California High School Exit Examination

Technical Report

July 2011–May 2012 Administrations

Prepared by

Educational Testing Service

for

The California Department of Education

December 2012

Copyright © 2012 by the California Department of Education

Table of Contents

EXECUTIVE SUMMARY **1**

CHAPTER 1: INTRODUCTION **9**

BACKGROUND 9
Test Purpose *9*
Content *9*
Target Population *10*
Intended Use and Purpose of Test Scores *10*
Schedule of Administrations and Participation Rules *10*
Significant Developments in the 2011–12 School Year *11*
Limitations of the Assessment *12*
Groups and Organizations *12*
OVERVIEW OF THE TECHNICAL REPORT 13

CHAPTER 2: TEST SPECIFICATIONS **15**

ITEM DEVELOPMENT 15
Item Formats *15*
Model for Generating Item Statistics *15*
Item Banking *16*
Item Refresh Rate and Released Test Questions *17*
TEST ASSEMBLY 17
Test Blueprint *17*
Test Length *17*
Number of Test Forms *18*
Content and Process Categories *18*
Content Rules and Item Selection *19*
Psychometric Criteria *19*
Item Arrangement *20*
TEST ADMINISTRATION 20
Test Security and Confidentiality *20*
Procedures to Maintain Standardization *20*
Demographic Distributions *21*
TEST VARIATIONS, ACCOMMODATIONS AND MODIFICATIONS 22
SCORES 26
Weighting of Scores *26*
Aggregation Procedures *27*
Individual Scores *27*
Group Scores *27*
EQUATING 28
APPENDIX 2.A: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—JULY 2011 29
APPENDIX 2.B: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—OCTOBER 2011 39
APPENDIX 2.C: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—NOVEMBER 2011 53
APPENDIX 2.D: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—DECEMBER 2011 68
APPENDIX 2.E: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—FEBRUARY 2012 75
APPENDIX 2.F: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MARCH 2012 89
APPENDIX 2.G: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MAY 2012 105

CHAPTER 3: ITEM DEVELOPMENT **118**

1.0 Item Specifications *118*

2.0 Prepare Item Development Plan *120*
3.0 Train Item Writers *120*
4.0 Order Items Based on Needs Assessment *121*
5.0 Internal Item Reviews (Educational Testing Service) *121*
6.0 External Item Reviews (California Educators and California Department of Education) *123*
7.0 Item Banking *126*
8.0 Create Field-Test Sets *127*
9.0 Field-Test Items *128*
10.0 Review and Process Field-Test Data *129*
11.0 Create Operational Forms with California Department of Education Review *130*
12.0 Administer Operational Test *130*

CHAPTER 4: TEST DEVELOPMENT 131

TEST LENGTH 131
RULES FOR ITEM SELECTION 131
Test Blueprint *131*
Content Rules and Item Selection *132*
Psychometric Criteria *132*
Rules for Item Sequence and Layout *137*

CHAPTER 5: TEST ADMINISTRATION 138

TEST SECURITY AND CONFIDENTIALITY 138
Educational Testing Service Office of Testing Integrity *138*
Test Development *138*
Item Review *139*
Item Bank *139*
Transfer of Forms and Items to the California Department of Education *140*
Printing *140*
Test Administration *140*
Test Delivery *140*
Processing Security *142*
Data Management *142*
Transfer of Files via Secure Data Exchange *142*
Scoring and Analysis *142*
Reporting and Posting Results *143*
Student Confidentiality *143*
Data Security *143*
PROCEDURES TO MAINTAIN STANDARDIZATION 144
Test Administrators *144*
CAHSEE Directions for Administration *145*
CAHSEE District and Test Site Coordinator's Manual *145*
CAHSEE Online *146*
Test Booklets *146*
STUDENTS WITH DISABILITIES 146
Identification of Students with Test Variations, Accommodations or Modifications *147*
Scoring *147*
DEMOGRAPHIC DATA CORRECTIONS 147
TESTING IRREGULARITIES 147
TEST ADMINISTRATION INCIDENTS 148

