |
| VR150120 | D | 2,578 | 0.77 | 0.80 | 7 | 16 | 77 | N/A | [no flag] |
| VR150125 | D | 2,578 | 0.73 | 0.71 | 7 | 20 | 73 | N/A | [no flag] |
| VR150129 | D | 2,578 | 0.81 | 0.74 | 7 | 12 | 81 | N/A | [no flag] |
| VR150131 | D | 2,578 | 0.61 | 0.74 | 9 | 30 | 61 | N/A | [no flag] |
| VR150493 | D | 2,347 | 0.76 | 0.74 | 8 | 16 | 76 | N/A | [no flag] |
| VR150497 | D | 2,347 | 0.66 | 0.76 | 9 | 25 | 66 | N/A | [no flag] |
| VR150501 | D | 2,347 | 0.65 | 0.65 | 9 | 26 | 65 | N/A | [no flag] |
| VR150530 | D | 2,347 | 0.39 | 0.56 | 10 | 52 | 39 | N/A | [no flag] |

Table 8.A.6 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Item Type | N | AIS | Item-Total Correlation | Percent Omit | Percent Score Point = 0 | Percent Score Point = 1 | Percent Score Point = 2 | Flag |
| VR154631 | D | 2,347 | 0.67 | 0.80 | 9 | 24 | 67 | N/A | [no flag] |
| VR154835 | D | 2,347 | 0.59 | 0.65 | 9 | 32 | 59 | N/A | [no flag] |
| VR155066 | D | 2,578 | 0.70 | 0.71 | 9 | 21 | 70 | N/A | [no flag] |
| VR155071 | D | 2,578 | 0.71 | 0.77 | 9 | 19 | 71 | N/A | [no flag] |
| VR155828 | D | 2,578 | 0.76 | 0.81 | 7 | 17 | 76 | N/A | [no flag] |
| VR169370 | D | 2,578 | 0.60 | 0.60 | 10 | 30 | 60 | N/A | [no flag] |
| VR191181 | D | 2,347 | 0.89 | 0.75 | 4 | 7 | 89 | N/A | [no flag] |
| VR191356 | D | 2,578 | 0.92 | 0.72 | 5 | 3 | 92 | N/A | [no flag] |
| VR224889 | D | 2,578 | 0.51 | 0.68 | 10 | 39 | 51 | N/A | O |
| VR132823 | P | 2,578 | 1.17 | 0.83 | 7 | 27 | 14 | 52 | [no flag] |
| VR132834 | P | 2,578 | 1.28 | 0.82 | 8 | 22 | 13 | 58 | [no flag] |
| VR144875 | P | 2,578 | 1.64 | 0.78 | 5 | 10 | 5 | 79 | H |
| VR145310 | P | 2,578 | 1.43 | 0.79 | 5 | 19 | 9 | 67 | [no flag] |
| VR150525 | P | 2,347 | 0.97 | 0.71 | 7 | 34 | 21 | 38 | [no flag] |
| VR154860 | P | 2,347 | 1.36 | 0.76 | 7 | 12 | 26 | 55 | [no flag] |
| VR154926 | P | 2,347 | 1.02 | 0.75 | 7 | 29 | 24 | 39 | [no flag] |
| VR155083 | P | 2,578 | 1.31 | 0.76 | 7 | 13 | 29 | 51 | [no flag] |
| VR155088 | P | 2,578 | 0.97 | 0.77 | 8 | 33 | 22 | 38 | [no flag] |
| VR191268 | P | 2,347 | 1.55 | 0.74 | 4 | 13 | 9 | 73 | [no flag] |
| VR191381 | P | 2,578 | 1.51 | 0.72 | 4 | 15 | 11 | 70 | [no flag] |
| VR218864 | P | 2,578 | 1.01 | 0.79 | 8 | 33 | 18 | 42 | [no flag] |

### Appendix 8.B: Differential Item Functioning Results

**Note:** In table 8.B.1 and table 8.B.2, “N/A” indicates that these items did not have sufficient sample to conduct DIF analyses or the Mantel-Haenszel statistics cannot be calculated because of the lack of variability in the data.

Table 8.B. Gender DIF Classifications Summary by Grade Level or Grade Span

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DIF Category | Kindergarten Number of items | Kindergarten Percent | Grade 1 Number of items | Grade 1 Percent | Grade 2 Number of items | Grade 2 Percent | Grade Span 3–5 Number of items | Grade Span 3–5 Percent | Grade Span 6–8 Number of items | Grade Span 6–8 Percent | Grade Span 9–12 Number of items | Grade Span 9–12 Percent |
| C- | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B- | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| A- | 16 | 47 | 16 | 47 | 16 | 47 | 25 | 51 | 25 | 50 | 24 | 48 |
| A+ | 18 | 53 | 17 | 50 | 18 | 53 | 24 | 49 | 25 | 50 | 24 | 48 |
| B+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 |
| C+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| N/A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Items Total:** | 34 | 100 | 34 | 100 | 34 | 100 | 49 | 100 | 50 | 100 | 50 | 100 |

Table 8.B. Hispanic or Latino versus Non-Hispanic and Non-Latino DIF Classifications Summary by Grade Level or Grade Span

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| DIF Category | Kindergarten Number of items | Kindergarten Percent | Grade 1 Number of items | Grade 1 Percent | Grade 2 Number of items | Grade 2 Percent | Grade Span 3–5 Number of items | Grade Span 3–5 Percent | Grade Span 6–8 Number of items | Grade Span 6–8 Percent | Grade Span 9–12 Number of items | Grade Span 9–12 Percent |
| C- | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B- | 3 | 9 | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 1 | 2 |
| A- | 16 | 47 | 16 | 47 | 16 | 47 | 23 | 47 | 25 | 50 | 22 | 44 |
| A+ | 14 | 41 | 18 | 53 | 16 | 47 | 24 | 49 | 24 | 48 | 27 | 54 |
| B+ | 1 | 3 | 0 | 0 | 2 | 6 | 1 | 2 | 1 | 2 | 0 | 0 |
| C+ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| N/A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **Items Total:** | 34 | 100 | 34 | 100 | 34 | 100 | 49 | 100 | 50 | 100 | 50 | 100 |

Note: In table 8.B.3, “N/A” is listed in the *SMD* column because all items in this table are dichotomous, and dichotomous items do not have SMD values.

Table 8.B. Items Exhibiting Significant DIF by Primary Disability Student Group

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level | Item ID | N Focal | N Reference | MHDIF | SMD | Comparison | In Favor Of |
| 6–8 | VR054031 | 141 | 762 | -4.36 | N/A | Intellectual disability–Specific learning disability | Intellectual disability |
| 9–12 | VR148029 | 723 | 1170 | 1.66 | N/A | Intellectual disability–Autism | Autism |
| 9–12 | VR150131 | 112 | 1356 | 4.85 | N/A | Intellectual disability–Specific learning disability | Specific learning disability |
| 9–12 | VR150497 | 723 | 1170 | 1.66 | N/A | Intellectual disability–Autism | Autism |

### Appendix 8.C: Item Response Theory Results

**Note:** In table 8.C.1 through table 8.C.6, “N/A” indicates that these items did not have *d*‑parameter estimates, and SE = standard error.

Table 8.C. IRT Item Statistics, Kindergarten

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR052871 | 1.49 | 0.12 | N/A | N/A |
| VR053826 | -1.79 | 0.13 | N/A | N/A |
| VR131310 | 2.51 | 0.14 | N/A | N/A |
| VR131342 | 0.62 | 0.11 | N/A | N/A |
| VR131419 | 1.01 | 0.09 | -0.60 | 0.60 |
| VR131450 | 0.60 | 0.11 | N/A | N/A |
| VR131594 | 1.74 | 0.12 | N/A | N/A |
| VR131679 | 1.26 | 0.12 | N/A | N/A |
| VR131683 | 0.87 | 0.11 | N/A | N/A |
| VR131687 | 1.27 | 0.12 | N/A | N/A |
| VR131711 | 0.97 | 0.09 | -0.43 | 0.43 |
| 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 |
| VR139042 | 0.71 | 0.12 | N/A | N/A |
| 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 |
| VR139975 | 1.08 | 0.09 | -0.48 | 0.48 |
| 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 |
| VR216450 | -0.35 | 0.09 | -0.81 | 0.81 |
| VR221674 | 0.92 | 0.11 | N/A | N/A |
| VR223164 | 1.28 | 0.09 | -0.22 | 0.22 |
| VR244385 | 1.23 | 0.12 | N/A | N/A |

Table 8.C. IRT Item Statistics, Grade One

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR053944 | -1.99 | 0.12 | N/A | N/A |
| VR053970 | -0.69 | 0.08 | -0.71 | 0.71 |
| 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 |
| VR130343 | 0.85 | 0.08 | -0.14 | 0.14 |
| VR130345 | 0.09 | 0.10 | N/A | N/A |
| VR130374 | 0.23 | 0.10 | N/A | N/A |
| VR130402 | -0.12 | 0.10 | N/A | N/A |
| VR130421 | 0.58 | 0.10 | N/A | N/A |
| VR130428 | 0.69 | 0.08 | -0.35 | 0.35 |
| 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 |
| VR134003 | 0.45 | 0.08 | -0.50 | 0.50 |
| 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 |
| VR138601 | 0.52 | 0.10 | N/A | N/A |
| VR138628 | 0.07 | 0.10 | N/A | N/A |
| VR150660 | 0.25 | 0.10 | N/A | N/A |
| VR150666 | -0.24 | 0.11 | 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.C. 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 |
| VR130298 | 0.62 | 0.08 | -0.10 | 0.10 |
| VR130352 | 0.31 | 0.11 | N/A | N/A |
| VR130367 | 0.10 | 0.11 | N/A | N/A |
| VR130401 | 0.71 | 0.11 | N/A | N/A |
| VR130438 | 0.15 | 0.11 | N/A | N/A |
| VR134637 | 0.04 | 0.11 | N/A | N/A |
| VR134649 | -0.10 | 0.11 | N/A | N/A |
| VR134664 | 0.90 | 0.08 | -0.12 | 0.12 |
| 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 |
| VR140523 | 0.91 | 0.11 | N/A | N/A |
| VR151565 | 0.61 | 0.11 | N/A | N/A |
| VR151573 | 0.38 | 0.11 | N/A | N/A |
| VR151587 | 1.45 | 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 |
| VR193816 | -1.80 | 0.12 | N/A | N/A |
| VR193828 | -0.23 | 0.08 | -0.31 | 0.31 |
| 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.C. IRT Item Statistics, Grade Span Three Through Five

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR053180 | -1.27 | 0.08 | N/A | N/A |
| VR053988 | -2.11 | 0.07 | N/A | N/A |
| VR053990 | -1.13 | 0.05 | -0.62 | 0.62 |
| VR054453 | -0.56 | 0.08 | N/A | N/A |
| VR054493 | -0.32 | 0.08 | N/A | N/A |
| VR060298 | -0.34 | 0.07 | N/A | N/A |
| VR130469 | -0.79 | 0.06 | N/A | N/A |
| VR130481 | -0.86 | 0.06 | N/A | N/A |
| VR130500 | 0.59 | 0.06 | N/A | N/A |
| VR130526 | -0.03 | 0.06 | 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 |
| VR140266 | 0.05 | 0.04 | -0.31 | 0.31 |
| VR144415 | -2.05 | 0.09 | N/A | N/A |
| VR144428 | -1.01 | 0.05 | -1.20 | 1.20 |
| 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 |
| VR146075 | 0.41 | 0.07 | 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 |
| VR151042 | 0.27 | 0.07 | N/A | N/A |
| VR151060 | 0.70 | 0.07 | N/A | N/A |
| VR151083 | 1.43 | 0.07 | N/A | N/A |
| VR151087 | 0.24 | 0.07 | N/A | N/A |
| VR151097 | 0.42 | 0.05 | -0.19 | 0.19 |
| VR155150 | -0.17 | 0.07 | N/A | N/A |
| VR155154 | -0.82 | 0.06 | 0.89 | -0.89 |

Table 8.C.4 *(continuation)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR155163 | -0.49 | 0.06 | 0.67 | -0.67 |
| VR155166 | -0.63 | 0.07 | N/A | N/A |
| VR168670 | 0.15 | 0.07 | N/A | N/A |
| VR168684 | -1.99 | 0.09 | N/A | N/A |
| VR194284 | -1.54 | 0.08 | N/A | N/A |
| VR194301 | -0.61 | 0.05 | -0.58 | 0.58 |
| VR210937 | 0.84 | 0.04 | -0.14 | 0.14 |
| 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.C. IRT Item Statistics, Grade Span Six Through Eight

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR054031 | -3.42 | 0.14 | N/A | N/A |
| VR131713 | -0.29 | 0.07 | N/A | N/A |
| VR131809 | -1.74 | 0.08 | N/A | N/A |
| VR132153 | -0.87 | 0.07 | N/A | N/A |
| VR132159 | -0.86 | 0.07 | N/A | N/A |
| 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 |
| VR133773 | -0.75 | 0.05 | -0.19 | 0.19 |
| VR133797 | -1.54 | 0.08 | N/A | N/A |
| VR133811 | -0.56 | 0.06 | 0.22 | -0.22 |
| VR135915 | -1.06 | 0.08 | N/A | N/A |
| VR146736 | -2.61 | 0.12 | N/A | N/A |
| VR146758 | -1.34 | 0.07 | -1.15 | 1.15 |
| VR147461 | -2.31 | 0.08 | N/A | N/A |
| VR147469 | -1.93 | 0.06 | -0.89 | 0.89 |
| VR148419 | -1.09 | 0.09 | N/A | N/A |
| VR148512 | 0.63 | 0.08 | N/A | N/A |
| VR148515 | -0.91 | 0.09 | N/A | N/A |
| VR148552 | 0.02 | 0.06 | -0.02 | 0.02 |
| VR148557 | 0.09 | 0.08 | N/A | N/A |
| VR148842 | -1.65 | 0.10 | N/A | N/A |
| 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 |
| VR150141 | -1.88 | 0.10 | N/A | N/A |
| VR150176 | -0.38 | 0.09 | N/A | N/A |
| VR150177 | -2.72 | 0.12 | N/A | N/A |
| VR150178 | -1.28 | 0.10 | N/A | N/A |
| VR150180 | -0.69 | 0.09 | 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 |

Table 8.C.5 *(continuation)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR150190 | -1.19 | 0.09 | N/A | N/A |
| VR166526 | -1.22 | 0.09 | N/A | N/A |
| VR166576 | -1.23 | 0.10 | N/A | N/A |
| VR166594 | -0.36 | 0.06 | -0.06 | 0.06 |
| VR166612 | -1.17 | 0.07 | 0.81 | -0.81 |
| 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 |
| VR213047 | -2.23 | 0.08 | -0.83 | 0.83 |

Table 8.C. IRT Item Statistics, Grade Span Nine Through Twelve

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR056029 | -1.24 | 0.08 | N/A | N/A |
| VR056049 | 0.16 | 0.07 | N/A | N/A |
| VR132587 | -2.18 | 0.07 | N/A | N/A |
| VR132598 | -1.24 | 0.06 | N/A | N/A |
| VR132612 | -1.27 | 0.06 | N/A | N/A |
| 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 |
| VR132834 | -0.95 | 0.05 | -0.53 | 0.53 |
| 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 |
| VR133981 | -1.01 | 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 |
| VR145306 | -2.62 | 0.07 | N/A | N/A |
| VR145310 | -1.42 | 0.05 | -0.93 | 0.93 |
| 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 |
| VR148051 | -1.36 | 0.09 | N/A | N/A |
| VR150116 | -2.11 | 0.09 | N/A | N/A |
| VR150120 | -1.96 | 0.09 | N/A | N/A |
| VR150125 | -1.66 | 0.08 | N/A | N/A |
| VR150129 | -2.37 | 0.09 | N/A | N/A |
| VR150131 | -0.83 | 0.08 | N/A | N/A |
| VR150493 | -1.72 | 0.09 | N/A | N/A |
| VR150497 | -0.99 | 0.08 | N/A | N/A |
| VR150501 | -0.91 | 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 |

Table 8.C.6 *(continuation)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item ID | Item Difficulty b | SE | d1 | d2 |
| VR155066 | -1.50 | 0.08 | N/A | N/A |
| VR155071 | -1.56 | 0.08 | N/A | N/A |
| VR155083 | -1.20 | 0.06 | 0.76 | -0.76 |
| VR155088 | -0.01 | 0.05 | 0.10 | -0.10 |
| VR155828 | -1.84 | 0.07 | N/A | N/A |
| VR169370 | -0.89 | 0.07 | N/A | N/A |
| VR191181 | -3.20 | 0.12 | N/A | N/A |
| VR191268 | -1.69 | 0.06 | -0.48 | 0.48 |
| VR191356 | -3.98 | 0.09 | N/A | N/A |
| VR191381 | -1.74 | 0.05 | -0.42 | 0.42 |
| VR218864 | -0.18 | 0.05 | -0.10 | 0.10 |
| VR224889 | -0.21 | 0.07 | N/A | N/A |

Table 8.C.7 through table 8.C.12 present the item–person map for each grade level or grade span. Note the following about these tables:

* In the *Student Theta Distribution* column, “X” represents 5 students, “.” represents a value in between 1 and 4 students, and no students are denoted as “-”.

**PL2**

* In the *Item Domain(s)* column, “R” represents a receptive item and “E” represents an expressive item.
* A hyphen (“-”) represents no correspondence with a theta value.
* The horizontal dashed lines divide the space on the logit scale into performance levels.