CHAPTER 6: ANALYSES 149

SAMPLES USED FOR THE ANALYSES 150
CLASSICAL ITEM ANALYSES 151

Summary of Item Statistics *154*
Procedures for Documenting Items That Fail to Meet the Desired Psychometric Criteria *155*
DIFFERENTIAL ITEM FUNCTIONING ANALYSES 155
ITEM RESPONSE THEORY ANALYSES (CALIBRATION, SCALING AND EQUATING) 158
Measurement Model *158*
Item Calibration and Scaling *158*
True-Score Equating *159*
Equating Braille, Large-Print and Audio CD Forms *160*
Raw-Score to Scale-Score Conversion Tables and Conditional Standard Errors of Measurement *160*
Scaling Field-Test Items *161*
Item Response Theory Model-Data Fit Analyses *161*
Summaries of Scaled Item Response Theory B-Values *162*
RELIABILITY ANALYSES 162
Test Reliabilities and Standard Errors of Measurement *162*
Strand Intercorrelations, Reliabilities and Standard Errors of Measurement *164*
Subgroup Reliabilities and Standard Errors of Measurement *164*
Writing Prompt and Rater Agreement Summary *165*
Decision Classification Analyses *166*
VALIDITY EVIDENCE 167
Test Purpose *168*
Constructs to Be Measured *168*
Intended Test Population *169*
Scores Generated and the Interpretations and Uses of These Scores *169*
Evidence Based on Content *169*
Evidence Based on Relations to Other Variables *172*
Evidence Based on Response Processes *174*
Evidence Based on Internal Structure *175*
Additional Validity Evidence *177*
Conclusions *177*
APPENDIX 6.A: CAHSEE ITEM REVIEW—DESCRIPTION AND EXAMPLES OF CLASSIFICATION CATEGORIES 179
APPENDIX 6.B: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—JULY 2011 182
APPENDIX 6.C: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—OCTOBER 2011 187
APPENDIX 6.D: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—NOVEMBER 2011 192
APPENDIX 6.E: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—DECEMBER 2011 197
APPENDIX 6.F: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—FEBRUARY 2012 202
APPENDIX 6.G: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MARCH 2012 207
APPENDIX 6.H: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MAY 2012 212
APPENDIX 6.I: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—FEBRUARY 2012 217
APPENDIX 6.J: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—MARCH 2012 223
APPENDIX 6.K: INTERCORRELATIONS, RELIABILITY ESTIMATES AND STANDARD ERRORS OF MEASUREMENT 233
APPENDIX 6.L: RATER AGREEMENT ANALYSES 242
APPENDIX 6.M: GENERALIZABILITY ANALYSES 249
APPENDIX 6.N: DECISION CLASSIFICATION RELIABILITY ANALYSES 251
APPENDIX 6.O: SCORING TABLES FOR OPERATIONAL AND SPECIAL TEST VERSIONS 265
APPENDIX 6.P: STANDARD ERRORS OF THETA BASED ON WEIGHTED RAW SCORES 285

CHAPTER 7: PERFORMANCE STANDARDS 286

BACKGROUND AND PROCEDURES 286
CAHSEE PASSING SCORES 287
RESULTS 288

CHAPTER 8: SCORING AND REPORTING 290

PROCEDURES FOR MAINTAINING AND RETRIEVING INDIVIDUAL SCORES 290

Scoring and Reporting Specifications *290*
Scanning and Scoring *291*
Types of Scores and Strand Scores 292
Raw Score *292*
Strand Score *292*
Scale Score *292*
Passing Scores *292*
Score for Writing Applications *292*
Score Verification Procedures 293
Scoring Key Verification Process *293*
Monitoring and Quality Control of Writing Scoring *294*
Quality Control in Raw-to-Scale Score Conversions *296*
Overview of Score Aggregation Procedures 297
Individual Scores *297*
Group Scores *297*
Reports to be Produced and Scores for Each Report 298
Types of Score Reports *299*
Score Report Contents *299*
Score Report Applications *300*
Criteria for Interpreting Test Scores 300
Criteria for Interpreting Score Reports 300
Appendix 8.A: Frequency Distributions and Demographic Summaries—July 2011 302
Appendix 8.B: Frequency Distributions and Demographic Summaries—October 2011 310
Appendix 8.C: Frequency Distributions and Demographic Summaries—November 2011 318
Appendix 8.D: Frequency Distributions and Demographic Summaries—December 2011 326
Appendix 8.E: Frequency Distributions and Demographic Summaries—February 2012 334
Appendix 8.F: Frequency Distributions and Demographic Summaries—March 2012 342
Appendix 8.G: Frequency Distributions and Demographic Summaries—May 2012 350

CHAPTER 9: QUALITY CONTROL PROCEDURES 358

Quality Control of Item Development 358
Item Specifications *358*
Item Writers *358*
Internal Contractor Reviews *358*
Content Expert Reviews *359*
Statewide Pupil Assessment Review Panel Review *359*
Data Review of Field-Tested Items *360*
Quality Control of the Item Bank 360
Quality Control of Test Materials 361
Collecting Test Materials *361*
Processing Test Materials *361*
Quality Control of Scanning 361
Post-Scanning Edits *362*
Quality Control of Image Editing 362
Quality Control of Answer Document Processing and Scoring 363
Processing of Answer Documents *363*
Scoring and Reporting Specifications *363*
Matching Information on CAHSEE Answer Documents *363*
Matching Multiple-Choice and Writing Scores for English-Language Arts *363*
Storing Answer Documents *363*
Quality Control of Psychometric Processes 364
Scoring Key Verification Process *364*
Quality Control of Item Analyses, Differential Item Functioning and Equating Process *364*