Table 8.C. Item–Person Map for Kindergarten

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 4 | 3 | . | 4.0 | - | 0 |
| 0 | 3 | - | 3.8 | - | 0 |
| 5 | 3 | x | 3.6 | - | 0 |
| 0 | 3 | - | 3.4 | - | 0 |
| 3 | 3 | . | 3.2 | - | 0 |
| 0 | 3 | - | 3.0 | - | 0 |
| 1 | 3 | . | 2.8 | - | 0 |
| 7 | 3 | .x | 2.6 | R | 1 |
| 7 | 3 | .x | 2.4 | - | 0 |
| 8 | 3 | .x | 2.2 | - | 0 |
| 21 | 3 | .xxxx | 2.0 | R | 1 |
| 12 | 3 | .xx | 1.8 | R | 1 |
| 26 | 3 | .xxxxx | 1.6 | E E | 2 |
| 25 | 2 | xxxxx | 1.4 | E R E | 3 |
| 33 | 2 | .xxxxxx | 1.2 | E R R E E E E | 7 |
| 64 | 2 | .xxxxxxxxxxxx | 1.0 | E R R E | 4 |
| 45 | 2 | xxxxxxxxx | 0.8 | E R E R | 4 |
| 47 | 2 | .xxxxxxxxx | 0.6 | E R E E | 4 |
| 47 | 2 | .xxxxxxxxx | 0.4 | E E E R R | 5 |

Table 8.C.7 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 34 | 2 | .xxxxxx | 0.2 | E E | 2 |
| 50 | 2 | xxxxxxxxxx | 0.0 | E R E R | 4 |
| 27 | 1 | .xxxxx | -0.2 | R | 1 |
| 26 | 1 | .xxxxx | -0.4 | - | 0 |
| 31 | 1 | .xxxxxx | -0.6 | E | 1 |
| 27 | 1 | .xxxxx | -0.8 | R | 1 |
| 15 | 1 | xxx | -1.0 | R E | 2 |
| 19 | 1 | .xxx | -1.2 | E | 1 |
| 20 | 1 | xxxx | -1.4 | - | 0 |
| 16 | 1 | .xxx | -1.6 | - | 0 |
| 0 | 1 | - | -1.8 | R | 1 |
| 16 | 1 | .xxx | -2.0 | - | 0 |
| 0 | 1 | - | -2.2 | - | 0 |
| 19 | 1 | .xxx | -2.4 | - | 0 |
| 0 | 1 | - | -2.6 | - | 0 |
| 0 | 1 | - | -2.8 | - | 0 |
| 0 | 1 | - | -3.0 | - | 0 |
| 29 | 1 | .xxxxx | -3.2 | - | 0 |
| 0 | 1 | - | -3.4 | - | 0 |
| 0 | 1 | - | -3.6 | - | 0 |
| 0 | 1 | - | -3.8 | - | 0 |
| 43 | 1 | .xxxxxxxx | -4.0 | - | 0 |

Table 8.C. Item–Person Map for Grade One

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 10 | 3 | xx | 4.0 | - | 0 |
| 11 | 3 | .xx | 3.8 | - | 0 |
| 0 | 3 | - | 3.6 | - | 0 |
| 0 | 3 | - | 3.4 | - | 0 |
| 0 | 3 | - | 3.2 | - | 0 |
| 1 | 3 | . | 3.0 | - | 0 |
| 0 | 3 | - | 2.8 | - | 0 |
| 9 | 3 | .x | 2.6 | - | 0 |
| 0 | 3 | - | 2.4 | - | 0 |
| 18 | 3 | .xxx | 2.2 | - | 0 |
| 7 | 3 | .x | 2.0 | - | 0 |
| 20 | 3 | xxxx | 1.8 | - | 0 |
| 33 | 3 | .xxxxxx | 1.6 | - | 0 |
| 13 | 3 | .xx | 1.4 | - | 0 |
| 26 | 3 | .xxxxx | 1.2 | R E | 2 |
| 34 | 2 | .xxxxxx | 1.0 | E E E E E E R | 7 |
| 46 | 2 | .xxxxxxxxx | 0.8 | E E E | 3 |
| 85 | 2 | xxxxxxxxxxxxxxxxx | 0.6 | E R R E | 4 |
| 59 | 2 | .xxxxxxxxxxx | 0.4 | E R R E | 4 |
| 61 | 2 | .xxxxxxxxxxxx | 0.2 | E R R R | 4 |
| 79 | 2 | .xxxxxxxxxxxxxxx | 0.0 | E E E R E | 5 |
| 44 | 2 | .xxxxxxxx | -0.2 | R R E R E | 5 |
| 47 | 1 | .xxxxxxxxx | -0.4 | E R | 2 |
| 38 | 1 | .xxxxxxx | -0.6 | R R E | 3 |
| 13 | 1 | .xx | -0.8 | R | 1 |
| 23 | 1 | .xxxx | -1.0 | - | 0 |
| 15 | 1 | xxx | -1.2 | - | 0 |
| 27 | 1 | .xxxxx | -1.4 | E R | 2 |
| 9 | 1 | .x | -1.6 | E | 1 |
| 9 | 1 | .x | -1.8 | - | 0 |

Table 8.C.8 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 0 | 1 | - | -2.0 | R | 1 |
| 15 | 1 | xxx | -2.2 | - | 0 |
| 11 | 1 | .xx | -2.4 | - | 0 |
| 0 | 1 | - | -2.6 | - | 0 |
| 0 | 1 | - | -2.8 | - | 0 |
| 12 | 1 | .xx | -3.0 | - | 0 |
| 0 | 1 | - | -3.2 | - | 0 |
| 0 | 1 | - | -3.4 | - | 0 |
| 26 | 1 | .xxxxx | -3.6 | - | 0 |
| 0 | 1 | - | -3.8 | - | 0 |
| 32 | 1 | .xxxxxx | -4.0 | - | 0 |

Table 8.C. Item–Person Map for Grade Two

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 4 | 3 | . | 4.0 | - | 0 |
| 6 | 3 | .x | 3.8 | - | 0 |
| 0 | 3 | - | 3.6 | - | 0 |
| 0 | 3 | - | 3.4 | - | 0 |
| 4 | 3 | . | 3.2 | - | 0 |
| 0 | 3 | - | 3.0 | - | 0 |
| 6 | 3 | .x | 2.8 | - | 0 |
| 0 | 3 | - | 2.6 | - | 0 |
| 13 | 3 | .xx | 2.4 | - | 0 |
| 15 | 3 | xxx | 2.2 | - | 0 |
| 19 | 3 | .xxx | 2.0 | - | 0 |
| 17 | 3 | .xxx | 1.8 | E | 1 |
| 16 | 3 | .xxx | 1.6 | - | 0 |
| 25 | 3 | xxxxx | 1.4 | R | 1 |
| 32 | 2 | .xxxxxx | 1.2 | - | 0 |
| 39 | 2 | .xxxxxxx | 1.0 | E E E E | 4 |
| 38 | 2 | .xxxxxxx | 0.8 | R E R E E E E | 7 |
| 51 | 2 | .xxxxxxxxxx | 0.6 | E R | 2 |
| 48 | 2 | .xxxxxxxxx | 0.4 | E R R R R E | 6 |
| 51 | 2 | .xxxxxxxxxx | 0.2 | E | 1 |
| 54 | 2 | .xxxxxxxxxx | 0.0 | E E R R R R | 6 |
| 50 | 1 | xxxxxxxxxx | -0.2 | R E R E | 4 |
| 34 | 1 | .xxxxxx | -0.4 | R E | 2 |
| 44 | 1 | .xxxxxxxx | -0.6 | E E | 2 |
| 38 | 1 | .xxxxxxx | -0.8 | R E | 2 |
| 35 | 1 | xxxxxxx | -1.0 | R | 1 |
| 12 | 1 | .xx | -1.2 | R | 1 |
| 8 | 1 | .x | -1.4 | - | 0 |
| 14 | 1 | .xx | -1.6 | - | 0 |
| 11 | 1 | .xx | -1.8 | E R E | 3 |

Table 8.C.9 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 9 | 1 | .x | -2.0 | - | 0 |
| 14 | 1 | .xx | -2.2 | - | 0 |
| 17 | 1 | .xxx | -2.4 | - | 0 |
| 0 | 1 | - | -2.6 | - | 0 |
| 0 | 1 | - | -2.8 | - | 0 |
| 9 | 1 | .x | -3.0 | - | 0 |
| 0 | 1 | - | -3.2 | - | 0 |
| 0 | 1 | - | -3.4 | - | 0 |
| 18 | 1 | .xxx | -3.6 | - | 0 |
| 0 | 1 | - | -3.8 | - | 0 |
| 22 | 1 | .xxxx | -4.0 | - | 0 |

Table 8.C. Item–Person Map for Grade Span Three Through Five

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 54 | 3 | .xxxxx | 4.0 | - | 0 |
| 0 | 3 | - | 3.8 | - | 0 |
| 0 | 3 | - | 3.6 | - | 0 |
| 20 | 3 | xx | 3.4 | - | 0 |
| 44 | 3 | .xxxx | 3.2 | - | 0 |
| 0 | 3 | - | 3.0 | - | 0 |
| 0 | 3 | - | 2.8 | - | 0 |
| 30 | 3 | xxx | 2.6 | - | 0 |
| 60 | 3 | xxxxxx | 2.4 | - | 0 |
| 47 | 3 | .xxxx | 2.2 | - | 0 |
| 96 | 3 | .xxxxxxxxx | 2.0 | - | 0 |
| 48 | 3 | .xxxx | 1.8 | E | 1 |
| 46 | 3 | .xxxx | 1.6 | - | 0 |
| 91 | 3 | .xxxxxxxxx | 1.4 | E | 1 |
| 185 | 3 | .xxxxxxxxxxxxxxxxxx | 1.2 | E | 1 |
| 80 | 3 | xxxxxxxx | 1.0 | E | 1 |
| 158 | 2 | .xxxxxxxxxxxxxxx | 0.8 | E E | 2 |
| 72 | 2 | .xxxxxxx | 0.6 | E R R R | 4 |
| 149 | 2 | .xxxxxxxxxxxxxx | 0.4 | E | 1 |
| 158 | 2 | .xxxxxxxxxxxxxxx | 0.2 | R E R R E | 5 |
| 174 | 2 | .xxxxxxxxxxxxxxxxx | 0.0 | E E E R E | 5 |
| 157 | 2 | .xxxxxxxxxxxxxxx | -0.2 | E R R E R E R E R | 9 |
| 169 | 2 | .xxxxxxxxxxxxxxxx | -0.4 | R E E E E | 5 |
| 220 | 2 | xxxxxxxxxxxxxxxxxxxxxx | -0.6 | R E E E R E | 6 |
| 139 | 1 | .xxxxxxxxxxxxx | -0.8 | R R R R E R | 6 |
| 131 | 1 | .xxxxxxxxxxxxx | -1.0 | E E E | 3 |
| 120 | 1 | xxxxxxxxxxxx | -1.2 | E E R | 3 |
| 72 | 1 | .xxxxxxx | -1.4 | R | 1 |
| 66 | 1 | .xxxxxx | -1.6 | R | 1 |
| 38 | 1 | .xxx | -1.8 | E R | 2 |

Table 8.C.10 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 46 | 1 | .xxxx | -2.0 | E R R | 3 |
| 13 | 1 | .x | -2.2 | R E | 2 |
| 29 | 1 | .xx | -2.4 | - | 0 |
| 12 | 1 | .x | -2.6 | - | 0 |
| 14 | 1 | .x | -2.8 | - | 0 |
| 20 | 1 | xx | -3.0 | - | 0 |
| 0 | 1 | - | -3.2 | - | 0 |
| 17 | 1 | .x | -3.4 | - | 0 |
| 16 | 1 | .x | -3.6 | - | 0 |
| 0 | 1 | - | -3.8 | - | 0 |
| 105 | 1 | .xxxxxxxxxx | -4.0 | - | 0 |

Table 8.C. Item–Person Map for Grade Span Six Through Eight

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 128 | 3 | .xxxxxxxxxxxx | 4.0 | - | 0 |
| 0 | 3 | - | 3.8 | - | 0 |
| 0 | 3 | - | 3.6 | - | 0 |
| 0 | 3 | - | 3.4 | - | 0 |
| 0 | 3 | - | 3.2 | - | 0 |
| 0 | 3 | - | 3.0 | - | 0 |
| 78 | 3 | .xxxxxxx | 2.8 | - | 0 |
| 134 | 3 | .xxxxxxxxxxxxx | 2.6 | - | 0 |
| 0 | 3 | - | 2.4 | - | 0 |
| 0 | 3 | - | 2.2 | - | 0 |
| 77 | 3 | .xxxxxxx | 2.0 | - | 0 |
| 102 | 3 | .xxxxxxxxxx | 1.8 | - | 0 |
| 74 | 3 | .xxxxxxx | 1.6 | - | 0 |
| 79 | 3 | .xxxxxxx | 1.4 | - | 0 |
| 50 | 3 | xxxxx | 1.2 | - | 0 |
| 123 | 3 | .xxxxxxxxxxxx | 1.0 | - | 0 |
| 89 | 3 | .xxxxxxxx | 0.8 | - | 0 |
| 94 | 3 | .xxxxxxxxx | 0.6 | E R | 2 |
| 140 | 2 | xxxxxxxxxxxxxx | 0.4 | E | 1 |
| 69 | 2 | .xxxxxx | 0.2 | R | 1 |
| 113 | 2 | .xxxxxxxxxxx | 0.0 | E E E E | 4 |
| 87 | 2 | .xxxxxxxx | -0.2 | R R E E E | 5 |
| 86 | 2 | .xxxxxxxx | -0.4 | E E E E R R | 6 |
| 112 | 2 | .xxxxxxxxxxx | -0.6 | E E | 2 |
| 114 | 2 | .xxxxxxxxxxx | -0.8 | R E E R | 4 |
| 122 | 2 | .xxxxxxxxxxxx | -1.0 | E R E E E E E R R R | 10 |
| 115 | 2 | .xxxxxxxxxxx | -1.2 | E E R R R R E | 7 |
| 87 | 1 | .xxxxxxxx | -1.4 | E R | 2 |
| 89 | 1 | .xxxxxxxx | -1.6 | R E R | 3 |
| 49 | 1 | .xxxx | -1.8 | R R R R E | 5 |

Table 8.C.11 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 52 | 1 | .xxxxx | -2.0 | E | 1 |
| 20 | 1 | xx | -2.2 | - | 0 |
| 33 | 1 | .xxx | -2.4 | R E | 2 |
| 17 | 1 | .x | -2.6 | R | 1 |
| 28 | 1 | .xx | -2.8 | E E R | 3 |
| 16 | 1 | .x | -3.0 | E | 1 |
| 14 | 1 | .x | -3.2 | - | 0 |
| 11 | 1 | .x | -3.4 | R R | 2 |
| 10 | 1 | x | -3.6 | - | 0 |
| 16 | 1 | .x | -3.8 | - | 0 |
| 102 | 1 | .xxxxxxxxxx | -4.0 | - | 0 |

Table 8.C. Item–Person Map for Grade Span Nine Through Twelve

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 360 | 3 | xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx | 4.0 | - | 0 |
| 0 | 3 | - | 3.8 | - | 0 |
| 0 | 3 | - | 3.6 | - | 0 |
| 0 | 3 | - | 3.4 | - | 0 |
| 0 | 3 | - | 3.2 | - | 0 |
| 0 | 3 | - | 3.0 | - | 0 |
| 0 | 3 | - | 2.8 | - | 0 |
| 133 | 3 | .xxxxxxxxxxxxx | 2.6 | - | 0 |
| 124 | 3 | .xxxxxxxxxxxx | 2.4 | - | 0 |
| 0 | 3 | - | 2.2 | - | 0 |
| 0 | 3 | - | 2.0 | - | 0 |
| 119 | 3 | .xxxxxxxxxxx | 1.8 | - | 0 |
| 114 | 3 | .xxxxxxxxxxx | 1.6 | - | 0 |
| 101 | 3 | .xxxxxxxxxx | 1.4 | - | 0 |
| 74 | 3 | .xxxxxxx | 1.2 | - | 0 |
| 175 | 3 | .xxxxxxxxxxxxxxxxx | 1.0 | E | 1 |
| 87 | 2 | .xxxxxxxx | 0.8 | - | 0 |
| 128 | 2 | .xxxxxxxxxxxx | 0.6 | - | 0 |
| 125 | 2 | .xxxxxxxxxxxx | 0.4 | E | 1 |
| 149 | 2 | .xxxxxxxxxxxxxx | 0.2 | E E | 2 |
| 130 | 2 | xxxxxxxxxxxxx | 0.0 | E E E | 3 |
| 177 | 2 | .xxxxxxxxxxxxxxxxx | -0.2 | E E R E | 4 |
| 79 | 2 | .xxxxxxx | -0.4 | E R E E E | 5 |
| 166 | 2 | .xxxxxxxxxxxxxxxx | -0.6 | E | 1 |
| 136 | 2 | .xxxxxxxxxxxxx | -0.8 | E R E E R | 5 |
| 146 | 2 | .xxxxxxxxxxxxxx | -1.0 | E R R E R R | 6 |
| 128 | 1 | .xxxxxxxxxxxx | -1.2 | E R R E R | 5 |
| 117 | 1 | .xxxxxxxxxxx | -1.4 | E R E R R E R | 7 |
| 117 | 1 | .xxxxxxxxxxx | -1.6 | R E R R | 4 |
| 103 | 1 | .xxxxxxxxxx | -1.8 | R E E R | 4 |

Table 8.C.12 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N Students | Performance Level | Student Theta Distribution | Logit | Item Domain(s) | N Items |
| 87 | 1 | .xxxxxxxx | -2.0 | R R E | 3 |
| 48 | 1 | .xxxx | -2.2 | E E R R | 4 |
| 54 | 1 | .xxxxx | -2.4 | E E | 2 |
| 45 | 1 | .xxxx | -2.6 | R | 1 |
| 23 | 1 | .xx | -2.8 | - | 0 |
| 22 | 1 | .xx | -3.0 | R | 1 |
| 23 | 1 | .xx | -3.2 | R | 1 |
| 24 | 1 | .xx | -3.4 | E | 1 |
| 4 | 1 | . | -3.6 | - | 0 |
| 12 | 1 | .x | -3.8 | - | 0 |
| 172 | 1 | .xxxxxxxxxxxxxxxxx | -4.0 | R | 1 |

### Appendix 8.D: Response Time Results

**Notes:**

* All students who completed the test and have an unrounded test time greater than zero (0) are included.
* Grade levels reflect students’ enrolled grade levels during the 2021–22 school year.
* “N/A” indicates that there was only one student in the quartile, and SD was not available.
* Response time percentiles are identified as follows:
* “Pt. 1” is the time taken by test takers in the first percentile of response time.
* “Pt. 10” is the time taken by test takers in the tenth percentile of response time.
* “Pt. 25” is the time taken by test takers in the twenty-fifth percentile of response time.
* “Pt. 50” is the time taken by test takers in the fiftieth percentile of response time.
* “Pt. 75” is the time taken by test takers in the seventy-fifth percentile of response time.
* Pt. 90” is the time taken by test takers in the ninetieth percentile of response time.
* “Pt. 99” is the time taken by test takers in the ninety-ninth percentile of response time.

Table 8.D. Total Testing Time (in Minutes) at Each Scale Score Interval

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Scale Score Interval Based on performance level | N | Mean | SD | Min. | Max. | Pt. 1 | Pt. 10 | Pt. 25 | Pt. 50 | Pt. 75 | Pt. 90 | Pt. 99 |
| Kindergarten | 201–243 | 556 | 22.76 | 19.26 | 0.93 | 141.77 | 2.18 | 5.66 | 10.38 | 17.19 | 27.99 | 45.92 | 99.84 |
| Kindergarten | 244–259 | 545 | 26.64 | 20.38 | 2.61 | 164.67 | 6.22 | 10.81 | 14.46 | 19.74 | 31.43 | 49.85 | 112.04 |
| Kindergarten | 260–299 | 197 | 21.52 | 14.91 | 2.13 | 119.40 | 6.09 | 9.92 | 13.17 | 17.91 | 25.14 | 34.30 | 103.13 |
| 1 | 301–343 | 489 | 21.96 | 20.22 | 1.64 | 213.06 | 1.93 | 6.54 | 10.36 | 16.09 | 25.89 | 44.18 | 98.07 |
| 1 | 344–359 | 569 | 26.25 | 19.16 | 3.93 | 153.68 | 5.05 | 11.45 | 14.69 | 20.40 | 30.04 | 46.47 | 113.75 |
| 1 | 360–399 | 209 | 20.90 | 12.34 | 6.47 | 83.41 | 6.88 | 9.65 | 12.73 | 17.90 | 24.50 | 36.04 | 71.69 |
| 2 | 401–443 | 537 | 23.69 | 17.96 | 0.66 | 141.75 | 2.82 | 7.38 | 11.25 | 18.64 | 30.74 | 45.38 | 88.25 |
| 2 | 444–459 | 474 | 25.57 | 14.90 | 1.35 | 124.04 | 7.19 | 12.11 | 16.25 | 21.99 | 30.55 | 43.47 | 87.39 |
| 2 | 460–499 | 208 | 23.12 | 14.34 | 1.67 | 100.80 | 5.62 | 10.23 | 13.81 | 18.40 | 27.83 | 40.78 | 69.58 |
| 3–5 | 501–543 | 1187 | 23.42 | 17.96 | 0.64 | 171.11 | 1.77 | 7.12 | 11.72 | 19.05 | 29.27 | 45.00 | 88.01 |
| 3–5 | 544–559 | 1917 | 25.12 | 16.38 | 1.40 | 202.76 | 5.28 | 12.11 | 15.60 | 20.88 | 29.16 | 43.83 | 88.89 |
| 3–5 | 560–599 | 1126 | 19.94 | 10.08 | 2.48 | 99.17 | 6.48 | 10.98 | 13.80 | 17.88 | 22.97 | 30.62 | 59.46 |

Table 8.D.1 *(continuation)*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Grade Level or Grade Span | Scale Score Interval Based on performance level | N | Mean | SD | Min. | Max. | Pt. 1 | Pt. 10 | Pt. 25 | Pt. 50 | Pt. 75 | Pt. 90 | Pt. 99 |
| 6–8 | 601–643 | 864 | 21.54 | 18.26 | 0.94 | 205.75 | 1.95 | 5.91 | 10.62 | 16.82 | 27.20 | 40.63 | 84.10 |
| 6–8 | 644–659 | 1317 | 20.78 | 14.66 | 1.51 | 228.89 | 3.46 | 9.30 | 12.69 | 17.64 | 24.10 | 34.25 | 71.53 |
| 6–8 | 660–699 | 1445 | 17.38 | 9.55 | 1.82 | 113.28 | 5.64 | 9.32 | 11.69 | 15.19 | 19.93 | 27.46 | 59.23 |
| 9–10 | 701–743 | 485 | 19.69 | 16.10 | 0.64 | 139.36 | 1.09 | 5.07 | 9.93 | 15.28 | 24.76 | 38.66 | 76.46 |
| 9–10 | 744–759 | 745 | 18.15 | 10.42 | 1.82 | 76.77 | 3.14 | 9.15 | 11.83 | 15.34 | 21.18 | 30.36 | 61.29 |
| 9–10 | 760–799 | 521 | 15.96 | 8.28 | 2.94 | 72.57 | 5.15 | 9.09 | 10.91 | 13.60 | 18.20 | 25.22 | 49.82 |
| 11–12 | 801–843 | 925 | 18.97 | 15.26 | 0.56 | 126.73 | 1.83 | 5.02 | 9.64 | 15.06 | 23.32 | 36.41 | 81.16 |
| 11–12 | 844–859 | 1281 | 19.29 | 12.69 | 1.88 | 181.63 | 3.30 | 9.10 | 11.97 | 16.16 | 22.44 | 32.44 | 65.77 |
| 11–12 | 860–899 | 1078 | 15.35 | 9.31 | 0.95 | 144.52 | 3.75 | 8.36 | 10.41 | 13.39 | 17.97 | 23.28 | 46.76 |

### Appendix 8.E: Factor Loading

Note the following about table 8.E.1 through table 8.E.6:

* “M1” refers to the one-factor model and “M2” refers to the bifactor model.
* “N/A” indicates that the statistic is not applicable to the item in the row.

Figure 8.E.1 for kindergarten presents the factor loading comparison plot chart that was created using the data in table 8.E.1, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y-axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a strong positive correlation between the two sets of factor loadings. All the data points cluster around a diagonal reference line with slope of 1 and y‑intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa.

The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for kindergarten.

Figure 8.E. Factor loading comparison for kindergarten

Table 8.E. Standardized Factor Loading for Kindergarten

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR216450 | Expressive | 0.61 | 0.55 | N/A | 0.26 |
| VR193113 | Expressive | 0.68 | 0.61 | N/A | 0.31 |
| VR139973 | Expressive | 0.73 | 0.68 | N/A | 0.23 |
| VR139022 | Expressive | 0.67 | 0.6 | N/A | 0.32 |
| VR139042 | Expressive | 0.7 | 0.66 | N/A | 0.15 |
| VR131711 | Expressive | 0.73 | 0.66 | N/A | 0.42 |
| VR131419 | Expressive | 0.69 | 0.61 | N/A | 0.41 |
| VR131450 | Expressive | 0.69 | 0.75 | N/A | -0.08 |
| VR154458 | Expressive | 0.77 | 0.69 | N/A | 0.38 |
| VR154465 | Expressive | 0.78 | 0.70 | N/A | 0.31 |
| VR215978 | Expressive | 0.75 | 0.67 | N/A | 0.35 |
| VR139729 | Expressive | 0.54 | 0.59 | N/A | -0.10 |
| VR139975 | Expressive | 0.64 | 0.57 | N/A | 0.36 |
| VR223164 | Expressive | 0.63 | 0.55 | N/A | 0.43 |
| VR244385 | Expressive | 0.58 | 0.64 | N/A | -0.15 |
| VR053826 | Receptive | 0.65 | 0.66 | 0.15 | N/A |
| VR193093 | Receptive | 0.67 | 0.68 | 0.17 | N/A |
| VR137085 | Receptive | 0.68 | 0.71 | -0.08 | N/A |
| VR138924 | Receptive | 0.59 | 0.61 | 0.28 | N/A |
| VR138950 | Receptive | 0.55 | 0.57 | 0.03 | N/A |
| VR138982 | Receptive | 0.73 | 0.74 | 0.20 | N/A |
| VR131679 | Receptive | 0.61 | 0.65 | -0.13 | N/A |
| VR131683 | Receptive | 0.62 | 0.65 | 0.14 | N/A |
| VR131687 | Receptive | 0.61 | 0.64 | -0.23 | N/A |
| VR131310 | Receptive | 0.52 | 0.56 | -0.01 | N/A |
| VR131342 | Receptive | 0.73 | 0.76 | -0.23 | N/A |
| VR131594 | Receptive | 0.47 | 0.50 | 0.36 | N/A |
| VR154406 | Receptive | 0.86 | 0.85 | -0.17 | N/A |
| VR154449 | Receptive | 0.56 | 0.61 | 0.41 | N/A |
| VR139666 | Receptive | 0.73 | 0.75 | -0.29 | N/A |
| VR139673 | Receptive | 0.72 | 0.73 | 0.04 | N/A |
| VR052871 | Receptive | 0.69 | 0.75 | 0.03 | N/A |
| VR170322 | Receptive | 0.72 | 0.77 | -0.23 | N/A |
| VR221674 | Receptive | 0.60 | 0.64 | 0.36 | N/A |

Figure 8.E.2 for grade one presents the factor loading comparison plot chart that was created using the data in table 8.E.2, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y-axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a strong positive correlation between the two sets of factor loadings. All the data points cluster around a diagonal reference line with slope of 1 and y‑intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa. The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for grade one.

Figure 8.E. Factor loading comparison for grade one

Table 8.E. Standardized Factor Loading for Grade One

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR053970 | Expressive | 0.65 | 0.62 | N/A | 0.19 |
| VR193651 | Expressive | 0.65 | 0.62 | N/A | 0.21 |
| VR137618 | Expressive | 0.51 | 0.49 | N/A | 0.14 |
| VR138601 | Expressive | 0.59 | 0.61 | N/A | 0.03 |
| VR138628 | Expressive | 0.71 | 0.73 | N/A | 0.00 |
| VR130428 | Expressive | 0.56 | 0.54 | N/A | 0.34 |
| VR130343 | Expressive | 0.60 | 0.56 | N/A | 0.36 |
| VR130345 | Expressive | 0.69 | 0.72 | N/A | 0.04 |
| VR154753 | Expressive | 0.52 | 0.50 | N/A | 0.51 |
| VR154755 | Expressive | 0.57 | 0.53 | N/A | 0.40 |
| VR133983 | Expressive | 0.56 | 0.55 | N/A | 0.51 |
| VR134003 | Expressive | 0.62 | 0.58 | N/A | 0.36 |
| VR134007 | Expressive | 0.75 | 0.76 | N/A | 0.05 |
| VR150709 | Expressive | 0.72 | 0.75 | N/A | -0.01 |
| VR150707 | Expressive | 0.62 | 0.60 | N/A | 0.45 |
| VR053944 | Receptive | 0.69 | 0.72 | -0.10 | N/A |
| VR193603 | Receptive | 0.69 | 0.71 | 0.11 | N/A |
| VR137615 | Receptive | 0.68 | 0.72 | -0.19 | N/A |
| VR138495 | Receptive | 0.60 | 0.62 | 0.19 | N/A |
| VR138505 | Receptive | 0.72 | 0.74 | 0.19 | N/A |
| VR138567 | Receptive | 0.74 | 0.78 | -0.13 | N/A |
| VR130374 | Receptive | 0.62 | 0.66 | -0.26 | N/A |
| VR130402 | Receptive | 0.60 | 0.60 | 0.35 | N/A |
| VR130421 | Receptive | 0.61 | 0.66 | -0.12 | N/A |
| VR130319 | Receptive | 0.74 | 0.77 | 0.24 | N/A |
| VR130326 | Receptive | 0.64 | 0.68 | -0.19 | N/A |
| VR130331 | Receptive | 0.58 | 0.59 | 0.37 | N/A |
| VR154742 | Receptive | 0.64 | 0.68 | -0.28 | N/A |
| VR154751 | Receptive | 0.71 | 0.71 | 0.26 | N/A |
| VR133917 | Receptive | 0.76 | 0.78 | 0.23 | N/A |
| VR133975 | Receptive | 0.67 | 0.68 | -0.09 | N/A |
| VR150660 | Receptive | 0.68 | 0.71 | 0.26 | N/A |
| VR150666 | Receptive | 0.72 | 0.75 | 0.20 | N/A |
| VR150685 | Receptive | 0.56 | 0.57 | -0.14 | N/A |

Figure 8.E.3 for grade two presents the factor loading comparison plot chart that was created using the data in table 8.E.3, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y-axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a strong positive correlation between the two sets of factor loadings. All the data points cluster around a diagonal reference line with slope of 1 and y-intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa. The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for grade two.

Figure 8.E. Factor loading comparison for grade two

Table 8.E. Standardized Factor Loading for Grade Two

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR193885 | Expressive | 0.60 | 0.56 | N/A | 0.23 |
| VR193828 | Expressive | 0.46 | 0.42 | N/A | 0.24 |
| VR140209 | Expressive | 0.58 | 0.54 | N/A | 0.24 |
| VR140520 | Expressive | 0.70 | 0.70 | N/A | 0.10 |
| VR140523 | Expressive | 0.61 | 0.62 | N/A | -0.03 |
| VR130438 | Expressive | 0.72 | 0.75 | N/A | -0.11 |
| VR130277 | Expressive | 0.76 | 0.76 | N/A | 0.11 |
| VR130298 | Expressive | 0.48 | 0.46 | N/A | 0.31 |
| VR223063 | Expressive | 0.64 | 0.6 | N/A | 0.61 |
| VR155670 | Expressive | 0.66 | 0.62 | N/A | 0.59 |
| VR134677 | Expressive | 0.58 | 0.55 | N/A | 0.28 |
| VR134668 | Expressive | 0.70 | 0.72 | N/A | -0.09 |
| VR134664 | Expressive | 0.52 | 0.51 | N/A | 0.36 |
| VR151624 | Expressive | 0.51 | 0.52 | N/A | 0.44 |
| VR151643 | Expressive | 0.44 | 0.44 | N/A | -0.03 |
| VR193873 | Receptive | 0.65 | 0.69 | -0.26 | N/A |
| VR193816 | Receptive | 0.72 | 0.74 | -0.09 | N/A |
| VR140204 | Receptive | 0.63 | 0.67 | 0.16 | N/A |
| VR140495 | Receptive | 0.73 | 0.75 | -0.13 | N/A |
| VR140498 | Receptive | 0.69 | 0.71 | 0.16 | N/A |
| VR140501 | Receptive | 0.76 | 0.79 | 0.07 | N/A |
| VR130352 | Receptive | 0.64 | 0.66 | 0.38 | N/A |
| VR130367 | Receptive | 0.74 | 0.77 | -0.03 | N/A |
| VR130401 | Receptive | 0.70 | 0.72 | 0.28 | N/A |
| VR130125 | Receptive | 0.69 | 0.70 | 0.26 | N/A |
| VR130141 | Receptive | 0.63 | 0.65 | 0.15 | N/A |
| VR130233 | Receptive | 0.63 | 0.65 | 0.32 | N/A |
| VR155513 | Receptive | 0.74 | 0.78 | 0.23 | N/A |
| VR155674 | Receptive | 0.66 | 0.65 | 0.23 | N/A |
| VR134637 | Receptive | 0.63 | 0.64 | -0.20 | N/A |
| VR134649 | Receptive | 0.76 | 0.78 | -0.20 | N/A |
| VR151565 | Receptive | 0.52 | 0.53 | -0.21 | N/A |
| VR151573 | Receptive | 0.69 | 0.69 | 0.34 | N/A |
| VR151587 | Receptive | 0.42 | 0.39 | 0.24 | N/A |

Figure 8.E.4 for grade span three through five presents the factor loading comparison plot chart that was created using the data in table 8.E.4, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y‑axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a very strong positive correlation between the two sets of factor loadings. All the data points cluster very closely around a diagonal reference line with slope of 1 and y-intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa. Additionally, the one-factor model tends to result in slightly higher factor loadings than the bifactor model. The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for grade span three through five, with a slight favor for the one-factor model due to its slightly higher factor loadings.

Figure 8.E. Factor loading comparison for grade span three through five

Table 8.E. Standardized Factor Loading for Grade Span Three Through Five

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR053990 | Expressive | 0.52 | 0.49 | N/A | 0.21 |
| VR194301 | Expressive | 0.53 | 0.51 | N/A | 0.21 |
| VR144686 | Expressive | 0.63 | 0.61 | N/A | 0.14 |
| VR146024 | Expressive | 0.74 | 0.75 | N/A | 0.04 |
| VR146075 | Expressive | 0.63 | 0.63 | N/A | 0.04 |
| VR131628 | Expressive | 0.80 | 0.80 | N/A | 0.06 |
| VR130526 | Expressive | 0.74 | 0.75 | N/A | -0.02 |
| VR210937 | Expressive | 0.45 | 0.46 | N/A | 0.38 |
| VR222573 | Expressive | 0.66 | 0.67 | N/A | 0.34 |
| VR222572 | Expressive | 0.69 | 0.67 | N/A | 0.33 |
| VR140221 | Expressive | 0.80 | 0.82 | N/A | -0.01 |
| VR140266 | Expressive | 0.56 | 0.57 | N/A | 0.36 |
| VR140236 | Expressive | 0.62 | 0.62 | N/A | 0.32 |
| VR151004 | Expressive | 0.47 | 0.50 | N/A | -0.09 |
| VR150995 | Expressive | 0.65 | 0.64 | N/A | 0.26 |
| VR144428 | Expressive | 0.65 | 0.62 | N/A | 0.21 |
| VR054453 | Expressive | 0.80 | 0.81 | N/A | 0.03 |
| VR054493 | Expressive | 0.80 | 0.81 | N/A | 0.02 |
| VR155154 | Expressive | 0.65 | 0.61 | N/A | 0.50 |
| VR155163 | Expressive | 0.64 | 0.62 | N/A | 0.56 |
| VR151083 | Expressive | 0.32 | 0.32 | N/A | 0.05 |
| VR151097 | Expressive | 0.46 | 0.46 | N/A | 0.34 |
| VR053988 | Receptive | 0.73 | 0.73 | -0.16 | N/A |
| VR194284 | Receptive | 0.79 | 0.79 | -0.16 | N/A |
| VR144676 | Receptive | 0.72 | 0.75 | 0.17 | N/A |
| VR145701 | Receptive | 0.79 | 0.81 | 0.07 | N/A |
| VR145817 | Receptive | 0.69 | 0.72 | 0.30 | N/A |
| VR145916 | Receptive | 0.64 | 0.64 | -0.19 | N/A |
| VR131591 | Receptive | 0.72 | 0.74 | -0.22 | N/A |
| VR131622 | Receptive | 0.63 | 0.65 | 0.17 | N/A |
| VR131627 | Receptive | 0.54 | 0.55 | 0.20 | N/A |
| VR130469 | Receptive | 0.83 | 0.85 | 0.15 | N/A |
| VR130481 | Receptive | 0.73 | 0.74 | -0.17 | N/A |
| VR130500 | Receptive | 0.72 | 0.73 | 0.12 | N/A |
| VR060298 | Receptive | 0.69 | 0.71 | 0.27 | N/A |
| VR223852 | Receptive | 0.82 | 0.82 | -0.27 | N/A |
| VR140200 | Receptive | 0.86 | 0.86 | -0.23 | N/A |
| VR140214 | Receptive | 0.80 | 0.82 | 0.12 | N/A |

Table 8.E.4 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR150967 | Receptive | 0.56 | 0.57 | -0.14 | N/A |
| VR150970 | Receptive | 0.77 | 0.79 | 0.23 | N/A |
| VR144415 | Receptive | 0.78 | 0.81 | -0.06 | N/A |
| VR168684 | Receptive | 0.82 | 0.82 | -0.01 | N/A |
| VR053180 | Receptive | 0.72 | 0.74 | -0.20 | N/A |
| VR168670 | Receptive | 0.76 | 0.76 | 0.17 | N/A |
| VR155166 | Receptive | 0.70 | 0.70 | -0.25 | N/A |
| VR155150 | Receptive | 0.78 | 0.78 | 0.16 | N/A |
| VR151042 | Receptive | 0.37 | 0.38 | -0.19 | N/A |
| VR151060 | Receptive | 0.68 | 0.69 | 0.31 | N/A |
| VR151087 | Receptive | 0.40 | 0.40 | -0.31 | N/A |

Figure 8.E.5 for grade span six through eight presents the factor loading comparison plot chart that was created using the data in table 8.E.5, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y‑axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a very strong positive correlation between the two sets of factor loadings. All the data points cluster very closely around a diagonal reference line with slope of 1 and y-intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa. The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for grade span six through eight, with a slight favor for the one-factor model due to its slightly higher factor loadings.