Score Verification Process *365*
Offloads to Test Development *365*
Independent Evaluation of the CAHSEE Program *365*
QUALITY CONTROL OF REPORTING 365
Excluding Student Scores from Summary Reports *366*

CHAPTER 10: HISTORICAL COMPARISONS 367

EXAMINEE PERFORMANCE 367
TEST CHARACTERISTICS 367
APPENDIX 10.A—HISTORICAL COMPARISONS ON STUDENT PERFORMANCE 369
APPENDIX 10.B—HISTORICAL COMPARISONS ON TEST CHARACTERISTICS 375

REFERENCES 377

List of Tables

Table E.1: Summary of Examinees Tested for Each Administration by Content....1
Table E.2: Summary of Passing Rates by Content Area and Test Administration....2
Table E.3.1: CAHSEE Summary Statistics—English-Language Arts (July and October 2011)....3
Table E.3.2: CAHSEE Summary Statistics—English-Language Arts (November and December 2011, February 2012)....4
Table E.3.3: CAHSEE Summary Statistics—English-Language Arts (March and May 2012)....5
Table E.4.1: CAHSEE Summary Statistics—Mathematics (July and October 2011)....6
Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2011, February 2012)....7
Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2012)....8
Table 1.1: Testing Date for Each Administration by Content: 2011–12....11
Table 2.1: Items Included in the 2011–12 Administrations....18
Table 2.2: English-Language Arts and Mathematics Strands....19
Table 2.3: Subgroup Definitions....22
Table 2.4: CAHSEE Modification/Accommodation Table: 2011–12....24
Table 2.5: Listing of Tables–Summary Statistics for Testing Variations and Disability....26
Table 2.A.1: Summary Statistics by Testing Variations and Disability—ELA, July 2011....29
Table 2.A.2: Summary Statistics by Testing Variations and Disability—Mathematics, July 2011....30
Table 2.A.3: Demographic Summary for All Examinees by Testing Variations—ELA, July 2011....31
Table 2.A.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, July 2011....32
Table 2.A.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, July 2011....33
Table 2.A.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, July 2011....34
Table 2.A.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, July 2011....35
Table 2.A.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, July 2011....36
Table 2.A.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, July 2011....37
Table 2.A.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, July 2011....38
Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2011....39
Table 2.B.2: Summary Statistics by Testing Variations and Disability—Mathematics, October 2011....40
Table 2.B.3: Demographic Summary for All Examinees by Testing Variations—ELA, October 2011....41
Table 2.B.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, October 2011....42
Table 2.B.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, October 2011....43
Table 2.B.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, October 2011....44
Table 2.B.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, October 2011....45
Table 2.B.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, October 2011....47
Table 2.B.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, October 2011....49
Table 2.B.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, October 2011....51
Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2011....53
Table 2.C.2: Summary Statistics by Testing Variations and Disability—Mathematics, November 2011....54
Table 2.C.3: Demographic Summary for All Examinees by Testing Variations—ELA, November 2011....55
Table 2.C.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, November 2011....56
Table 2.C.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, November 2011....57
Table 2.C.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, November 2011....58
Table 2.C.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, November 2011....59
Table 2.C.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, November 2011....62
Table 2.C.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, November 2011....64
Table 2.C.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, November 2011....66
Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2011....68
Table 2.D.2: Summary Statistics by Testing Variations and Disability—Mathematics, December 2011....68
Table 2.D.3: Demographic Summary for All Examinees by Testing Variations—ELA, December 2011....69
Table 2.D.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, December 2011....70
Table 2.D.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, December 2011....71