Figure 8.E. Factor loading comparison for grade span six through eight

Table 8.E. Standardized Factor Loading for Grade Span Six Through Eight

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR196675 | Expressive | 0.71 | 0.69 | N/A | 0.23 |
| VR213047 | Expressive | 0.69 | 0.66 | N/A | 0.27 |
| VR147469 | Expressive | 0.68 | 0.65 | N/A | 0.22 |
| VR150190 | Expressive | 0.70 | 0.68 | N/A | 0.07 |
| VR150189 | Expressive | 0.88 | 0.89 | N/A | -0.06 |
| VR132159 | Expressive | 0.88 | 0.88 | N/A | 0.01 |
| VR132305 | Expressive | 0.89 | 0.90 | N/A | -0.05 |
| VR135915 | Expressive | 0.93 | 0.93 | N/A | -0.02 |
| VR166594 | Expressive | 0.62 | 0.64 | N/A | 0.36 |
| VR166612 | Expressive | 0.67 | 0.66 | N/A | 0.45 |
| VR133773 | Expressive | 0.64 | 0.65 | N/A | 0.33 |
| VR133797 | Expressive | 0.86 | 0.85 | N/A | 0.11 |
| VR133811 | Expressive | 0.68 | 0.70 | N/A | 0.41 |
| VR148864 | Expressive | 0.60 | 0.62 | N/A | 0.34 |
| VR148916 | Expressive | 0.56 | 0.56 | N/A | 0.04 |
| VR146758 | Expressive | 0.65 | 0.63 | N/A | 0.15 |
| VR150178 | Expressive | 0.88 | 0.88 | N/A | 0.04 |
| VR150180 | Expressive | 0.87 | 0.87 | N/A | 0.02 |
| VR167959 | Expressive | 0.59 | 0.61 | N/A | 0.42 |
| VR167974 | Expressive | 0.71 | 0.69 | N/A | 0.38 |
| VR148557 | Expressive | 0.57 | 0.57 | N/A | 0.08 |
| VR148552 | Expressive | 0.64 | 0.66 | N/A | 0.39 |
| VR196656 | Receptive | 0.84 | 0.84 | 0.07 | N/A |
| VR054031 | Receptive | 0.82 | 0.81 | 0.12 | N/A |
| VR147461 | Receptive | 0.81 | 0.83 | -0.15 | N/A |
| VR150186 | Receptive | 0.87 | 0.84 | 0.26 | N/A |
| VR150187 | Receptive | 0.84 | 0.84 | -0.02 | N/A |
| VR150188 | Receptive | 0.78 | 0.75 | 0.36 | N/A |
| VR131809 | Receptive | 0.87 | 0.87 | 0.23 | N/A |
| VR131713 | Receptive | 0.82 | 0.84 | -0.21 | N/A |
| VR132153 | Receptive | 0.66 | 0.66 | 0.36 | N/A |
| VR132196 | Receptive | 0.88 | 0.90 | -0.20 | N/A |
| VR132209 | Receptive | 0.88 | 0.88 | 0.21 | N/A |
| VR132300 | Receptive | 0.33 | 0.33 | 0.53 | N/A |
| VR166526 | Receptive | 0.83 | 0.82 | 0.22 | N/A |
| VR166576 | Receptive | 0.87 | 0.87 | -0.04 | N/A |

Table 8.E.5 *(continuation)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR133740 | Receptive | 0.85 | 0.86 | -0.14 | N/A |
| VR133759 | Receptive | 0.79 | 0.80 | -0.08 | N/A |
| VR148853 | Receptive | 0.82 | 0.83 | -0.21 | N/A |
| VR148842 | Receptive | 0.90 | 0.89 | 0.18 | N/A |
| VR148858 | Receptive | 0.82 | 0.82 | -0.08 | N/A |
| VR146736 | Receptive | 0.89 | 0.89 | 0.13 | N/A |
| VR150141 | Receptive | 0.86 | 0.86 | 0.18 | N/A |
| VR150176 | Receptive | 0.81 | 0.81 | 0.00 | N/A |
| VR150177 | Receptive | 0.86 | 0.84 | 0.23 | N/A |
| VR166709 | Receptive | 0.65 | 0.66 | 0.03 | N/A |
| VR167935 | Receptive | 0.74 | 0.74 | 0.21 | N/A |
| VR148512 | Receptive | 0.46 | 0.47 | -0.02 | N/A |
| VR148419 | Receptive | 0.73 | 0.73 | 0.34 | N/A |
| VR148515 | Receptive | 0.86 | 0.86 | -0.08 | N/A |

Figure 8.E.6 for high school (grade span nine through twelve) presents the factor loading comparison plot chart that was created using the data in table 8.E.6, which immediately follows. The graph’s x-axis shows factor loadings on the bifactor model from 0 to 1 in intervals of 0.1, and its y-axis shows factor loadings on the one-factor model from 0 to 1 in intervals of 0.1.

The graph shows that there is a very strong positive correlation between the two sets of factor loadings. All the data points cluster around a diagonal reference line with slope of 1 and y-intercept of 0. Items with high loadings on one model tend to have high loadings on the other, and vice-versa. Additionally, the one-factor model tends to result in slightly higher factor loadings than the bifactor model. The graph indicates that the two models allow for a similar interpretation of the dimensionality structure of the assessment for high school (grade span nine through twelve), with a slight favor for the one-factor model due to its slightly higher factor loadings.

Figure 8.E. Factor loading comparison for high school

Table 8.E. Standardized Factor Loading for Grade Span Nine Through Twelve

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR191381 | Expressive | 0.57 | 0.57 | N/A | 0.33 |
| VR191268 | Expressive | 0.62 | 0.63 | N/A | 0.37 |
| VR145310 | Expressive | 0.67 | 0.67 | N/A | 0.23 |
| VR148050 | Expressive | 0.84 | 0.86 | N/A | -0.04 |
| VR148051 | Expressive | 0.85 | 0.84 | N/A | 0.02 |
| VR132612 | Expressive | 0.71 | 0.75 | N/A | -0.05 |
| VR134023 | Expressive | 0.86 | 0.90 | N/A | -0.04 |
| VR133981 | Expressive | 0.84 | 0.85 | N/A | -0.03 |
| VR154926 | Expressive | 0.56 | 0.61 | N/A | 0.41 |
| VR154860 | Expressive | 0.68 | 0.70 | N/A | 0.39 |
| VR132823 | Expressive | 0.67 | 0.71 | N/A | 0.39 |
| VR132834 | Expressive | 0.70 | 0.75 | N/A | 0.40 |
| VR132848 | Expressive | 0.81 | 0.85 | N/A | 0.00 |
| VR150525 | Expressive | 0.49 | 0.54 | N/A | 0.39 |
| VR150530 | Expressive | 0.53 | 0.56 | N/A | -0.01 |
| VR144875 | Expressive | 0.71 | 0.69 | N/A | 0.27 |
| VR150131 | Expressive | 0.80 | 0.79 | N/A | 0.04 |
| VR150129 | Expressive | 0.84 | 0.87 | N/A | -0.03 |
| VR155088 | Expressive | 0.57 | 0.59 | N/A | 0.31 |
| VR155083 | Expressive | 0.68 | 0.69 | N/A | 0.38 |
| VR218864 | Expressive | 0.61 | 0.63 | N/A | 0.32 |
| VR056049 | Expressive | 0.70 | 0.72 | N/A | 0.04 |
| VR191356 | Receptive | 0.81 | 0.82 | 0.11 | N/A |
| VR191181 | Receptive | 0.84 | 0.85 | 0.07 | N/A |
| VR145306 | Receptive | 0.82 | 0.81 | 0.27 | N/A |
| VR148031 | Receptive | 0.73 | 0.76 | -0.17 | N/A |
| VR148029 | Receptive | 0.86 | 0.83 | 0.34 | N/A |
| VR147932 | Receptive | 0.86 | 0.86 | 0.20 | N/A |
| VR132587 | Receptive | 0.89 | 0.88 | 0.27 | N/A |
| VR155828 | Receptive | 0.90 | 0.90 | 0.20 | N/A |
| VR132598 | Receptive | 0.84 | 0.85 | 0.17 | N/A |
| VR133507 | Receptive | 0.78 | 0.77 | 0.27 | N/A |
| VR133887 | Receptive | 0.82 | 0.86 | -0.15 | N/A |
| VR133915 | Receptive | 0.86 | 0.85 | 0.32 | N/A |
| VR154631 | Receptive | 0.87 | 0.88 | 0.09 | N/A |
| VR154835 | Receptive | 0.65 | 0.71 | -0.32 | N/A |

Table 8.E.6 *(continuation one)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Item ID | Communication Mode | M1 | M2 General | M2 Receptive | M2 Expressive |
| VR132681 | Receptive | 0.79 | 0.79 | 0.11 | N/A |
| VR132695 | Receptive | 0.86 | 0.85 | 0.23 | N/A |
| VR150493 | Receptive | 0.80 | 0.84 | -0.21 | N/A |
| VR150497 | Receptive | 0.82 | 0.81 | 0.32 | N/A |
| VR150501 | Receptive | 0.70 | 0.77 | -0.28 | N/A |
| VR144835 | Receptive | 0.91 | 0.93 | -0.01 | N/A |
| VR150116 | Receptive | 0.91 | 0.90 | 0.23 | N/A |
| VR150120 | Receptive | 0.89 | 0.89 | 0.17 | N/A |
| VR150125 | Receptive | 0.78 | 0.82 | -0.10 | N/A |
| VR155066 | Receptive | 0.77 | 0.82 | -0.22 | N/A |
| VR155071 | Receptive | 0.85 | 0.84 | 0.25 | N/A |
| VR056029 | Receptive | 0.80 | 0.79 | 0.30 | N/A |
| VR224889 | Receptive | 0.67 | 0.69 | 0.13 | N/A |
| VR169370 | Receptive | 0.62 | 0.68 | -0.35 | N/A |

### Appendix 8.F: Reliability of Alternate ELPAC Performance and Decision Classification

**Note:** SEM is standard error of measurement. In table 8.F.1 through table 8.F.7, to protect student privacy, when the number of students in a student group is 10 or fewer, the reliability and SEM statistics are not reported and are replaced by “N/A.”

Table 8.F. Reliability Estimates by Student Group for Kindergarten

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 1,166 | 0.87 | 5.28 |
| Male | 845 | 0.87 | 5.31 |
| Female | 320 | 0.86 | 5.21 |
| Nonbinary | 1 | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A |
| Asian | 211 | 0.86 | 5.37 |
| Native Hawaiian or Other Pacific Islander | 4 | N/A | N/A |
| Filipino | 18 | 0.90 | 5.58 |
| Hispanic or Latino | 843 | 0.87 | 5.26 |
| Black or African American | 5 | N/A | N/A |
| White | 62 | 0.87 | 5.38 |
| Two or more races | 21 | 0.83 | 5.03 |
| Intellectual disability | 273 | 0.85 | 5.24 |
| Hearing impairment | 4 | N/A | N/A |
| Speech or language impairment | 56 | 0.81 | 4.84 |
| Visual impairment | 2 | N/A | N/A |
| Emotional impairment | 0 | N/A | N/A |
| Orthopedic impairment | 29 | 0.87 | 6.07 |
| Other health impairment | 47 | 0.85 | 4.98 |
| Specific learning disability | 4 | N/A | N/A |
| Deaf-blindness | 0 | N/A | N/A |
| Multiple disabilities | 48 | 0.82 | 6.39 |
| Autism | 702 | 0.87 | 5.23 |
| Traumatic brain injury | 1 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 340 | 0.87 | 5.48 |
| Economically disadvantaged | 826 | 0.87 | 5.20 |
| In US schools less than 12 months | 932 | 0.86 | 5.32 |
| In US schools 12 months or more | 209 | 0.87 | 5.11 |
| Duration unknown | 25 | 0.85 | 5.02 |
| Migrant education | 6 | N/A | N/A |
| Not migrant education | 1,160 | 0.87 | 5.29 |

Table 8.F. Reliability Estimates by Student Group for Grade One

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 1,163 | 0.86 | 5.17 |
| Male | 818 | 0.86 | 5.17 |
| Female | 345 | 0.86 | 5.17 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A |
| Asian | 186 | 0.87 | 5.27 |
| Native Hawaiian or Other Pacific Islander | 3 | N/A | N/A |
| Filipino | 18 | 0.86 | 5.49 |
| Hispanic or Latino | 857 | 0.86 | 5.13 |
| Black or African American | 13 | 0.87 | 5.54 |
| White | 66 | 0.86 | 5.20 |
| Two or more races | 18 | 0.85 | 5.49 |
| Intellectual disability | 342 | 0.84 | 5.06 |
| Hearing impairment | 7 | N/A | N/A |
| Speech or language impairment | 38 | 0.75 | 5.13 |
| Visual impairment | 2 | N/A | N/A |
| Emotional impairment | 2 | N/A | N/A |
| Orthopedic impairment | 24 | 0.88 | 5.97 |
| Other health impairment | 56 | 0.84 | 4.76 |
| Specific learning disability | 3 | N/A | N/A |
| Deaf-blindness | 1 | N/A | N/A |
| Multiple disabilities | 43 | 0.85 | 6.14 |
| Autism | 643 | 0.86 | 5.18 |
| Traumatic brain injury | 2 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 308 | 0.86 | 5.10 |
| Economically disadvantaged | 855 | 0.86 | 5.20 |
| In US schools less than 12 months | 56 | 0.85 | 5.39 |
| In US schools 12 months or more | 1,087 | 0.86 | 5.16 |
| Duration unknown | 20 | 0.88 | 5.16 |
| Migrant education | 11 | 0.88 | 4.94 |
| Not migrant education | 1,152 | 0.86 | 5.17 |

Table 8.F. Reliability Estimates by Student Group for Grade Two

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 1,131 | 0.87 | 5.50 |
| Male | 803 | 0.88 | 5.60 |
| Female | 328 | 0.86 | 5.22 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 2 | N/A | N/A |
| Asian | 175 | 0.88 | 5.74 |
| Native Hawaiian or Other Pacific Islander | 8 | N/A | N/A |
| Filipino | 18 | 0.88 | 5.52 |
| Hispanic or Latino | 868 | 0.87 | 5.40 |
| Black or African American | 8 | N/A | N/A |
| White | 39 | 0.90 | 6.59 |
| Two or more races | 13 | 0.62 | 4.28 |
| Intellectual disability | 388 | 0.86 | 5.14 |
| Hearing impairment | 10 | N/A | N/A |
| Speech or language impairment | 38 | 0.80 | 5.47 |
| Visual impairment | 1 | N/A | N/A |
| Emotional impairment | 0 | N/A | N/A |
| Orthopedic impairment | 22 | 0.88 | 5.57 |
| Other health impairment | 52 | 0.87 | 5.20 |
| Specific learning disability | 5 | N/A | N/A |
| Deaf-blindness | 0 | N/A | N/A |
| Multiple disabilities | 54 | 0.87 | 5.85 |
| Autism | 560 | 0.88 | 5.74 |
| Traumatic brain injury | 1 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 298 | 0.88 | 5.52 |
| Economically disadvantaged | 833 | 0.87 | 5.49 |
| In US schools less than 12 months | 11 | 0.89 | 6.51 |
| In US schools 12 months or more | 1,111 | 0.87 | 5.49 |
| Duration unknown | 9 | N/A | N/A |
| Migrant education | 8 | N/A | N/A |
| Not migrant education | 1,123 | 0.87 | 5.50 |

Table 8.F. Reliability Estimates by Student Group for Grade Span Three Through Five

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 3,936 | 0.88 | 4.59 |
| Male | 2,706 | 0.88 | 4.61 |
| Female | 1,230 | 0.88 | 4.55 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 7 | N/A | N/A |
| Asian | 560 | 0.87 | 4.34 |
| Native Hawaiian or Other Pacific Islander | 11 | 0.88 | 4.68 |
| Filipino | 87 | 0.88 | 4.38 |
| Hispanic or Latino | 3,032 | 0.88 | 4.64 |
| Black or African American | 29 | 0.88 | 4.32 |
| White | 154 | 0.88 | 4.80 |
| Two or more races | 56 | 0.88 | 4.70 |
| Intellectual disability | 1,577 | 0.87 | 4.45 |
| Hearing impairment | 22 | 0.86 | 4.49 |
| Speech or language impairment | 75 | 0.72 | 5.32 |
| Visual impairment | 3 | N/A | N/A |
| Emotional impairment | 5 | N/A | N/A |
| Orthopedic impairment | 89 | 0.89 | 5.09 |
| Other health impairment | 189 | 0.84 | 4.92 |
| Specific learning disability | 133 | 0.42 | 5.98 |
| Deaf-blindness | 0 | N/A | N/A |
| Multiple disabilities | 147 | 0.87 | 4.62 |
| Autism | 1,691 | 0.87 | 4.48 |
| Traumatic brain injury | 5 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 970 | 0.87 | 4.44 |
| Economically disadvantaged | 2,966 | 0.87 | 4.64 |
| In US schools less than 12 months | 38 | 0.88 | 4.44 |
| In US schools 12 months or more | 3,868 | 0.87 | 4.60 |
| Duration unknown | 30 | 0.87 | 4.29 |
| Migrant education | 48 | 0.88 | 4.39 |
| Not migrant education | 3,888 | 0.88 | 4.60 |

Table 8.F. Reliability Estimates by Student Group for Grade Span Six Through Eight

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 3,070 | 0.86 | 5.30 |
| Male | 2,000 | 0.87 | 5.30 |
| Female | 1,070 | 0.86 | 5.28 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 7 | N/A | N/A |
| Asian | 433 | 0.87 | 5.18 |
| Native Hawaiian or Other Pacific Islander | 8 | N/A | N/A |
| Filipino | 55 | 0.82 | 4.91 |
| Hispanic or Latino | 2,382 | 0.86 | 5.33 |
| Black or African American | 22 | 0.88 | 5.68 |
| White | 133 | 0.87 | 5.14 |
| Two or more races | 30 | 0.89 | 5.61 |
| Intellectual disability | 1,502 | 0.86 | 5.07 |
| Hearing impairment | 17 | 0.77 | 5.25 |
| Speech or language impairment | 32 | 0.52 | 6.02 |
| Visual impairment | 7 | N/A | N/A |
| Emotional impairment | 7 | N/A | N/A |
| Orthopedic impairment | 67 | 0.89 | 5.41 |
| Other health impairment | 132 | 0.81 | 6.19 |
| Specific learning disability | 156 | 0.29 | 7.06 |
| Deaf-blindness | 2 | N/A | N/A |
| Multiple disabilities | 141 | 0.90 | 5.63 |
| Autism | 1,003 | 0.86 | 5.08 |
| Traumatic brain injury | 4 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 697 | 0.87 | 5.15 |
| Economically disadvantaged | 2,373 | 0.86 | 5.34 |
| In US schools less than 12 months | 25 | 0.91 | 4.75 |
| In US schools 12 months or more | 3,027 | 0.86 | 5.30 |
| Duration unknown | 18 | 0.87 | 6.11 |
| Migrant education | 54 | 0.84 | 5.74 |
| Not migrant education | 3,016 | 0.87 | 5.29 |

Table 8.F. Reliability Estimates by Student Group for Grade Span Nine and Ten

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 1,514 | 0.87 | 4.51 |
| Male | 979 | 0.87 | 4.58 |
| Female | 535 | 0.87 | 4.38 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 1 | N/A | N/A |
| Asian | 209 | 0.87 | 4.20 |
| Native Hawaiian or Other Pacific Islander | 7 | N/A | N/A |
| Filipino | 30 | 0.87 | 4.23 |
| Hispanic or Latino | 1,170 | 0.87 | 4.54 |
| Black or African American | 8 | N/A | N/A |
| White | 69 | 0.88 | 4.97 |
| Two or more races | 20 | 0.91 | 4.81 |
| Intellectual disability | 805 | 0.85 | 4.34 |
| Hearing impairment | 11 | 0.66 | 3.99 |
| Speech or language impairment | 16 | 0.57 | 5.16 |
| Visual impairment | 7 | N/A | N/A |
| Emotional impairment | 12 | 0.21 | 5.88 |
| Orthopedic impairment | 31 | 0.91 | 5.13 |
| Other health impairment | 43 | 0.83 | 5.00 |
| Specific learning disability | 62 | 0.47 | 5.69 |
| Deaf-blindness | 0 | N/A | N/A |
| Multiple disabilities | 99 | 0.89 | 4.69 |
| Autism | 422 | 0.86 | 4.37 |
| Traumatic brain injury | 6 | N/A | N/A |
| Not classified | 0 | N/A | N/A |
| Not economically disadvantaged | 376 | 0.88 | 4.36 |
| Economically disadvantaged | 1,138 | 0.86 | 4.56 |
| In US schools less than 12 months | 9 | N/A | N/A |
| In US schools 12 months or more | 1,499 | 0.87 | 4.51 |
| Duration unknown | 6 | N/A | N/A |
| Migrant education | 10 | N/A | N/A |
| Not migrant education | 1,504 | 0.87 | 4.52 |

Table 8.F. Reliability Estimates by Student Group for Grade Span Eleven and Twelve

|  |  |  |  |
| --- | --- | --- | --- |
| Student Group | N | Reliability | SEM |
| All | 2,761 | 0.87 | 4.62 |
| Male | 1,787 | 0.87 | 4.65 |
| Female | 974 | 0.88 | 4.57 |
| Nonbinary | 0 | N/A | N/A |
| American Indian or Alaska Native | 3 | N/A | N/A |
| Asian | 401 | 0.88 | 4.55 |
| Native Hawaiian or Other Pacific Islander | 13 | 0.85 | 4.80 |
| Filipino | 51 | 0.88 | 4.83 |
| Hispanic or Latino | 2,144 | 0.87 | 4.64 |
| Black or African American | 23 | 0.89 | 4.54 |
| White | 111 | 0.88 | 4.44 |
| Two or more races | 15 | 0.89 | 4.58 |
| Intellectual disability | 1,466 | 0.87 | 4.45 |
| Hearing impairment | 28 | 0.81 | 4.37 |
| Speech or language impairment | 10 | N/A | N/A |
| Visual impairment | 14 | 0.91 | 5.74 |
| Emotional impairment | 16 | 0.76 | 5.83 |
| Orthopedic impairment | 98 | 0.89 | 4.75 |
| Other health impairment | 72 | 0.78 | 4.83 |
| Specific learning disability | 78 | 0.39 | 6.06 |
| Deaf-blindness | 3 | N/A | N/A |
| Multiple disabilities | 133 | 0.90 | 4.60 |
| Autism | 824 | 0.88 | 4.66 |
| Traumatic brain injury | 18 | 0.82 | 4.95 |
| Not classified | 1 | N/A | N/A |
| Not economically disadvantaged | 769 | 0.88 | 4.55 |
| Economically disadvantaged | 1,992 | 0.87 | 4.65 |
| In US schools less than 12 months | 5 | N/A | N/A |
| In US schools 12 months or more | 2,732 | 0.87 | 4.62 |
| Duration unknown | 24 | 0.88 | 4.24 |
| Migrant education | 28 | 0.64 | 4.61 |
| Not migrant education | 2,733 | 0.87 | 4.62 |

Table 8.F. 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 | 0.88 | 0.84 | 0.91 | 0.88 | 0.78 | 0.72 |
| 1 | 0.87 | 0.83 | 0.90 | 0.87 | 0.77 | 0.70 |
| 2 | 0.88 | 0.84 | 0.91 | 0.88 | 0.79 | 0.72 |
| 3–5 | 0.90 | 0.86 | 0.89 | 0.85 | 0.79 | 0.71 |
| 6–8 | 0.90 | 0.87 | 0.88 | 0.85 | 0.78 | 0.71 |
| 9–10 | 0.89 | 0.85 | 0.89 | 0.86 | 0.78 | 0.71 |
| 11–12 | 0.89 | 0.86 | 0.89 | 0.86 | 0.78 | 0.72 |

### Appendix 8.G: Raw-to-Scale-Score Conversions

Table 8.G. Raw-to-Scale-Score Conversion Table for Kindergarten

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 201 | 19 | Level 1 |
| 1 | 214 | 11 | Level 1 |
| 2 | 222 | 8 | Level 1 |
| 3 | 227 | 7 | Level 1 |
| 4 | 230 | 6 | Level 1 |
| 5 | 233 | 5 | Level 1 |
| 6 | 235 | 5 | Level 1 |
| 7 | 237 | 5 | Level 1 |
| 8 | 239 | 4 | Level 1 |
| 9 | 241 | 4 | Level 1 |
| 10 | 242 | 4 | Level 1 |
| 11 | 244 | 4 | Level 2 |
| 12 | 245 | 4 | Level 2 |
| 13 | 247 | 4 | Level 2 |
| 14 | 248 | 4 | Level 2 |
| 15 | 249 | 4 | Level 2 |
| 16 | 251 | 4 | Level 2 |
| 17 | 252 | 4 | Level 2 |
| 18 | 253 | 4 | Level 2 |
| 19 | 255 | 4 | Level 2 |
| 20 | 256 | 4 | Level 2 |
| 21 | 258 | 4 | Level 2 |
| 22 | 260 | 4 | Level 3 |
| 23 | 261 | 4 | Level 3 |
| 24 | 263 | 5 | Level 3 |
| 25 | 265 | 5 | Level 3 |
| 26 | 268 | 5 | Level 3 |
| 27 | 271 | 6 | Level 3 |
| 28 | 274 | 7 | Level 3 |
| 29 | 279 | 8 | Level 3 |
| 30 | 287 | 11 | Level 3 |
| 31 | 299 | 18 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 301 | 17 | Level 1 |
| 1 | 311 | 11 | Level 1 |
| 2 | 319 | 8 | Level 1 |
| 3 | 324 | 7 | Level 1 |
| 4 | 328 | 6 | Level 1 |
| 5 | 331 | 5 | Level 1 |
| 6 | 334 | 5 | Level 1 |
| 7 | 336 | 5 | Level 1 |
| 8 | 338 | 5 | Level 1 |
| 9 | 340 | 4 | Level 1 |
| 10 | 341 | 4 | Level 1 |
| 11 | 343 | 4 | Level 1 |
| 12 | 345 | 4 | Level 2 |
| 13 | 346 | 4 | Level 2 |
| 14 | 347 | 4 | Level 2 |
| 15 | 349 | 4 | Level 2 |
| 16 | 350 | 4 | Level 2 |
| 17 | 352 | 4 | Level 2 |
| 18 | 353 | 4 | Level 2 |
| 19 | 355 | 4 | Level 2 |
| 20 | 356 | 4 | Level 2 |
| 21 | 358 | 4 | Level 2 |
| 22 | 359 | 4 | Level 2 |
| 23 | 361 | 5 | Level 3 |
| 24 | 363 | 5 | Level 3 |
| 25 | 365 | 5 | Level 3 |
| 26 | 368 | 6 | Level 3 |
| 27 | 372 | 7 | Level 3 |
| 28 | 376 | 8 | Level 3 |
| 29 | 384 | 11 | Level 3 |
| 30 | 399 | 22 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 401 | 16 | Level 1 |
| 1 | 409 | 11 | Level 1 |
| 2 | 417 | 8 | Level 1 |
| 3 | 422 | 7 | Level 1 |
| 4 | 425 | 6 | Level 1 |
| 5 | 428 | 5 | Level 1 |
| 6 | 430 | 5 | Level 1 |
| 7 | 432 | 5 | Level 1 |
| 8 | 434 | 5 | Level 1 |
| 9 | 436 | 4 | Level 1 |
| 10 | 438 | 4 | Level 1 |
| 11 | 439 | 4 | Level 1 |
| 12 | 441 | 4 | Level 1 |
| 13 | 442 | 4 | Level 1 |
| 14 | 444 | 4 | Level 2 |
| 15 | 445 | 4 | Level 2 |
| 16 | 447 | 4 | Level 2 |
| 17 | 448 | 4 | Level 2 |
| 18 | 450 | 4 | Level 2 |
| 19 | 451 | 4 | Level 2 |
| 20 | 453 | 4 | Level 2 |
| 21 | 455 | 4 | Level 2 |
| 22 | 457 | 5 | Level 2 |
| 23 | 459 | 5 | Level 2 |
| 24 | 461 | 5 | Level 3 |
| 25 | 463 | 5 | Level 3 |
| 26 | 466 | 6 | Level 3 |
| 27 | 470 | 7 | Level 3 |
| 28 | 475 | 8 | Level 3 |
| 29 | 484 | 12 | Level 3 |
| 30 | 499 | 22 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Three Through Five Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 501 | 20 | Level 1 |
| 1 | 515 | 10 | Level 1 |
| 2 | 522 | 7 | Level 1 |
| 3 | 526 | 6 | Level 1 |
| 4 | 529 | 5 | Level 1 |
| 5 | 532 | 4 | Level 1 |
| 6 | 534 | 4 | Level 1 |
| 7 | 536 | 4 | Level 1 |
| 8 | 537 | 4 | Level 1 |
| 9 | 539 | 4 | Level 1 |
| 10 | 540 | 3 | Level 1 |
| 11 | 541 | 3 | Level 1 |
| 12 | 543 | 3 | Level 1 |
| 13 | 544 | 3 | Level 2 |
| 14 | 545 | 3 | Level 2 |
| 15 | 546 | 3 | Level 2 |
| 16 | 547 | 3 | Level 2 |
| 17 | 548 | 3 | Level 2 |
| 18 | 549 | 3 | Level 2 |
| 19 | 551 | 3 | Level 2 |
| 20 | 552 | 3 | Level 2 |
| 21 | 553 | 3 | Level 2 |
| 22 | 555 | 4 | Level 2 |
| 23 | 556 | 4 | Level 2 |
| 24 | 557 | 4 | Level 2 |
| 25 | 559 | 4 | Level 2 |
| 26 | 561 | 4 | Level 3 |
| 27 | 563 | 5 | Level 3 |
| 28 | 566 | 5 | Level 3 |
| 29 | 570 | 7 | Level 3 |
| 30 | 576 | 9 | Level 3 |
| 31 | 599 | 31 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Three Through Five Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 501 | 19 | Level 1 |
| 1 | 515 | 9 | Level 1 |
| 2 | 522 | 7 | Level 1 |
| 3 | 526 | 6 | Level 1 |
| 4 | 529 | 5 | Level 1 |
| 5 | 532 | 4 | Level 1 |
| 6 | 534 | 4 | Level 1 |
| 7 | 535 | 4 | Level 1 |
| 8 | 537 | 4 | Level 1 |
| 9 | 539 | 4 | Level 1 |
| 10 | 540 | 4 | Level 1 |
| 11 | 542 | 4 | Level 1 |
| 12 | 543 | 3 | Level 1 |
| 13 | 544 | 3 | Level 2 |
| 14 | 545 | 3 | Level 2 |
| 15 | 547 | 3 | Level 2 |
| 16 | 548 | 3 | Level 2 |
| 17 | 549 | 3 | Level 2 |
| 18 | 551 | 3 | Level 2 |
| 19 | 552 | 3 | Level 2 |
| 20 | 553 | 4 | Level 2 |
| 21 | 555 | 4 | Level 2 |
| 22 | 556 | 4 | Level 2 |
| 23 | 558 | 4 | Level 2 |
| 24 | 559 | 4 | Level 2 |
| 25 | 561 | 4 | Level 3 |
| 26 | 563 | 4 | Level 3 |
| 27 | 565 | 5 | Level 3 |
| 28 | 568 | 5 | Level 3 |
| 29 | 572 | 7 | Level 3 |
| 30 | 579 | 9 | Level 3 |
| 31 | 599 | 27 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Six Through Eight Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 601 | 17 | Level 1 |
| 1 | 613 | 9 | Level 1 |
| 2 | 620 | 7 | Level 1 |
| 3 | 624 | 6 | Level 1 |
| 4 | 627 | 5 | Level 1 |
| 5 | 629 | 5 | Level 1 |
| 6 | 631 | 4 | Level 1 |
| 7 | 633 | 4 | Level 1 |
| 8 | 635 | 4 | Level 1 |
| 9 | 636 | 4 | Level 1 |
| 10 | 638 | 4 | Level 1 |
| 11 | 639 | 3 | Level 1 |
| 12 | 640 | 3 | Level 1 |
| 13 | 642 | 3 | Level 1 |
| 14 | 643 | 3 | Level 1 |
| 15 | 644 | 3 | Level 2 |
| 16 | 645 | 3 | Level 2 |
| 17 | 647 | 3 | Level 2 |
| 18 | 648 | 3 | Level 2 |
| 19 | 649 | 3 | Level 2 |
| 20 | 651 | 4 | Level 2 |
| 21 | 652 | 4 | Level 2 |
| 22 | 654 | 4 | Level 2 |
| 23 | 655 | 4 | Level 2 |
| 24 | 657 | 4 | Level 2 |
| 25 | 659 | 4 | Level 2 |
| 26 | 661 | 5 | Level 3 |
| 27 | 664 | 5 | Level 3 |
| 28 | 668 | 7 | Level 3 |
| 29 | 675 | 9 | Level 3 |
| 30 | 699 | 34 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Six Through Eight Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 601 | 17 | Level 1 |
| 1 | 613 | 9 | Level 1 |
| 2 | 620 | 7 | Level 1 |
| 3 | 624 | 6 | Level 1 |
| 4 | 627 | 5 | Level 1 |
| 5 | 630 | 4 | Level 1 |
| 6 | 632 | 4 | Level 1 |
| 7 | 634 | 4 | Level 1 |
| 8 | 635 | 4 | Level 1 |
| 9 | 637 | 4 | Level 1 |
| 10 | 638 | 4 | Level 1 |
| 11 | 640 | 4 | Level 1 |
| 12 | 641 | 4 | Level 1 |
| 13 | 642 | 4 | Level 1 |
| 14 | 644 | 4 | Level 2 |
| 15 | 645 | 4 | Level 2 |
| 16 | 647 | 4 | Level 2 |
| 17 | 648 | 4 | Level 2 |
| 18 | 649 | 4 | Level 2 |
| 19 | 651 | 4 | Level 2 |
| 20 | 652 | 4 | Level 2 |
| 21 | 654 | 4 | Level 2 |
| 22 | 656 | 4 | Level 2 |
| 23 | 657 | 4 | Level 2 |
| 24 | 659 | 4 | Level 2 |
| 25 | 661 | 4 | Level 3 |
| 26 | 664 | 5 | Level 3 |
| 27 | 667 | 6 | Level 3 |
| 28 | 671 | 7 | Level 3 |
| 29 | 677 | 9 | Level 3 |
| 30 | 699 | 30 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Nine and Ten Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 701 | 18 | Level 1 |
| 1 | 714 | 9 | Level 1 |
| 2 | 721 | 6 | Level 1 |
| 3 | 725 | 5 | Level 1 |
| 4 | 728 | 4 | Level 1 |
| 5 | 730 | 4 | Level 1 |
| 6 | 732 | 4 | Level 1 |
| 7 | 734 | 4 | Level 1 |
| 8 | 735 | 3 | Level 1 |
| 9 | 737 | 3 | Level 1 |
| 10 | 738 | 3 | Level 1 |
| 11 | 739 | 3 | Level 1 |
| 12 | 740 | 3 | Level 1 |
| 13 | 741 | 3 | Level 1 |
| 14 | 743 | 3 | Level 1 |
| 15 | 744 | 3 | Level 2 |
| 16 | 745 | 3 | Level 2 |
| 17 | 746 | 3 | Level 2 |
| 18 | 747 | 3 | Level 2 |
| 19 | 748 | 3 | Level 2 |
| 20 | 750 | 3 | Level 2 |
| 21 | 751 | 3 | Level 2 |
| 22 | 752 | 3 | Level 2 |
| 23 | 754 | 4 | Level 2 |
| 24 | 755 | 4 | Level 2 |
| 25 | 757 | 4 | Level 2 |
| 26 | 759 | 4 | Level 2 |
| 27 | 762 | 5 | Level 3 |
| 28 | 766 | 6 | Level 3 |
| 29 | 771 | 8 | Level 3 |
| 30 | 799 | 44 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Nine and Ten Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 701 | 17 | Level 1 |
| 1 | 713 | 9 | Level 1 |
| 2 | 720 | 6 | Level 1 |
| 3 | 724 | 5 | Level 1 |
| 4 | 727 | 4 | Level 1 |
| 5 | 729 | 4 | Level 1 |
| 6 | 731 | 4 | Level 1 |
| 7 | 733 | 4 | Level 1 |
| 8 | 734 | 3 | Level 1 |
| 9 | 736 | 3 | Level 1 |
| 10 | 737 | 3 | Level 1 |
| 11 | 738 | 3 | Level 1 |
| 12 | 739 | 3 | Level 1 |
| 13 | 740 | 3 | Level 1 |
| 14 | 742 | 3 | Level 1 |
| 15 | 743 | 3 | Level 1 |
| 16 | 744 | 3 | Level 2 |
| 17 | 745 | 3 | Level 2 |
| 18 | 746 | 3 | Level 2 |
| 19 | 747 | 3 | Level 2 |
| 20 | 748 | 3 | Level 2 |
| 21 | 750 | 3 | Level 2 |
| 22 | 751 | 3 | Level 2 |
| 23 | 752 | 3 | Level 2 |
| 24 | 754 | 4 | Level 2 |
| 25 | 756 | 4 | Level 2 |
| 26 | 758 | 4 | Level 2 |
| 27 | 760 | 5 | Level 3 |
| 28 | 764 | 6 | Level 3 |
| 29 | 769 | 8 | Level 3 |
| 30 | 799 | 51 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Eleven and Twelve Form One

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 801 | 18 | Level 1 |
| 1 | 814 | 9 | Level 1 |
| 2 | 821 | 6 | Level 1 |
| 3 | 825 | 5 | Level 1 |
| 4 | 828 | 4 | Level 1 |
| 5 | 830 | 4 | Level 1 |
| 6 | 832 | 4 | Level 1 |
| 7 | 834 | 4 | Level 1 |
| 8 | 835 | 3 | Level 1 |
| 9 | 837 | 3 | Level 1 |
| 10 | 838 | 3 | Level 1 |
| 11 | 839 | 3 | Level 1 |
| 12 | 840 | 3 | Level 1 |
| 13 | 841 | 3 | Level 1 |
| 14 | 843 | 3 | Level 1 |
| 15 | 844 | 3 | Level 2 |
| 16 | 845 | 3 | Level 2 |
| 17 | 846 | 3 | Level 2 |
| 18 | 847 | 3 | Level 2 |
| 19 | 848 | 3 | Level 2 |
| 20 | 850 | 3 | Level 2 |
| 21 | 851 | 3 | Level 2 |
| 22 | 852 | 3 | Level 2 |
| 23 | 854 | 4 | Level 2 |
| 24 | 855 | 4 | Level 2 |
| 25 | 857 | 4 | Level 2 |
| 26 | 859 | 4 | Level 2 |
| 27 | 862 | 5 | Level 3 |
| 28 | 866 | 6 | Level 3 |
| 29 | 871 | 8 | Level 3 |
| 30 | 899 | 44 | Level 3 |

Table 8.G. Raw-to-Scale-Score Conversion Table for Grade Span Eleven and Twelve Form Two

|  |  |  |  |
| --- | --- | --- | --- |
| Raw Score | Scale Score | CSEM | Level |
| 0 | 801 | 17 | Level 1 |
| 1 | 813 | 9 | Level 1 |
| 2 | 820 | 6 | Level 1 |
| 3 | 824 | 5 | Level 1 |
| 4 | 827 | 4 | Level 1 |
| 5 | 829 | 4 | Level 1 |
| 6 | 831 | 4 | Level 1 |
| 7 | 833 | 4 | Level 1 |
| 8 | 834 | 3 | Level 1 |
| 9 | 836 | 3 | Level 1 |
| 10 | 837 | 3 | Level 1 |
| 11 | 838 | 3 | Level 1 |
| 12 | 839 | 3 | Level 1 |
| 13 | 840 | 3 | Level 1 |
| 14 | 842 | 3 | Level 1 |
| 15 | 843 | 3 | Level 1 |
| 16 | 844 | 3 | Level 2 |
| 17 | 845 | 3 | Level 2 |
| 18 | 846 | 3 | Level 2 |
| 19 | 847 | 3 | Level 2 |
| 20 | 848 | 3 | Level 2 |
| 21 | 850 | 3 | Level 2 |
| 22 | 851 | 3 | Level 2 |
| 23 | 852 | 3 | Level 2 |
| 24 | 854 | 4 | Level 2 |
| 25 | 856 | 4 | Level 2 |
| 26 | 858 | 4 | Level 2 |
| 27 | 860 | 5 | Level 3 |
| 28 | 864 | 6 | Level 3 |
| 29 | 869 | 8 | Level 3 |
| 30 | 899 | 51 | Level 3 |