Table 2.D.6: Percentiles of Scale Scores for Students with Testing Variations— Mathematics, December 2011 72
Table 2.D.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, December 2011 73
Table 2.D.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, December 2011 73
Table 2.D.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, December 2011 74
Table 2.D.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, December 2011 74
Table 2.E.1: Summary Statistics by Testing Variations and Disability—ELA, February 2012 75
Table 2.E.2: Summary Statistics by Testing Variations and Disability—Mathematics, February 2012.....76
Table 2.E.3: Demographic Summary for All Examinees by Testing Variations—ELA, February 2012.....77
Table 2.E.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, February 2012 78
Table 2.E.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, February 2012 79
Table 2.E.6: Percentiles of Scale Scores for Students with Testing Variations — Mathematics, February 2012.....80
Table 2.E.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, February 2012 81
Table 2.E.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, February 2012.....83
Table 2.E.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, February 2012.....85
Table 2.E.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, February 2012 87
Table 2.F.1: Summary Statistics by Testing Variations and Disability—ELA, March 2012.....89
Table 2.F.2: Summary Statistics by Testing Variations and Disability—Mathematics, March 2012 90
Table 2.F.3: Demographic Summary for All Examinees by Testing Variations—ELA, March 2012 91
Table 2.F.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, March 2012.....92
Table 2.F.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, March 2012.....93
Table 2.F.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, March 2012 94
Table 2.F.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, March 2012.95
Table 2.F.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, March 2012.....98
Table 2.F.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, March 2012.....101
Table 2.F.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, March 2012.....103
Table 2.G.1: Summary Statistics by Testing Variations and Disability—ELA, May 2012 105
Table 2.G.2: Summary Statistics by Testing Variations and Disability—Mathematics, May 2012.....106
Table 2.G.3: Demographic Summary for All Examinees by Testing Variations—ELA, May 2012.....107
Table 2.G.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, May 2012 108
Table 2.G.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, May 2012 109
Table 2.G.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, May 2012.....110
Table 2.G.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, May 2012 .111
Table 2.G.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, May 2012.....113
Table 2.G.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, May 2012.....115
Table 2.G.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, May 2012.....117
Table 3.1 Field-Test Percentages for the CAHSEE 120
Table 3.2 Summary of Items and Forms Presented in 2012 CAHSEE 129
Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items.....135
Table 4.2: Number of Items and Mean B-Value Ranges by Content Strand for ELA 135
Table 4.3: Difficulty (B) and Discrimination (R-bis) Specifications for Mathematics 136
Table 4.4: Number of Items and Mean B-Value Ranges by Content Strand for Mathematics.....136
Table 6.1: Listing of Summary Tables for Items.....150
Table 6.2: Flagging Criteria for Classical Item Analyses.....155
Table 6.3: DIF Categories.....157
Table 6.4: Field-Test Items Recommended for Future Forms.....162
Table 6.B.1: Summary of Operational Item Statistics—ELA 182
Table 6.B.2: Summary of Operational Item Statistics—Mathematics.....182
Table 6.B.3: IRT Model Data Fit Distribution of Operational Items—ELA 183

Table 6.B.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 183
Table 6.B.5: Operational Items Containing Significant DIF 183
Table 6.B.6: Distribution of Operational Item DIF Classifications—ELA 184
Table 6.B.7: Distribution of Operational Item DIF Classifications—Mathematics 185
Table 6.B.8: Listing of CR Item Statistics—ELA 186
Table 6.C.1: Summary of Operational Item Statistics—ELA 187
Table 6.C.2: Summary of Operational Item Statistics—Mathematics 187
Table 6.C.3: IRT Model Data Fit Distribution of Operational Items—ELA 188
Table 6.C.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 188
Table 6.C.5: Operational Items Containing Significant DIF 188
Table 6.C.6: Distribution of Operational Item DIF Classifications—ELA 189
Table 6.C.7: Distribution of Operational Item DIF Classifications—Mathematics 190
Table 6.C.8: Listing of CR Item Statistics—ELA 191
Table 6.D.1: Summary of Operational Item Statistics—ELA 192
Table 6.D.2: Summary of Operational Item Statistics—Mathematics 192
Table 6.D.3: IRT Model Data Fit Distribution of Operational Items—ELA 193
Table 6.D.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 193
Table 6.D.5: Operational Items Containing Significant DIF 193
Table 6.D.6: Distribution of Operational Item DIF Classifications—ELA 194
Table 6.D.7: Distribution of Operational Item DIF Classifications—Mathematics 195
Table 6.D.8: Listing of CR Item Statistics—ELA 196
Table 6.E.1: Summary of Operational Item Statistics—ELA 197
Table 6.E.2: Summary of Operational Item Statistics—Mathematics 197
Table 6.E.3: IRT Model Data Fit Distribution of Operational Items—ELA 198
Table 6.E.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 198
Table 6.E.5: Operational Items Containing Significant DIF 198
Table 6.E.6: Distribution of Operational Item DIF Classifications—ELA 199
Table 6.E.7: Distribution of Operational Item DIF Classifications—Mathematics 200
Table 6.E.8: Listing of CR Item Statistics—ELA 201
Table 6.F.1: Summary of Operational Item Statistics—ELA 202
Table 6.F.2: Summary of Operational Item Statistics—Mathematics 202
Table 6.F.3: IRT Model Data Fit Distribution of Operational Items—ELA 203
Table 6.F.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 203
Table 6.F.5: Operational Items Containing Significant DIF 203
Table 6.F.6: Distribution of Operational Item DIF Classifications—ELA 204
Table 6.F.7: Distribution of Operational Item DIF Classifications—Mathematics 205
Table 6.F.8: Listing of CR Item Statistics—ELA 206
Table 6.G.1: Summary of Operational Item Statistics—ELA 207
Table 6.G.2: Summary of Operational Item Statistics—Mathematics 207
Table 6.G.3: IRT Model Data Fit Distribution of Operational Items—ELA 208
Table 6.G.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 208
Table 6.G.5: Operational Items Containing Significant DIF 208
Table 6.G.6: Distribution of Operational Item DIF Classifications—ELA 209
Table 6.G.7: Distribution of Operational Item DIF Classifications—Mathematics 210
Table 6.G.8: Listing of CR Item Statistics—ELA 211
Table 6.H.1: Summary of Operational Item Statistics—ELA 212
Table 6.H.2: Summary of Operational Item Statistics—Mathematics 212
Table 6.H.3: IRT Model Data Fit Distribution of Operational Items—ELA 213
Table 6.H.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 213
Table 6.H.5: Operational Items Containing Significant DIF 213
Table 6.H.6: Distribution of Operational Item DIF Classifications—ELA 214
Table 6.H.7: Distribution of Operational Item DIF Classifications—Mathematics 215
Table 6.H.8: Listing of CR Item Statistics—ELA 216
Table 6.I.1: Summary of Field-Test Item Statistics—ELA 217
Table 6.I.2: Summary of Field-Test Item Statistics—Mathematics 217
Table 6.I.3: IRT Model Data Fit Distribution of Field-Test Items—ELA 218
Table 6.I.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics 218
Table 6.I.5: Field-Test Items Containing Significant DIF 219
Table 6.I.6: Distribution of Field-Test Item DIF Classifications—ELA 221
Table 6.I.7: Distribution of Field-Test Item DIF Classifications—Mathematics 222
Table 6.J.1: Summary of Field-Test Item Statistics—ELA 223
Table 6.J.2: Summary of Field-Test Item Statistics—Mathematics 223