### Appendix 8.H: Interrater Reliability

Table 8.H. Interrater Reliability for Kindergarten

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR131419 | 2 | 201 | 0.83 | 0.84 | 90 | 5 | 4 |
| VR131711 | 2 | 200 | 0.87 | 0.87 | 92 | 4 | 4 |
| VR139022 | 2 | 210 | 0.86 | 0.87 | 91 | 4 | 4 |
| VR139973 | 2 | 223 | 0.79 | 0.82 | 88 | 6 | 6 |
| VR139975 | 2 | 192 | 0.85 | 0.81 | 91 | 2 | 7 |
| VR154458 | 2 | 196 | 0.85 | 0.90 | 91 | 7 | 2 |
| VR154465 | 2 | 196 | 0.83 | 0.88 | 91 | 6 | 3 |
| VR193113 | 2 | 233 | 0.83 | 0.84 | 90 | 5 | 5 |
| VR215978 | 2 | 194 | 0.85 | 0.85 | 91 | 5 | 5 |
| VR216450 | 2 | 233 | 0.77 | 0.78 | 88 | 4 | 8 |
| VR223164 | 2 | 194 | 0.85 | 0.84 | 91 | 5 | 4 |
| **AVERAGE:** | **N/A** | **207** | **0.84** | **0.85** | **90** | **5** | **5** |

Table 8.H. Interrater Reliability for Grade One

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR053970 | 2 | 221 | 0.81 | 0.84 | 90 | 5 | 5 |
| VR130343 | 2 | 203 | 0.81 | 0.87 | 88 | 9 | 2 |
| VR130428 | 2 | 203 | 0.84 | 0.86 | 90 | 6 | 4 |
| VR133983 | 2 | 199 | 0.88 | 0.90 | 92 | 5 | 3 |
| VR134003 | 2 | 202 | 0.82 | 0.85 | 89 | 7 | 4 |
| VR137618 | 2 | 218 | 0.82 | 0.86 | 89 | 7 | 4 |
| VR150707 | 2 | 199 | 0.94 | 0.95 | 96 | 3 | 2 |
| VR154753 | 2 | 203 | 0.83 | 0.86 | 89 | 7 | 3 |
| VR154755 | 2 | 202 | 0.85 | 0.85 | 90 | 7 | 3 |
| VR193651 | 2 | 221 | 0.74 | 0.81 | 87 | 8 | 5 |
| **AVERAGE:** | **N/A** | **207** | **0.83** | **0.87** | **90** | **6** | **3** |

Table 8.H. Interrater Reliability for Grade Two

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR130298 | 2 | 199 | 0.84 | 0.90 | 89 | 9 | 2 |
| VR134664 | 2 | 193 | 0.77 | 0.84 | 85 | 11 | 4 |
| VR134677 | 2 | 192 | 0.84 | 0.90 | 90 | 8 | 2 |
| VR140209 | 2 | 214 | 0.77 | 0.81 | 89 | 5 | 6 |
| VR151624 | 2 | 190 | 0.87 | 0.90 | 92 | 6 | 3 |
| VR155670 | 2 | 196 | 0.82 | 0.85 | 88 | 10 | 2 |
| VR193828 | 2 | 214 | 0.82 | 0.82 | 90 | 5 | 5 |
| VR193885 | 2 | 215 | 0.81 | 0.79 | 92 | 2 | 6 |
| VR223063 | 2 | 195 | 0.85 | 0.86 | 90 | 8 | 2 |
| **AVERAGE:** | **N/A** | **201** | **0.82** | **0.85** | **89** | **7** | **3** |

Table 8.H. Interrater Reliability for Grade Span Three Through Five

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR053990 | 2 | 831 | 0.81 | 0.83 | 91 | 4 | 4 |
| VR140236 | 2 | 780 | 0.87 | 0.90 | 92 | 6 | 3 |
| VR140266 | 2 | 779 | 0.85 | 0.88 | 91 | 6 | 4 |
| VR144428 | 2 | 438 | 0.82 | 0.85 | 92 | 4 | 5 |
| VR144686 | 2 | 822 | 0.83 | 0.86 | 92 | 4 | 4 |
| VR150995 | 2 | 368 | 0.89 | 0.92 | 94 | 3 | 3 |
| VR151097 | 2 | 405 | 0.81 | 0.84 | 88 | 8 | 4 |
| VR155154 | 2 | 418 | 0.82 | 0.81 | 89 | 8 | 3 |
| VR155163 | 2 | 418 | 0.86 | 0.88 | 91 | 7 | 2 |
| VR194301 | 2 | 393 | 0.85 | 0.89 | 92 | 5 | 3 |
| VR210937 | 2 | 788 | 0.85 | 0.87 | 91 | 6 | 3 |
| VR222572 | 2 | 370 | 0.86 | 0.85 | 92 | 5 | 3 |
| VR222573 | 2 | 369 | 0.90 | 0.94 | 94 | 5 | 1 |
| **AVERAGE:** | **N/A** | **552** | **0.85** | **0.87** | **91** | **5** | **3** |

Table 8.H. Interrater Reliability for Grade Span Six Through Eight

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR133773 | 2 | 653 | 0.84 | 0.88 | 91 | 6 | 3 |
| VR133811 | 2 | 647 | 0.81 | 0.85 | 89 | 7 | 4 |
| VR146758 | 2 | 389 | 0.74 | 0.78 | 89 | 4 | 7 |
| VR147469 | 2 | 687 | 0.75 | 0.76 | 91 | 3 | 6 |
| VR148552 | 2 | 370 | 0.80 | 0.85 | 87 | 9 | 4 |
| VR148864 | 2 | 278 | 0.89 | 0.92 | 94 | 3 | 3 |
| VR166594 | 2 | 281 | 0.84 | 0.83 | 90 | 4 | 6 |
| VR166612 | 2 | 279 | 0.86 | 0.85 | 92 | 5 | 3 |
| VR167959 | 2 | 375 | 0.79 | 0.82 | 86 | 10 | 4 |
| VR167974 | 2 | 376 | 0.82 | 0.84 | 89 | 8 | 2 |
| VR196675 | 2 | 691 | 0.74 | 0.77 | 91 | 4 | 5 |
| VR213047 | 2 | 301 | 0.75 | 0.76 | 93 | 2 | 5 |
| **AVERAGE:** | **N/A** | **444** | **0.80** | **0.83** | **90** | **6** | **4** |

Table 8.H. Interrater Reliability for Grade Span Nine Through Twelve

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Item ID | Maximum Points | Number of Responses | Kappa | QWK | Percent Exact | Percent Adjacent | Percent Discrepant |
| VR132823 | 2 | 790 | 0.80 | 0.85 | 89 | 7 | 4 |
| VR132834 | 2 | 785 | 0.79 | 0.84 | 89 | 6 | 5 |
| VR144875 | 2 | 494 | 0.66 | 0.72 | 88 | 6 | 6 |
| VR145310 | 2 | 836 | 0.74 | 0.78 | 87 | 7 | 6 |
| VR150525 | 2 | 320 | 0.81 | 0.82 | 88 | 8 | 5 |
| VR154860 | 2 | 325 | 0.78 | 0.77 | 88 | 8 | 4 |
| VR154926 | 2 | 324 | 0.82 | 0.86 | 89 | 8 | 3 |
| VR155083 | 2 | 468 | 0.76 | 0.78 | 87 | 9 | 4 |
| VR155088 | 2 | 473 | 0.77 | 0.85 | 85 | 11 | 3 |
| VR191268 | 2 | 342 | 0.69 | 0.71 | 88 | 5 | 7 |
| VR191381 | 2 | 839 | 0.72 | 0.77 | 89 | 7 | 5 |
| VR218864 | 2 | 470 | 0.77 | 0.83 | 86 | 10 | 4 |
| **AVERAGE:** | **N/A** | **539** | **0.76** | **0.80** | **88** | **8** | **5** |

## Quality Control

The California Department of Education (CDE) and ETS implemented rigorous quality control procedures throughout the test development, administration, scoring, analyses, and reporting processes for the 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 Alternate ELPAC. Items developed for the Alternate ELPAC were subject to an extensive item review process. The item writers responsible for developing 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_1).

During administrations of the 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 using the statistical criteria described in subsection [*8.2.6 Summary of Classical Item Analysis Flagging Criteria*](#_Summary_of_Classical_1) 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 and Learning Technology Research & Development (ALTRD) staff 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 Alternate ELPAC form to ensure it met the form-building specifications. Both ETS ALTRD 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 item options were parallel in structure;
* language was grade-level appropriate;
* no more than three multiple-choice 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 Materials

Brief descriptions of the types of materials used for and during testing appear in the following subsections.

#### Test Administration Manuals

ETS staff reviewed to verify 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 *Alternate ELPAC Directions for Administration (DFAs)*, were made available to designated local educational agency (LEA) staff through the Test Operations Management System (TOMS), which required a secure logon.

The manuals used in the administration of the ELPAC are listed in subsection [*5.3.4 Instructions for Test Administration*](#_Instructions_for_Test_1).

#### Processing Test Materials

The Alternate ELPAC is delivered as a computer-based assessment and administered one-on-one. All test questions are scored in the moment by the test examiner, including the expressive items scored according to a rubric. The rubrics are in the *DFAs*. There are no test materials that must be scanned or scored remotely.

Computer-based tests that were submitted and 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 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 quality assurance 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_1).

### Quality Control of Scoring

ETS conforms to high standards of quality and fairness when scoring tests 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 student demographic groups.

#### Machine-Scoring Procedures

To ensure valid item-level scoring for the 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 Alternate ELPAC forms by producing sample results for tests. 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 a test 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 test contained information for each item, keys for multiple-choice 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 Alternate ELPAC were scored by local test examiners when they administered the test. Every LEA that had an eligible Alternate ELPAC student in California was required to either complete the online LEA Certification course on the Moodle Training Site, 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 Alternate ELPAC, refer to subsection [*5.2. Administration and Scoring Training*](#_Administration_and_Scoring).

To ensure the quality of the rubric-scored items, second scoring was conducted by a secondary test examiner on approximately 10 percent of students’ responses on expressive items. Ideally, the secondary test examiner was an educator familiar with the student’s individual testing needs and preferred modes of communication to accurately score the responses.

Results of the interrater reliability were included in section [*8.8.6 Interrater Agreement*](#_Interrater__)*.*

#### 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 a test 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 ALTRD 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 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 a test 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*](#_Development_of_Scoring). 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 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 Alternate ELPAC results data, including item analyses, differential item functioning analyses, item response theory (IRT) calibration, equating, and scaling.

Detailed checklists were developed by members of the team for each of the statistical procedures. Classical item analyses were performed to evaluate the performance of the operational items. Classical item statistics included item difficulty and correlations between item scores and total scores. Items that were flagged for questionable statistical attributes were sent to ETS ALTRD staff for review; their comments were then reviewed by the psychometricians before the review by the CDE. The ETS ALTRD and PAR teams worked together to evaluate and make recommendations to the CDE about any problematic items that should be removed from IRT calibration.

IRT calibration of field test items included checks to ascertain that the input files were established accurately. Checks were also made on the number of items, number of students with valid scores, IRT item difficulty and discrimination estimates, standard errors for the item difficulty estimates, and the equating and scaling process. Two psychometricians conducted parallel calibration processing and compared the results to check for any inconsistency. Psychometricians also performed detailed reviews of relevant statistics to determine whether the chosen IRT model fits the data. ETS then presented and reviewed the calibration results with the CDE for approval.

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 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 and California Educator Reporting System 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, content areas, 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.

#### Exclusion of Student Scores from Summary Reports

ETS provided the CDE with reporting specifications that documented when to exclude student scores from summary reports. These specifications included the logic for handling submitted assessments that, for example, indicated the student tested but responded to no items, was absent, was not tested because of parent/guardian request, or did not complete the assessment because of illness. The methods for handling other anomalies were also covered in the specifications. These anomalies are described in more detail in [*7.3.2 Special Cases*](#_Special_Cases).

### 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 spans and domains. These responses were entered for fictitious students in selected schools and across several LEAs. Each student’s test 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.

### Reference

Educational Testing Service. (2014). *ETS Standards for Quality and Fairness*. Princeton, NJ: Educational Testing Service.

## In-Test Survey

This chapter describes the development and administration of the in-test survey, the survey questionnaires for test examiners, and the results from the analyses of test examiners’ responses.

### Survey Design and Development

The Alternate English Language Proficiency Assessments for California (ELPAC) in-test survey was developed by research staff at ETS in consultation with national experts, technical advisors, and the California Department of Education. The various groups provided guidance in terms of the length of the survey and the questions to consider.

The goal of the survey was to gather validity evidence, for peer review and other purposes, on the following aspects of the Alternate ELPAC:

* **English Language Proficiency (ELP):** To gather an external measure of student ELP and provide concurrent and future checks on the validity of the test
* **Communication modes:** To gather evidence on whether the assessment design allowed students to use a preferred communication mode
* **Accessibility:**
* To gather evidence about how accessibility resources were used during test administration (This allows comparisons between student groups using certain accessibility resources to those who are not using it and any discrepancies on test scores between student groups.)
* To inform future test administration training (e.g., to refine the training content or experience)

Ten survey questions were developed. Survey questions 1 through 5 were designed to collect data about the students’ ELP. The questions support the interpretation of the test scores by establishing a range of skills in the domains of Listening and Reading (receptive communication modes) and Speaking and Writing (expressive communication mode) as shown during classroom instruction. Questions 6 and 7 asked about students’ primary communication mode used during classroom instruction and during the test administration. Questions 8 through 10 were asked to gather information about the student’s use (if any) of universal tools, designated supports, and accommodations during the administration of the Alternate ELPAC. The 10 survey questions with the details on the options and student response frequencies are presented in [appendix 10.A](#_Appendix_10.A:_Student_1).

### Survey Administration

All test examiners were required to respond to the in-test survey questions via the test delivery system. Six questions were to be answered prior to testing the student, and four questions were to be answered after testing the student. Test examiners and local educational agencies were provided with access to the survey questions in the web-based *Alternate ELPAC Operational Field Test Manual* and other test administration resources prior to entering the test delivery system to limit the time a test examiner sat with a student for test administration.

### Summary of Survey Responses

Survey question 1 asked a test examiner to rate the student’s overall ELP level (high or fluent, medium or intermediate, low or novice) based on interactions with the student during classroom instruction. The frequency of the responses to this survey question is presented by grade level and grade span in table 10.A.1 through table 10.A.7 in [appendix 10.A](#_Appendix_10.A:_Student_1). Table 10.A.8 through table 10.A.14 provide the mean and standard deviation of the scale scores from the Alternate ELPAC for students in each of the levels rated by the test examiners. Results showed 73.3 percent agreement between the Alternate ELPAC ELP level based on threshold scores and the teachers’ ratings of students as English learners or as fluent English proficient across grades (ETS, 2023).

Survey questions 2 through 5 asked the test examiner to rate the student’s language skills in Listening, Speaking, Reading, and Writing. Responses to these questions are presented in table 10.A.15 through table 10.A.42 in [appendix 10.A](#_Appendix_10.A:_Student_1). Results indicate that the approximate ordering of skills by domain level is reasonable. As grade level or grade span increases, a higher percentage of students can perform higher-level, more complex skills.

Survey questions 6 and 7 asked about the student’s primary communication mode used during classroom instruction and during the administration of the Alternate ELPAC. Responses to these questions are presented in table 10.A.43 through table 10.A.56 in [appendix 10.A](#_Appendix_10.A:_Student_1). Results indicate that the primary communication modes used by students were consistent for instruction and assessment.

Survey questions 8 through 10 asked about the student’s use of the available universal tools, designated supports, or accommodations. Responses to these questions are presented in table 10.A.57 through table 10.A.77 in [appendix 10.A](#_Appendix_10.A:_Student_1). About 10 percent of students used universal tools during the Alternate ELPAC administration. The most-used designated supports were separate setting and simplified test directions. The most used accommodations were breaks and alternate response options.

### Reference

ETS. (2023). *Summative Alternate English Language Proficiency Assessments for California threshold score validation study report* [Unpublished report.] Princeton, NJ: ETS.

### Appendix 10.A: Student Survey Results

**Note:** “N/A” is not applicable, and “SD” is standard deviation.

Table 10.A. Responses to Question 1 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 167 | 13 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 361 | 27 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 784 | 59 |
| Missing | 23 | 2 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 1 for Grade One

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 190 | 15 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 366 | 28 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 726 | 56 |
| Missing | 15 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 1 for Grade Two

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 220 | 18 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 397 | 32 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 600 | 48 |
| Missing | 22 | 2 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 1 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 960 | 22 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 1,536 | 36 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 1,734 | 40 |
| Missing | 67 | 2 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 1 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 1,069 | 29 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 1,317 | 36 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 1,236 | 34 |
| Missing | 52 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 1 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 627 | 35 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 569 | 32 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 550 | 31 |
| Missing | 33 | 2 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 1 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | N | Percent |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 1,280 | 38 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 969 | 29 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 1,038 | 31 |
| Missing | 59 | 2 |
| **Total:** | 3,346 | 100 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Kindergarten

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 253 | 17 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 250 | 15 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 236 | 19 |
| Missing | 236 | 20 |
| **Total:** | 242 | 19 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade One

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 356 | 18 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 351 | 15 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 339 | 19 |
| Missing | 340 | 14 |
| **Total:** | 345 | 19 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade Two

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 456 | 17 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 448 | 17 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 436 | 18 |
| Missing | 448 | 21 |
| **Total:** | 444 | 19 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 561 | 16 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 555 | 15 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 541 | 17 |
| Missing | 551 | 18 |
| **Total:** | 551 | 18 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 668 | 20 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 660 | 19 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 642 | 20 |
| Missing | 646 | 26 |
| **Total:** | 656 | 23 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 763 | 19 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 756 | 18 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 737 | 18 |
| Missing | 745 | 15 |
| **Total:** | 752 | 21 |

Table 10.A. Scale Score Summary by Response to English Language Proficiency Question for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Based on your interactions with this student during classroom instruction, which of the following best characterizes this student’s current level of overall English language proficiency? | Scale Score Mean | Scale Score SD |
| High or fluent English proficient (Students at this level have sufficient English language proficiency. They may need occasional linguistic support to enable them to access adapted grade-level content in English.) | 865 | 21 |
| Medium or intermediate English learner (Students at this level have moderate English language proficiency. They may need frequent linguistic support to enable them to access adapted grade-level content in English.) | 855 | 17 |
| Low or novice English learner (Students at this level have minimal English language proficiency. They need substantial linguistic support to enable them to access adapted grade-level content in English.) | 837 | 19 |
| Missing | 854 | 25 |
| **Total:** | 853 | 23 |