Table 6.J.3: IRT Model Data Fit Distribution of Field-Test Items—ELA....224
Table 6.J.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics....224
Table 6.J.5: Field-Test Items Containing Significant DIF....225
Table 6.J.6: Distribution of Field-Test Item DIF Classifications—ELA....231
Table 6.J.7: Distribution of Field-Test Item DIF Classifications—Mathematics....232
Table 6.K.1: Intercorrelations and Reliability Estimates by Section—July 2011....233
Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2011....234
Table 6.K.3: Intercorrelations and Reliability Estimates by Section—November 2011....235
Table 6.K.4: Intercorrelations and Reliability Estimates by Section—December 2011....236
Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 2012....237
Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 2012....238
Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 2012....239
Table 6.K.8: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—February 2012....240
Table 6.K.9: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—February 2012....240
Table 6.K.10: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—March 2012....241
Table 6.K.11: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—March 2012....241
Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2011....242
Table 6.L.2: Summary Statistics for the ELA Essay Item—July 2010....242
Table 6.L.3: Agreement of First and Second Ratings on the ELA Essay Item—October 2011....243
Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2010....243
Table 6.L.5: Agreement of First and Second Ratings on the ELA Essay Item—November 2011....244
Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2010....244
Table 6.L.7: Agreement of First and Second Ratings on the ELA Essay Item—December 2011....245
Table 6.L.8: Summary Statistics for the ELA Essay Item—December 2010....245
Table 6.L.9: Agreement of First and Second Ratings on the ELA Essay Item—February 2012....246
Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2011....246
Table 6.L.11: Agreement of First and Second Ratings on the ELA Essay Item—March 2012....247
Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2011....247
Table 6.L.13: Agreement of First and Second Ratings on the ELA Essay Item—May 2012....248
Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2011....248
Table 6.M.1: Generalizability Results—July 2011....249
Table 6.M.2: Generalizability Results—October 2011....249
Table 6.M.3: Generalizability Results—November 2011....249
Table 6.M.4: Generalizability Results—December 2011....249
Table 6.M.5: Generalizability Results—February 2012....250
Table 6.M.6: Generalizability Results—March 2012....250
Table 6.M.7: Generalizability Results—May 2012....250
Table 6.N.1: ESEA Reliability Classifications—July 2011....251
Table 6.N.2: Pass/Not Pass Classifications—July 2011....252
Table 6.N.3: ESEA Reliability Classifications—October 2011....253
Table 6.N.4: Pass/Not Pass Classifications—October 2011....254
Table 6.N.5: ESEA Reliability Classifications—November 2011....255
Table 6.N.6: Pass/Not Pass Classifications—November 2011....256
Table 6.N.7: ESEA Reliability Classifications—December 2011....257
Table 6.N.8: Pass/Not Pass Classifications—December 2011....258
Table 6.N.9: ESEA Reliability Classifications—February 2012....259
Table 6.N.10: Pass/Not Pass Classifications—February 2012....260
Table 6.N.11: ESEA Reliability Classifications—March 2012....261
Table 6.N.12: Pass/Not Pass Classifications—March 2012....262
Table 6.N.13: ESEA Reliability Classifications—May 2012....263
Table 6.N.14: Pass/Not Pass Classifications—May 2012....264
Table 6.O.1: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, July 2011....265
Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2011....266
Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2011....267
Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2011....268
Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2011....269
Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2011....270
Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2011....271

Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2011 ..272
Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2012.....273
Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2012 ..274
Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2012 275
Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2012.....276
Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2012 277
Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2012.....278
Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA— July, October and November 2011 279
Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July, October and November 2011 280
Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—December 2011 and February 2012 281
Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— December 2011 and February 2012.....282
Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—March and May 2012 283
Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—March and May 2012.....284
Table 7.1: Passing Scores on the February 2004 CAHSEE 287
Table 7.2: Summary of Cut Scores and Passing Rates Based on Equating Samples 289
Table 8.1: Listing of Frequency Distribution and Demographic Summary Tables.....298
Table 8.2: Types of CAHSEE Reports 299
Table 8.A.1: Frequency Distributions, ELA—July 2011.....302
Table 8.A.2: Frequency Distributions, Mathematics—July 2011 302
Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2011.....303
Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2011 303
Table 8.A.5: Demographic Summary for All Examinees ELA—July 2011.....304
Table 8.A.6: Demographic Summary for All Examinees, Mathematics—July 2011 305
Table 8.A.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA— July 2011.....306
Table 8.A.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—July 2011 307
Table 8.A.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—July 2011.....308
Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2011.309
Table 8.B.1: Frequency Distributions, ELA—October 2011 310
Table 8.B.2: Frequency Distributions, Mathematics—October 2011.....310
Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2011.....311
Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2011 311
Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2011 312
Table 8.B.6: Demographic Summary for All Examinees, Mathematics—October 2011.....313
Table 8.B.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—October 2011 314
Table 8.B.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—October 2011.....315
Table 8.B.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—October 2011 316
Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2011 317
Table 8.C.1: Frequency Distributions, ELA—November 2011 318
Table 8.C.2: Frequency Distributions, Mathematics—November 2011.....318
Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2011.....319
Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2011 319
Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2011 320
Table 8.C.6: Demographic Summary for All Examinees, Mathematics—November 2011.....321
Table 8.C.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—November 2011 322
Table 8.C.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—November 2011.....323
Table 8.C.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—November 2011.....324
Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2011.....325
Table 8.D.1: Frequency Distributions, ELA—December 2011 326
Table 8.D.2: Frequency Distributions, Mathematics—December 2011.....326
Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2011.....327
Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2011 327
Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2011 328
Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2011.....329
Table 8.D.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA —December 2011 330
Table 8.D.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—December 2011.....331
Table 8.D.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—December 2011.....332
Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2011.....333

Table 8.E.1: Frequency Distributions, ELA—February 2012 334
Table 8.E.2: Frequency Distributions, Mathematics—February 2012 334
Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2012 335
Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2012 335
Table 8.E.5: Demographic Summary for All Examinees, ELA February 2012 336
Table 8.E.6: Demographic Summary for All Examinees, Mathematics—February 2012 337
Table 8.E.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—February 2012 338
Table 8.E.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— February 2012 339
Table 8.E.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—February 2012 340
Table 8.E.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—February 2012 341
Table 8.F.1: Frequency Distributions, ELA—March 2012 342
Table 8.F.2: Frequency Distributions, Mathematics—March 2012 342
Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2012 343
Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2012 343
Table 8.F.5: Demographic Summary for All Examinees, ELA—March 2012 344
Table 8.F.6: Demographic Summary for All Examinees, Mathematics—March 2012 345
Table 8.F.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—March 2012 346
Table 8.F.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—March 2012 347
Table 8.F.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—March 2012 348
Table 8.F.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—March 2012 349
Table 8.G.1: Frequency Distributions, ELA—May 2012 350
Table 8.G.2: Frequency Distributions, Mathematics—May 2012 350
Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2012 351
Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2012 351
Table 8.G.5: Demographic Summary for All Examinees, ELA—May 2012 352
Table 8.G.6: Demographic Summary for All Examinees, Mathematics—May 2012 353
Table 8.G.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—May 2012 354
Table 8.G.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— May 2012 355
Table 8.G.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—May 2012 356
Table 8.G.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—May 2012 357
Table 10.A.1: Number of Examinees Tested, Scale Score Means, and Standard Deviations of the CAHSEE Across 2009–10, 2010–11 and 2011–12 369
Table 10.A.2: Percentage of Students Passing Each Content Area Across 2009–10, 2010–11 and 2011–12 369
Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced Across 2009–10, 2010–11 and 2011–12 370
Table 10.A.4: Scale Score Distributions Across 2009, 2010 and 2011 for ELA (July to December) 371
Table 10.A.5: Scale Score Distributions Across 2010, 2011 and 2012 for ELA (February to May) 372
Table 10.A.6: Scale Score Distributions Across 2009, 2010 and 2011 for Mathematics (July to December) 373
Table 10.A.7: Scale Score Distributions Across 2010, 2011 and 2012 for Mathematics (February to May) 374
Table 10.B.1: Average Proportion Correct of Operational Test Items Across 2009–10, 2010–11 and 2011–12 375
Table 10.B.2: Average IRT b-values of Operational Test Items Across 2009–10, 2010–11 and 2011–12 375
Table 10.B.3: Average Point-Biserial Correlation of Operational Test Items Across 2009–10, 2010–11 and 2011–12 376
Table 10.B.4: Reliabilities and Standard Errors of Measurement (SEMs) of Operational Test Forms Across 2009–10, 2010–11 and 2011–12 376

List of Figures

Figure 3.1 The ETS Item Development Process for the CAHSEE 118
Figure 4.1: Example of Test Characteristic Curves 134
Figure 6.1: CAHSEE Item Analysis — Mathematics Item 154
Figure 6.2: CAHSEE Item Analysis — ELA Item 154
Figure 6.3: Samples of CAHSEE Item-Fit Rating Categories 181

This page intentionally left blank.

Executive Summary

The California High School Exit Examination (CAHSEE) has been developed by the California Department of Education (CDE) to assess achievement of content standards for English-language arts (ELA) and mathematics set by the California State Board of Education (SBE). The CAHSEE is administered seven times per year between July and May to allow several testing opportunities for those students who have not yet passed the examination. This report covers the seven CAHSEE test administrations given in July, October, November, and December 2011, and February, March, and May 2012.

Each test form was constructed from items that had been previously administered and placed onto the operational scale. There were 80 operational multiple-choice (MC) items in each mathematics form and 72 MC and one constructed-response (CR) item in each ELA form. Each test form also included a set of anchor items[1] that were used to maintain the operational scale across administrations. All items included on operational test forms had been evaluated for bias and sensitivity and for alignment of the content standards. In addition, each test form was reviewed and approved by the CDE.

Table E.1 presents the administration dates and the total number of examinees taking one or both CAHSEE content areas during the July, October, November, and December 2011, and February, March, and May 2012 administrations. The majority of examinees in February and March were first-time examinees. Grade ten students are only allowed to take the test in the February, March, or May administrations. Because students are allowed to take either ELA or mathematics in separate administrations, not all students took the ELA and mathematics examinations in a single administration. Examinees taking only one content area mostly were repeat examinees who did not pass that content area during a previous administration.

Table E.1: Summary of Examinees Tested for Each Administration by Content

Administration	Total Examinees[1]	ELA & Mathematics	ELA Only	Mathematics Only
July 2011	11,095	2,981	4,185	3,929
October 2011	51,981	20,042	16,301	15,638
November 2011	132,316	54,653	38,948	38,715
December 2011	3,583	823	1,496	1,264
February 2012	178,506	132,947	24,325	21,234
March 2012	413,449	358,051	30,053	25,345
May 2012	56,144	19,390	19,776	16,978

[1] Total number of examinees consists of examinees taking both ELA and mathematics, and either ELA or mathematics for each administration. These examinees include only students who received Passed or Not Passed status, and do not include students who took the modifications, were absent, previously passed, or did not attempt the examination.

[1] Anchor items, also called linking items, are used to link the scores on the current administration's test form to scores obtained on the base forms to adjust for the difficulty level of the forms across administrations. This is accomplished during the equating process, as discussed in Chapters 2 and 6.

The passing rates for all students taking each content area by test administration are presented in Table E.2.

Table E.2: Summary of Passing Rates by Content Area and Test Administration

Administration	English-Language Arts		Mathematics	
	N Tested	N Passed (%)	N Tested	N Passed (%)
July 2011	7,166	1,944 (27)	6,910	1,867 (27)
October 2011	36,343	13,185 (36)	35,680	13,541 (38)
November 2011	93,601	36,586 (39)	93,368	38,841 (42)
December 2011	2,319	662 (29)	2,087	718 (34)
February 2012	157,272	107,570 (68)	154,181	107,023 (69)
March 2012	388,104	300,952 (78)	383,396	302,193 (79)
May 2012	39,166	11,369 (29)	36,368	12,553 (35)

Educational Testing Service (ETS) conducted a series of statistical analyses to evaluate the items and test forms for each CAHSEE administration. The following analyses were completed for each administration: classical item analyses, differential item functioning (DIF), item response theory (IRT) calibration, scaling, and equating. Scoring tables were also created for each operational test form and reliability indices were calculated. This technical report also includes results from the following studies: inter-rater agreement and generalizability for the ELA CR items, and decision accuracy and consistency for the Pass/Not Pass and proficiency level classifications. Additional summary analyses conducted for students having special accommodation needs are included in this report.

All item analyses, including calibration, equating, and scaling, were completed using the Generalized Analysis System (GENASYS; ETS proprietary software) or commercially available software (e.g., SAS, SPSS, and GENOVA). In all cases, analyses were conducted on valid cases in each content area (e.g., students must have attempted at least the first 5 items on the test form). Individuals who entered invalid form numbers, left fields blank, or double-marked fields were excluded for the purposes of the analyses reported in this technical report. Students who left sections blank were excluded from the equating samples. Summary information for students who tested without modifications is presented in the Executive Summary and Chapter 8, and summary information for students who tested with modifications is presented in Chapter 2.