Table 10.A. Responses to Question 2 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 192 | 14 |
| Follows 1-step directions | 353 | 26 |
| Attends and responds to simple commands | 262 | 20 |
| Indicates a choice when offered an array of items | 166 | 12 |
| Points to or touches objects upon request | 213 | 16 |
| Does not yet attend to sound | 138 | 10 |
| Missing | 11 | 1 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 2 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 248 | 19 |
| Follows 1-step directions | 395 | 30 |
| Attends and responds to simple commands | 255 | 20 |
| Indicates a choice when offered an array of items | 128 | 10 |
| Points to or touches objects upon request | 177 | 14 |
| Does not yet attend to sound | 84 | 6 |
| Missing | 10 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 2 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 335 | 27 |
| Follows 1-step directions | 366 | 30 |
| Attends and responds to simple commands | 236 | 19 |
| Indicates a choice when offered an array of items | 93 | 8 |
| Points to or touches objects upon request | 140 | 11 |
| Does not yet attend to sound | 60 | 5 |
| Missing | 9 | 1 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 2 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 1,566 | 36 |
| Follows 1-step directions | 1,252 | 29 |
| Attends and responds to simple commands | 698 | 16 |
| Indicates a choice when offered an array of items | 257 | 6 |
| Points to or touches objects upon request | 359 | 8 |
| Does not yet attend to sound | 132 | 3 |
| Missing | 33 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 2 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 1,592 | 43 |
| Follows 1-step directions | 919 | 25 |
| Attends and responds to simple commands | 540 | 15 |
| Indicates a choice when offered an array of items | 198 | 5 |
| Points to or touches objects upon request | 286 | 8 |
| Does not yet attend to sound | 111 | 3 |
| Missing | 28 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 2 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 780 | 44 |
| Follows 1-step directions | 423 | 24 |
| Attends and responds to simple commands | 260 | 15 |
| Indicates a choice when offered an array of items | 100 | 6 |
| Points to or touches objects upon request | 134 | 8 |
| Does not yet attend to sound | 69 | 4 |
| Missing | 13 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 2 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s listening skills in English? | N | Percent |
| Follows 2-step directions | 1,645 | 49 |
| Follows 1-step directions | 730 | 22 |
| Attends and responds to simple commands | 411 | 12 |
| Indicates a choice when offered an array of items | 176 | 5 |
| Points to or touches objects upon request | 248 | 7 |
| Does not yet attend to sound | 116 | 3 |
| Missing | 20 | 1 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 3 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 63 | 5 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 196 | 15 |
| Verbally speaks 1-2 word phrases | 217 | 16 |
| Verbally speaks 1 word at a time | 149 | 11 |
| Uses touch and gestures by pointing and head nodding | 182 | 14 |
| Uses a sign language (American Sign Language or other) | 4 | 0 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 237 | 18 |
| Uses eye gaze with intentionality | 31 | 2 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 64 | 5 |
| Intentionally communicative, when interpreted by a familiar individual | 40 | 3 |
| Not yet intentionally communicative | 143 | 11 |
| Missing | 9 | 1 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 3 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 92 | 7 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 253 | 20 |
| Verbally speaks 1-2 word phrases | 199 | 15 |
| Verbally speaks 1 word at a time | 154 | 12 |
| Uses touch and gestures by pointing and head nodding | 192 | 15 |
| Uses a sign language (American Sign Language or other) | 7 | 1 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 172 | 13 |
| Uses eye gaze with intentionality | 20 | 2 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 72 | 6 |
| Intentionally communicative, when interpreted by a familiar individual | 23 | 2 |
| Not yet intentionally communicative | 105 | 8 |
| Missing | 8 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 3 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 136 | 11 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 257 | 21 |
| Verbally speaks 1-2 word phrases | 203 | 16 |
| Verbally speaks 1 word at a time | 135 | 11 |
| Uses touch and gestures by pointing and head nodding | 141 | 11 |
| Uses a sign language (American Sign Language or other) | 7 | 1 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 155 | 13 |
| Uses eye gaze with intentionality | 16 | 1 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 78 | 6 |
| Intentionally communicative, when interpreted by a familiar individual | 27 | 2 |
| Not yet intentionally communicative | 75 | 6 |
| Missing | 9 | 1 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 3 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 889 | 21 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 1,160 | 27 |
| Verbally speaks 1-2 word phrases | 687 | 16 |
| Verbally speaks 1 word at a time | 341 | 8 |
| Uses touch and gestures by pointing and head nodding | 357 | 8 |
| Uses a sign language (American Sign Language or other) | 19 | 0 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 306 | 7 |
| Uses eye gaze with intentionality | 46 | 1 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 245 | 6 |
| Intentionally communicative, when interpreted by a familiar individual | 56 | 1 |
| Not yet intentionally communicative | 158 | 4 |
| Missing | 33 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 3 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 1,084 | 30 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 942 | 26 |
| Verbally speaks 1-2 word phrases | 465 | 13 |
| Verbally speaks 1 word at a time | 247 | 7 |
| Uses touch and gestures by pointing and head nodding | 241 | 7 |
| Uses a sign language (American Sign Language or other) | 25 | 1 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 218 | 6 |
| Uses eye gaze with intentionality | 41 | 1 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 178 | 5 |
| Intentionally communicative, when interpreted by a familiar individual | 64 | 2 |
| Not yet intentionally communicative | 141 | 4 |
| Missing | 28 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 3 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 579 | 33 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 386 | 22 |
| Verbally speaks 1-2 word phrases | 218 | 12 |
| Verbally speaks 1 word at a time | 130 | 7 |
| Uses touch and gestures by pointing and head nodding | 109 | 6 |
| Uses a sign language (American Sign Language or other) | 12 | 1 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 106 | 6 |
| Uses eye gaze with intentionality | 30 | 2 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 106 | 6 |
| Intentionally communicative, when interpreted by a familiar individual | 29 | 2 |
| Not yet intentionally communicative | 65 | 4 |
| Missing | 9 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 3 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s speaking skills in English? | N | Percent |
| Verbally speaks 3 or more words in complete sentences using grammatical rules | 1,218 | 36 |
| Verbally speaks 2-3 or more words in sentences or phrases without consistently following grammatical rules | 648 | 19 |
| Verbally speaks 1-2 word phrases | 397 | 12 |
| Verbally speaks 1 word at a time | 213 | 6 |
| Uses touch and gestures by pointing and head nodding | 206 | 6 |
| Uses a sign language (American Sign Language or other) | 35 | 1 |
| Uses vocalizations, gestures, and facial expressions to communicate intentionally | 231 | 7 |
| Uses eye gaze with intentionality | 48 | 1 |
| Uses an Augmentative and Alternative Communication (AAC) system (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 147 | 4 |
| Intentionally communicative, when interpreted by a familiar individual | 64 | 2 |
| Not yet intentionally communicative | 125 | 4 |
| Missing | 14 | 0 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 4 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 17 | 1 |
| Reads text without symbol support but without comprehension | 30 | 2 |
| Identifies individual words without picture support | 59 | 4 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 78 | 6 |
| Recognizes letter sounds (knows sounds associated with letters) | 113 | 8 |
| Recognizes letters (can identify them by name) | 205 | 15 |
| Matches objects to pictures | 280 | 21 |
| Identifies and names objects | 84 | 6 |
| Does not yet have an understanding of print or text | 459 | 34 |
| Missing | 10 | 1 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 4 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 39 | 3 |
| Reads text without symbol support but without comprehension | 64 | 5 |
| Identifies individual words without picture support | 61 | 5 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 127 | 10 |
| Recognizes letter sounds (knows sounds associated with letters) | 117 | 9 |
| Recognizes letters (can identify them by name) | 208 | 16 |
| Matches objects to pictures | 270 | 21 |
| Identifies and names objects | 62 | 5 |
| Does not yet have an understanding of print or text | 340 | 26 |
| Missing | 9 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 4 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 51 | 4 |
| Reads text without symbol support but without comprehension | 93 | 8 |
| Identifies individual words without picture support | 79 | 6 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 136 | 11 |
| Recognizes letter sounds (knows sounds associated with letters) | 130 | 10 |
| Recognizes letters (can identify them by name) | 165 | 13 |
| Matches objects to pictures | 265 | 21 |
| Identifies and names objects | 61 | 5 |
| Does not yet have an understanding of print or text | 252 | 20 |
| Missing | 7 | 1 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 4 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 450 | 10 |
| Reads text without symbol support but without comprehension | 420 | 10 |
| Identifies individual words without picture support | 380 | 9 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 685 | 16 |
| Recognizes letter sounds (knows sounds associated with letters) | 410 | 10 |
| Recognizes letters (can identify them by name) | 475 | 11 |
| Matches objects to pictures | 704 | 16 |
| Identifies and names objects | 145 | 3 |
| Does not yet have an understanding of print or text | 588 | 14 |
| Missing | 40 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 4 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 683 | 19 |
| Reads text without symbol support but without comprehension | 512 | 14 |
| Identifies individual words without picture support | 308 | 8 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 629 | 17 |
| Recognizes letter sounds (knows sounds associated with letters) | 205 | 6 |
| Recognizes letters (can identify them by name) | 251 | 7 |
| Matches objects to pictures | 486 | 13 |
| Identifies and names objects | 110 | 3 |
| Does not yet have an understanding of print or text | 469 | 13 |
| Missing | 22 | 1 |
| **Total:** | 3,678 | 100 |

Table 10.A. Responses to Question 4 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 437 | 25 |
| Reads text without symbol support but without comprehension | 209 | 12 |
| Identifies individual words without picture support | 140 | 8 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 301 | 17 |
| Recognizes letter sounds (knows sounds associated with letters) | 62 | 3 |
| Recognizes letters (can identify them by name) | 107 | 6 |
| Matches objects to pictures | 208 | 12 |
| Identifies and names objects | 51 | 3 |
| Does not yet have an understanding of print or text | 252 | 14 |
| Missing | 12 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 4 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s reading skills in English? | N | Percent |
| Reads text without any symbol support with comprehension | 941 | 28 |
| Reads text without symbol support but without comprehension | 372 | 11 |
| Identifies individual words without picture support | 214 | 6 |
| Reads words, phrases, or sentences when pictures/symbols are provided for support | 499 | 15 |
| Recognizes letter sounds (knows sounds associated with letters) | 130 | 4 |
| Recognizes letters (can identify them by name) | 184 | 5 |
| Matches objects to pictures | 453 | 14 |
| Identifies and names objects | 92 | 3 |
| Does not yet have an understanding of print or text | 441 | 13 |
| Missing | 21 | 1 |
| **Total:** | 3,347 | 100 |

Table 10.A. Responses to Question 5 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 2 | 0 |
| Writes a simple sentence or phrase (spelling not always correct) | 7 | 1 |
| Writes words (spelling not always correct) | 38 | 3 |
| Writes using word banks | 7 | 1 |
| Selects letters or symbols to express meaning | 34 | 3 |
| Copies letters and words, but does not produce independent writing | 289 | 22 |
| Randomly selects letters or symbols when asked to write | 85 | 6 |
| Makes random marks or scribbles | 542 | 41 |
| Does not yet demonstrate expressive writing skills | 321 | 24 |
| Missing | 13 | 1 |
| **Total:** | 1,338 | 100 |

Table 10.A. Responses to Question 5 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 7 | 1 |
| Writes a simple sentence or phrase (spelling not always correct) | 46 | 4 |
| Writes words (spelling not always correct) | 67 | 5 |
| Writes using word banks | 33 | 3 |
| Selects letters or symbols to express meaning | 33 | 3 |
| Copies letters and words, but does not produce independent writing | 342 | 26 |
| Randomly selects letters or symbols when asked to write | 88 | 7 |
| Makes random marks or scribbles | 425 | 33 |
| Does not yet demonstrate expressive writing skills | 250 | 19 |
| Missing | 11 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 5 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 18 | 1 |
| Writes a simple sentence or phrase (spelling not always correct) | 71 | 6 |
| Writes words (spelling not always correct) | 86 | 7 |
| Writes using word banks | 43 | 3 |
| Selects letters or symbols to express meaning | 37 | 3 |
| Copies letters and words, but does not produce independent writing | 381 | 31 |
| Randomly selects letters or symbols when asked to write | 77 | 6 |
| Makes random marks or scribbles | 335 | 27 |
| Does not yet demonstrate expressive writing skills | 182 | 15 |
| Missing | 9 | 1 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 5 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 258 | 6 |
| Writes a simple sentence or phrase (spelling not always correct) | 493 | 11 |
| Writes words (spelling not always correct) | 419 | 10 |
| Writes using word banks | 278 | 6 |
| Selects letters or symbols to express meaning | 115 | 3 |
| Copies letters and words, but does not produce independent writing | 1,313 | 31 |
| Randomly selects letters or symbols when asked to write | 218 | 5 |
| Makes random marks or scribbles | 690 | 16 |
| Does not yet demonstrate expressive writing skills | 473 | 11 |
| Missing | 40 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 5 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 547 | 15 |
| Writes a simple sentence or phrase (spelling not always correct) | 643 | 18 |
| Writes words (spelling not always correct) | 381 | 10 |
| Writes using word banks | 227 | 6 |
| Selects letters or symbols to express meaning | 103 | 3 |
| Copies letters and words, but does not produce independent writing | 811 | 22 |
| Randomly selects letters or symbols when asked to write | 127 | 3 |
| Makes random marks or scribbles | 420 | 11 |
| Does not yet demonstrate expressive writing skills | 383 | 10 |
| Missing | 32 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 5 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 359 | 20 |
| Writes a simple sentence or phrase (spelling not always correct) | 298 | 17 |
| Writes words (spelling not always correct) | 200 | 11 |
| Writes using word banks | 86 | 5 |
| Selects letters or symbols to express meaning | 38 | 2 |
| Copies letters and words, but does not produce independent writing | 355 | 20 |
| Randomly selects letters or symbols when asked to write | 42 | 2 |
| Makes random marks or scribbles | 174 | 10 |
| Does not yet demonstrate expressive writing skills | 212 | 12 |
| Missing | 15 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 5 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following best describes the student’s writing skills in English? | N | Percent |
| Writes 1-3 sentences (spelling not always correct) | 804 | 24 |
| Writes a simple sentence or phrase (spelling not always correct) | 570 | 17 |
| Writes words (spelling not always correct) | 287 | 9 |
| Writes using word banks | 150 | 4 |
| Selects letters or symbols to express meaning | 88 | 3 |
| Copies letters and words, but does not produce independent writing | 618 | 18 |
| Randomly selects letters or symbols when asked to write | 72 | 2 |
| Makes random marks or scribbles | 356 | 11 |
| Does not yet demonstrate expressive writing skills | 374 | 11 |
| Missing | 27 | 1 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 6 for Kindergarten

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 126 | 9 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 174 | 13 |
| Verbal (i.e., spoken language) – single word responses | 234 | 18 |
| Writing | 0 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 378 | 28 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 94 | 7 |
| Eye gaze | 48 | 4 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 4 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 100 | 7 |
| Other | 16 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 130 | 10 |
| I am not sure what communication modes the student uses in the classroom | 21 | 2 |
| Missing | 10 | 1 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 6 for Grade One

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 177 | 14 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 226 | 17 |
| Verbal (i.e., spoken language) – single word responses | 233 | 18 |
| Writing | 2 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 340 | 26 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 92 | 7 |
| Eye gaze | 35 | 3 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 5 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 80 | 6 |
| Other | 14 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 70 | 5 |
| I am not sure what communication modes the student uses in the classroom | 10 | 1 |
| Missing | 13 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 6 for Grade Two

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 232 | 19 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 220 | 18 |
| Verbal (i.e., spoken language) – single word responses | 229 | 18 |
| Writing | 0 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 270 | 22 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 88 | 7 |
| Eye gaze | 29 | 2 |
| Braille (either contracted or uncontracted) | 1 | 0 |
| American Sign Language (ASL) or other signed response | 7 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 79 | 6 |
| Other | 9 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 55 | 4 |
| I am not sure what communication modes the student uses in the classroom | 12 | 1 |
| Missing | 8 | 1 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 6 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,402 | 33 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 875 | 20 |
| Verbal (i.e., spoken language) – single word responses | 670 | 16 |
| Writing | 3 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 609 | 14 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 320 | 7 |
| Eye gaze | 50 | 1 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 20 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 151 | 4 |
| Other | 19 | 0 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 121 | 3 |
| I am not sure what communication modes the student uses in the classroom | 19 | 0 |
| Missing | 38 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 6 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,501 | 41 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 670 | 18 |
| Verbal (i.e., spoken language) – single word responses | 472 | 13 |
| Writing | 5 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 431 | 12 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 208 | 6 |
| Eye gaze | 58 | 2 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 33 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 122 | 3 |
| Other | 22 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 99 | 3 |
| I am not sure what communication modes the student uses in the classroom | 24 | 1 |
| Missing | 29 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 6 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 723 | 41 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 334 | 19 |
| Verbal (i.e., spoken language) – single word responses | 205 | 12 |
| Writing | 8 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 203 | 11 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 116 | 7 |
| Eye gaze | 28 | 2 |
| Braille (either contracted or uncontracted) | 1 | 0 |
| American Sign Language (ASL) or other signed response | 12 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 56 | 3 |
| Other | 9 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 57 | 3 |
| I am not sure what communication modes the student uses in the classroom | 13 | 1 |
| Missing | 14 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 6 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| During classroom instruction, what is the primary communication mode that the student uses to communicate, either in English or in another language? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,499 | 45 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 529 | 16 |
| Verbal (i.e., spoken language) – single word responses | 366 | 11 |
| Writing | 7 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 401 | 12 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 168 | 5 |
| Eye gaze | 68 | 2 |
| Braille (either contracted or uncontracted) | 1 | 0 |
| American Sign Language (ASL) or other signed response | 37 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 110 | 3 |
| Other | 22 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 97 | 3 |
| I am not sure what communication modes the student uses in the classroom | 21 | 1 |
| Missing | 20 | 1 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 7 for Kindergarten

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 67 | 5 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 131 | 10 |
| Verbal (i.e., spoken language) – single word responses | 253 | 19 |
| Writing | 1 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 565 | 42 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 49 | 4 |
| Eye gaze | 34 | 3 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 2 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 45 | 3 |
| Other | 14 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 151 | 11 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 23 | 2 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 7 for Grade One

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 112 | 9 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 198 | 15 |
| Verbal (i.e., spoken language) – single word responses | 260 | 20 |
| Writing | 0 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 485 | 37 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 53 | 4 |
| Eye gaze | 30 | 2 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 0 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 34 | 3 |
| Other | 13 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 101 | 8 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 11 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 7 for Grade Two

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 152 | 12 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 171 | 14 |
| Verbal (i.e., spoken language) – single word responses | 246 | 20 |
| Writing | 1 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 458 | 37 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 35 | 3 |
| Eye gaze | 21 | 2 |
| Braille (either contracted or uncontracted) | 2 | 0 |
| American Sign Language (ASL) or other signed response | 4 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 25 | 2 |
| Other | 14 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 80 | 6 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 30 | 2 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 7 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,059 | 25 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 801 | 19 |
| Verbal (i.e., spoken language) – single word responses | 765 | 18 |
| Writing | 2 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 1,130 | 26 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 161 | 4 |
| Eye gaze | 53 | 1 |
| Braille (either contracted or uncontracted) | 1 | 0 |
| American Sign Language (ASL) or other signed response | 13 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 66 | 2 |
| Other | 29 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 164 | 4 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 53 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 7 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,255 | 34 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 665 | 18 |
| Verbal (i.e., spoken language) – single word responses | 550 | 15 |
| Writing | 11 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 763 | 21 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 113 | 3 |
| Eye gaze | 46 | 1 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 16 | 0 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 47 | 1 |
| Other | 27 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 127 | 3 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 54 | 1 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 7 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 647 | 36 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 275 | 15 |
| Verbal (i.e., spoken language) – single word responses | 250 | 14 |
| Writing | 2 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 368 | 21 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 49 | 3 |
| Eye gaze | 40 | 2 |
| Braille (either contracted or uncontracted) | 0 | 0 |
| American Sign Language (ASL) or other signed response | 9 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 39 | 2 |
| Other | 11 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 68 | 4 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 21 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 7 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| During the Alternate ELPAC administration, what was the primary communication mode that the student used to respond to test questions? | N | Percent |
| Verbal (i.e., spoken language) – responses of three words or more | 1,257 | 38 |
| Verbal (i.e., spoken language) – two words responses or fixed phrases | 501 | 15 |
| Verbal (i.e., spoken language) – single word responses | 445 | 13 |
| Writing | 11 | 0 |
| Gesture (e.g., pointing, nodding, touching, arranging) | 675 | 20 |
| Augmentative and Alternative Communication (AAC) systems (e.g., communication board, picture cards, Big Mack Switch, Proloquo2 Go APP on iPad) | 94 | 3 |
| Eye gaze | 58 | 2 |
| Braille (either contracted or uncontracted) | 1 | 0 |
| American Sign Language (ASL) or other signed response | 35 | 1 |
| Vocalizations (i.e., sounds made orally but not recognizable as words) | 47 | 1 |
| Other | 34 | 1 |
| The student does not yet have an established communication mode and does not yet demonstrate communicative intent. | 128 | 4 |
| I am not sure what communication modes the student uses in the classroom | 0 | 0 |
| Missing | 60 | 2 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 8 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 166 | 12 |
| No | 1,143 | 86 |
| Missing | 26 | 2 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 8 for Grade One

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 140 | 11 |
| No | 1,148 | 89 |
| Missing | 9 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 8 for Grade Two

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 141 | 11 |
| No | 1,072 | 87 |
| Missing | 26 | 2 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 8 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 439 | 10 |
| No | 3,810 | 89 |
| Missing | 48 | 1 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 8 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 296 | 8 |
| No | 3,322 | 90 |
| Missing | 56 | 2 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 8 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 155 | 9 |
| No | 1,606 | 90 |
| Missing | 18 | 1 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 8 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Did the student use any additional Universal Tools, other than expanding the items and passages as recommended for all students, during the Alternate ELPAC administration? (That is, did the student use other tools such as Zoom, the Highlighter, the Digital Notepad, etc.) | N | Percent |
| Yes | 310 | 9 |
| No | 2,987 | 89 |
| Missing | 49 | 1 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 9 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 13 | 1 |
| Color contrast | 16 | 1 |
| Color overlay | 7 | 1 |
| Designated Interface Assistant | 21 | 2 |
| Magnification | 96 | 7 |
| Masking | 84 | 6 |
| Medical supports | 6 | 0 |
| Mouse pointer | 72 | 5 |
| Noise buffer | 43 | 3 |
| Permissive Mode | 3 | 0 |
| Print on demand | 19 | 1 |
| Print Size | 48 | 4 |
| Separate Setting | 546 | 41 |
| Simplified test directions | 375 | 28 |
| Streamline | 19 | 1 |
| Translated test directions | 35 | 3 |
| Turn off any universal tool(s) | 7 | 1 |
| No designated supports used | 493 | 37 |
| Missing | 70 | 5 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 9 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 30 | 2 |
| Color contrast | 16 | 1 |
| Color overlay | 1 | 0 |
| Designated Interface Assistant | 15 | 1 |
| Magnification | 85 | 7 |
| Masking | 97 | 7 |
| Medical supports | 3 | 0 |
| Mouse pointer | 92 | 7 |
| Noise buffer | 49 | 4 |
| Permissive Mode | 0 | 0 |
| Print on demand | 23 | 2 |
| Print Size | 31 | 2 |
| Separate Setting | 503 | 39 |
| Simplified test directions | 373 | 29 |
| Streamline | 11 | 1 |
| Translated test directions | 31 | 2 |
| Turn off any universal tool(s) | 3 | 0 |
| No designated supports used | 492 | 38 |
| Missing | 63 | 5 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 9 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 19 | 2 |
| Color contrast | 4 | 0 |
| Color overlay | 4 | 0 |
| Designated Interface Assistant | 20 | 2 |
| Magnification | 95 | 8 |
| Masking | 52 | 4 |
| Medical supports | 5 | 0 |
| Mouse pointer | 83 | 7 |
| Noise buffer | 65 | 5 |
| Permissive Mode | 1 | 0 |
| Print on demand | 11 | 1 |
| Print Size | 35 | 3 |
| Separate Setting | 507 | 41 |
| Simplified test directions | 380 | 31 |
| Streamline | 24 | 2 |
| Translated test directions | 20 | 2 |
| Turn off any universal tool(s) | 2 | 0 |
| No designated supports used | 400 | 32 |
| Missing | 81 | 7 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 9 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 62 | 1 |
| Color contrast | 40 | 1 |
| Color overlay | 19 | 0 |
| Designated Interface Assistant | 66 | 2 |
| Magnification | 270 | 6 |
| Masking | 175 | 4 |
| Medical supports | 12 | 0 |
| Mouse pointer | 291 | 7 |
| Noise buffer | 271 | 6 |
| Permissive Mode | 7 | 0 |
| Print on demand | 36 | 1 |
| Print Size | 74 | 2 |
| Separate Setting | 1,863 | 43 |
| Simplified test directions | 1,237 | 29 |
| Streamline | 74 | 2 |
| Translated test directions | 57 | 1 |
| Turn off any universal tool(s) | 16 | 0 |
| No designated supports used | 1,505 | 35 |
| Missing | 234 | 5 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 9 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 60 | 2 |
| Color contrast | 36 | 1 |
| Color overlay | 15 | 0 |
| Designated Interface Assistant | 51 | 1 |
| Magnification | 197 | 5 |
| Masking | 90 | 2 |
| Medical supports | 13 | 0 |
| Mouse pointer | 279 | 8 |
| Noise buffer | 166 | 5 |
| Permissive Mode | 13 | 0 |
| Print on demand | 18 | 0 |
| Print Size | 103 | 3 |
| Separate Setting | 1,485 | 40 |
| Simplified test directions | 878 | 24 |
| Streamline | 33 | 1 |
| Translated test directions | 43 | 1 |
| Turn off any universal tool(s) | 7 | 0 |
| No designated supports used | 1,537 | 42 |
| Missing | 173 | 5 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 9 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 23 | 1 |
| Color contrast | 11 | 1 |
| Color overlay | 10 | 1 |
| Designated Interface Assistant | 22 | 1 |
| Magnification | 107 | 6 |
| Masking | 43 | 2 |
| Medical supports | 10 | 1 |
| Mouse pointer | 125 | 7 |
| Noise buffer | 56 | 3 |
| Permissive Mode | 0 | 0 |
| Print on demand | 14 | 1 |
| Print Size | 37 | 2 |
| Separate Setting | 613 | 34 |
| Simplified test directions | 356 | 20 |
| Streamline | 13 | 1 |
| Translated test directions | 32 | 2 |
| Turn off any universal tool(s) | 6 | 0 |
| No designated supports used | 822 | 46 |
| Missing | 78 | 4 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 9 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following Designated Supports, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| American Sign Language/Manually Coded English (ASL/MCE) for test directions | 59 | 2 |
| Color contrast | 38 | 1 |
| Color overlay | 14 | 0 |
| Designated Interface Assistant | 52 | 2 |
| Magnification | 171 | 5 |
| Masking | 84 | 3 |
| Medical supports | 15 | 0 |
| Mouse pointer | 249 | 7 |
| Noise buffer | 113 | 3 |
| Permissive Mode | 5 | 0 |
| Print on demand | 27 | 1 |
| Print Size | 78 | 2 |
| Separate Setting | 935 | 28 |
| Simplified test directions | 638 | 19 |
| Streamline | 51 | 2 |
| Translated test directions | 61 | 2 |
| Turn off any universal tool(s) | 8 | 0 |
| No designated supports used | 1,635 | 49 |
| Missing | 201 | 6 |
| **Total:** | 3,346 | 100 |

Table 10.A. Responses to Question 10 for Kindergarten

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 202 | 15 |
| Alternate response options | 427 | 32 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 18 | 1 |
| Breaks | 691 | 52 |
| Scribe | 27 | 2 |
| Speech-to-text | 26 | 2 |
| No accommodations used | 370 | 28 |
| Missing | 31 | 2 |
| **Total:** | 1,335 | 100 |

Table 10.A. Responses to Question 10 for Grade One

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 208 | 16 |
| Alternate response options | 412 | 32 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 39 | 3 |
| Breaks | 657 | 51 |
| Scribe | 35 | 3 |
| Speech-to-text | 27 | 2 |
| No accommodations used | 373 | 29 |
| Missing | 14 | 1 |
| **Total:** | 1,297 | 100 |

Table 10.A. Responses to Question 10 for Grade Two

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 171 | 14 |
| Alternate response options | 400 | 32 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 32 | 3 |
| Breaks | 554 | 45 |
| Scribe | 44 | 4 |
| Speech-to-text | 24 | 2 |
| No accommodations used | 391 | 32 |
| Missing | 29 | 2 |
| **Total:** | 1,239 | 100 |

Table 10.A. Responses to Question 10 for Grade Span Three Through Five

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 557 | 13 |
| Alternate response options | 1,086 | 25 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 87 | 2 |
| Breaks | 1,667 | 39 |
| Scribe | 138 | 3 |
| Speech-to-text | 139 | 3 |
| No accommodations used | 1,770 | 41 |
| Missing | 82 | 2 |
| **Total:** | 4,297 | 100 |

Table 10.A. Responses to Question 10 for Grade Span Six Through Eight

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 377 | 10 |
| Alternate response options | 703 | 19 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 79 | 2 |
| Breaks | 1,113 | 30 |
| Scribe | 87 | 2 |
| Speech-to-text | 107 | 3 |
| No accommodations used | 1,942 | 53 |
| Missing | 68 | 2 |
| **Total:** | 3,674 | 100 |

Table 10.A. Responses to Question 10 for Grade Span Nine and Ten

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 160 | 9 |
| Alternate response options | 352 | 20 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 31 | 2 |
| Breaks | 411 | 23 |
| Scribe | 35 | 2 |
| Speech-to-text | 61 | 3 |
| No accommodations used | 1,015 | 57 |
| Missing | 27 | 2 |
| **Total:** | 1,779 | 100 |

Table 10.A. Responses to Question 10 for Grade Span Eleven and Twelve

|  |  |  |
| --- | --- | --- |
| Which of the following Accommodations, if any, did the student use during the Alternate ELPAC administration? (Please choose as many as are applicable.) | N | Percent |
| Additional Instructional Supports and Resources for Alternate Assessments | 303 | 9 |
| Alternate response options | 627 | 19 |
| American Sign Language/Manually Coded English (ASL/MCE) for content and responses | 70 | 2 |
| Breaks | 701 | 21 |
| Scribe | 60 | 2 |
| Speech-to-text | 81 | 2 |
| No accommodations used | 1,989 | 59 |
| Missing | 63 | 2 |
| **Total:** | 3,346 | 100 |

## Continuous and Systematic Improvement

This chapter discusses the various procedures used to gather information to improve the Alternate English Language Proficiency Assessments for California (ELPAC) as well as strategies to implement possible future improvements. Some sections within this chapter include feedback from the Alternate ELPAC Operational Field Test Post-Test Survey, which focuses on feedback for continuous and systematic improvement.

### 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 (IEP) 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 Operational Field Test 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 test met each guiding principle. While the feedback shows evidence that educators found that the test design was successful at achieving the guiding principles, ETS will continue to work in collaboration with the California Department of Education (CDE) in evaluating and improving the Alternate ELPAC test design.

### Item Development and Item Pool Diversity

The Alternate ELPAC consists of two separate assessments—the Initial Alternate ELPAC, for initial identification for EL programs and services, and the annual Summative Alternate ELPAC, to measure student progress toward ELP and support decisions on student reclassification. (The Initial Alternate ELPAC had not yet had its inaugural administration when the Summative Alternate ELPAC operational field test was administered.) These assessments follow a single test blueprint, as described in the high-level test design for the Alternate ELPAC (CDE, 2019). Each item development cycle builds upon feedback received in previous item development cycles, including information received at item review meetings and data review meetings. Early development cycles for the Alternate ELPAC program also relied heavily on information from the Alternate ELPAC Pilot cognitive labs study (CDE, 2020).

Analysis of the pilot results led to modifications of the proposed Alternate ELPAC test blueprint. For example, one of the piloted task types was removed. In addition, test design features were added to increase the accessibility of expressive (constructed response) items. Optional individualization was added to the test to allow students who use alternate response options, such as picture cards and realia, as a form of expressive language to communicate, as they do within the classroom setting.

As the Alternate ELPAC program develops items year after year, CDE and educator feedback will continue to inform aspects of item development such as item type appropriateness, accessibility, item difficulty, readability, graphics selection, and diversity of topics. The items will continue to be developed on the basis of item development plans that rely on analysis of the existing item bank. The decisions made during item development will continue to support the fulfillment of the test blueprint for each summative form administration and the growth and diversification of the operational item bank, as well as the incremental release of test items to the public in training and practice tests.

### Administration and Test Delivery

#### Post-Test Survey

The Alternate ELPAC Operational Field Test Post-Test Survey was developed by ETS in consultation with the CDE and focused on gathering information and data from educators who were part of the Alternate ELPAC operational field test administration to highlight successes and identify areas for immediate and long-term improvement. The survey focused on actionable improvement in the areas of test preparation, training, and test administration. Feedback on questions specifically about test design, accessibility, and training were included, as appropriate.

Survey respondents were asked a series of questions regarding testing administration materials, manuals, communications, training, test administration systems, and the assignment and implementation of accessibility resources. Survey respondents were also asked for suggestions for further improvement.

Of the 300 respondents, a majority administered the test themselves; the remaining respondents were involved in the administration at the coordination level only. Most of the respondents who administered the assessment indicated that they administered it to multiple students.

Educators involved in administering the Alternate ELPAC operational field test expressed positive experiences in their preparations for testing. Despite this being the first Alternate ELPAC administration, educators felt that the resources and training materials they were given were useful in preparing them and their students. Their feedback generally described adequate preparation, training, support, and assessment administration experiences. The materials offered in preparation for the Alternate ELPAC operational field test were found to be helpful or very helpful to most respondents. When asked about issues encountered with various systems and technology that were a part of the Alternate ELPAC test administration, most respondents reported never having issues.

With regard to second scoring, only a very small percentage of respondents that conducted second scoring said that the process provided to them to conduct second scoring was not clear and sufficient. More than one half of the respondents who conducted second scoring reported no challenges while completing second scoring. Not having staff available to second score was the most common challenge respondents reported with completing second scoring. Second scoring materials that were offered in preparation for second scoring were also seen as helpful by almost all respondents that accessed them.

Educators provided valuable feedback for potential improvements to the future administration of the Alternate ELPAC by reporting some lessons they learned this year. In response to the local educational agency (LEA) feedback, ETS will implement the following improvements for the 2022–23 operational administration:

* Updating the Alternate ELPAC Administration and Scoring Training (AST) to include shorter, more focused modules and videos
* Creating a *Preparation for Administration* document that includes the information that was formerly at the beginning of the *Directions for Administration (DFAs),* focusing on the purpose of the documents
* Reevaluating the second scoring assignment process in an attempt to minimize burden on LEAs
* Creating new demonstration videos and updating existing videos, where possible, with suggestions provided by LEAs

#### Training and Communication

As ETS continues work on the Summative Alternate ELPAC, recruitment, training, and communication will be a focal point moving forward. ETS will continue to provide timely communications for each critical component of the Alternate ELPAC administration, including information on key dates, availability of materials, and training schedules. Additionally, as a final attempt to ensure that all LEAs complete the trainings required to administer the 2022–23 Summative Alternate ELPAC, ETS will send communications to superintendents of LEAs whose certifications are overdue. Other communications will also encourage LEAs to use the practice and training tests to prepare students and test examiners.

ETS will continue to work with the Sacramento County Office of Education to emphasize the importance and necessity of training, along with providing statewide training to LEA staff so they are prepared to administer the test. Training will focus on proper administration and scoring of open-ended, rubric-scored items as well as the use of the various individualization options available. ETS will also continue to support familiarizing students with the Alternate ELPAC items using practice and training tests and informational videos.

The results of the Alternate ELPAC Operational Field Test Post-Test Survey show that a majority of respondents—79 percent—found the AST via the Moodle Training Site helpful or very helpful. Suggestions from the LEAs included shortening the length of the videos, having resources open in pop-up windows, and having the ability to rewind videos and play videos at a faster pace. These will be incorporated into improvements to the AST.

#### Accessibility

With the launch of the Alternate ELPAC, students have access to a full suite of accessibility resources during testing. The pilot cognitive lab provided an opportunity to evaluate the embedded and non-embedded universal tools and designated supports, as well as to consider the embedded and non-embedded accommodations that were referenced in the ELPAC accessibility framework. The LEA staff assigns and verifies designated supports and accommodations in the Test Operations Management System (TOMS) prior to the student testing. Universal tools are available to all students in the computer-based interface.

During the 2022–23 Alternate ELPAC administration, the following additional options will be added to the color contrast designated support:

* Yellow font on a black background
* Red font on a white background
* White font on a red background

Most of the educators who specifically administered the test to students who are deaf or hard of hearing responded that they had no recommendations to make the test more accessible to these students. The majority of respondents who used picture cards or real objects (manipulatives) agreed or strongly agreed that the use of picture cards or real objects (manipulatives) was helpful in enabling their students to access the test.

Almost all respondents agreed or strongly agreed that, in general, they were able to provide accessibility resources as appropriate to their student.

#### *Directions for Administration*

Test examiners used grade level– and grade span–specific forms of the secure *DFAs* for the Alternate ELPAC to administer the Alternate ELPAC to students. *DFAs* were downloaded from the secure TOMS website.

Test examiner training directed them to follow all directions and guidelines and read, word-for-word, the instructions to students in the administration script to ensure standardization of test administration. *DFAs* also included scoring rubrics where warranted. The publicly available Alternate ELPAC practice and training tests also included *DFAs* available to anyone accessing the practice and training tests.

There were lessons from the pilot applied to operational test administration, such as identification of the need for better support around the administration of rubric-scored items and, most notably, development of items that allow for modeling. Almost all respondents of the Alternate ELPAC Operational Field Test Post-Test Survey found that the *DFAs* provided all the information they needed to administer the test. Most respondents agreed or strongly agreed that the scoring rubrics within the *DFA* were clear and sufficient to score student responses, and almost all respondents agreed or strongly agreed that the *DFA* was clear and helpful in explaining how to present test questions that allowed “modeling” of student responses.

In the transition from the operational field test to the Initial Alternate ELPAC, the *DFA* content that was more focused on preparation and planning, as opposed to the script, will be removed from the *DFA* and developed as a separate document. ETS and the CDE will continue to collaborate on evaluating the *DFAs* for continuous improvement as additional feedback is obtained.

### Psychometric Analyses

As the computer-based Alternate ELPAC transitioned from an operational field test to an operational administration beginning in spring 2020, the ETS Psychometric Analysis & Research team continued to maintain best practices to ensure quality of psychometric results and looked for ways to streamline and improve psychometric processes. ETS will implement a plan to automate some of the manual psychometric reviews (e.g., reviewing item analysis results to evaluate items flagged because of out-of-range classical item statistics). Automation of manual procedures will help to facilitate timely reporting of student scores. ETS will continue to look for ways to provide more validity evidence with each operational administration of the test.

### Reference

California Department of Education. (2019). *Proposed high-level test design for the Alternate English Language Proficiency Assessments for California.* Approved by the California State Board of Education in May 2019. Sacramento, CA: California Department of Education.

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

California Department of Education. (2023). *Alternate ELPAC operational field test post-test survey summary report*. [Unpublished report]. California Department of Education.

1. Data for 2021–22 was retrieved from the *CalEdFacts* web page on the CDE website. [↑](#footnote-ref-2)
2. This definition was retrieved from the CDE California Longitudinal Pupil Achievement Data System (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 2021–22 ELPAC administration. [↑](#footnote-ref-4)