Highlights of the results for 2011–12 CAHSEE administrations included in this report are presented in Tables E.3.1-3.3 and E.4.1-4.3 on the following pages. These statistics indicate satisfactory psychometric properties of the test form constructed for these examinations.

Table E.3.1: CAHSEE Summary Statistics—English-Language Arts (July and October 2011)

Administration	July 2011	October 2011
Scale Score Information		
Number of Examinees	7,166	36,343
Mean	337	341
SD[1]	25	30
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	336	340
Raw Score Information		
Number of Examinees	7,166	36,343
Mean	47.07	49.83
SD	11.98	13.90
Possible Range	0-90	0-90
Obtained Range	1-89	0-90
Median	47	50
Test Information		
Reliability	0.87	0.90
Raw Score Standard Error of Measurement	4.40	4.37
Mean Omits	0.45	0.72
SD Omits	3.98	5.21
Percentage Responding to:		
All Items	90	89
All Items - 1 Item	97	97
All Items - 2 Items	98	98
All Items - 3 Items	99	98
All Items - 4 Items	99	98
All Items - 5 Items	99	98
Item Information[2]		
Number of Items	72	72
Mean Observed Average Item Score (AIS)	0.53	0.56
Equated Mean Rasch B-Value	-0.01	-0.05
Mean R-biserial	0.40	0.46
SD R-biserial	0.09	0.09
Examinee Information		
ELA-Only Examinees		
Number of Examinees	4,185	16,301
Mean Scale Score	337	341
SD Scale Score	21	24
Median Scale Score	338	342
ELA and Mathematics Examinees		
Number of Examinees	2,981	20,042
Mean Scale Score	336	342
SD Scale Score	30	34
Median Scale Score	334	338

[1] Standard deviation.

[2] Means and standard deviations for the item information section are computed on 72 MC items.

Table E.3.2: CAHSEE Summary Statistics—English-Language Arts (November and December 2011, February 2012)

Administration	November 2011	December 2011	February 2012
Scale Score Information			
Number of Examinees	93,601	2,319	157,272
Mean	344	339	371
SD	32	26	41
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	343	339	372
Raw Score Information			
Number of Examinees	93,601	2,319	157,272
Mean	51.07	48.41	60.87
SD	14.68	11.97	16.61
Possible Range	0-90	0-90	0-90
Obtained Range	0-90	0-89	0-90
Median	52	49	64
Test Information			
Reliability	0.91	0.87	0.94
Raw Score Standard Error of Measurement	4.36	4.25	4.05
Mean Omits	0.74	0.27	0.41
SD Omits	5.33	3.00	3.80
Percentage Responding to:			
All Items	88	92	92
All Items - 1 Item	97	98	98
All Items - 2 Items	98	99	99
All Items - 3 Items	98	100	99
All Items - 4 Items	98	100	99
All Items - 5 Items	98	100	99
Item Information[1]			
Number of Items	72	72	72
Mean Observed Average Item Score (AIS)	0.58	0.54	0.70
Equated Mean Rasch B-Value	-0.09	-0.04	-0.01
Mean R-biserial	0.48	0.41	0.56
SD R-biserial	0.08	0.11	0.09
Examinee Information			
ELA-Only Examinees			
Number of Examinees	38,948	1,496	24,325
Mean Scale Score	344	339	338
SD Scale Score	26	23	27
Median Scale Score	345	339	338
ELA and Mathematics Examinees			
Number of Examinees	54,653	823	132,947
Mean Scale Score	344	338	377
SD Scale Score	36	31	40
Median Scale Score	341	335	379

[1] Means and standard deviations for the item information section are computed on 72 MC items.

Table E.3.3: CAHSEE Summary Statistics—English-Language Arts (March and May 2012)

Administration	March 2012	May 2012
Scale Score Information		
Number of Examinees	388,104	39,166
Mean	378	337
SD	39	34
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	382	334
Raw Score Information		
Number of Examinees	388,104	39,166
Mean	65.23	48.58
SD	14.63	15.87
Possible Range	0-90	0-90
Obtained Range	0-90	0-90
Median	69	48
Test Information		
Reliability	0.93	0.92
Raw Score Standard Error of Measurement	3.81	4.48
Mean Omits	0.28	1.02
SD Omits	3.05	6.53
Percentage Responding to:		
All Items	94	88
All Items - 1 Item	98	96
All Items - 2 Items	99	97
All Items - 3 Items	99	98
All Items - 4 Items	99	98
All Items - 5 Items	99	98
Item Information[1]		
Number of Items	72	72
Mean Observed Average Item Score (AIS)	0.76	0.55
Equated Mean Rasch B-Value	-0.16	-0.11
Mean R-biserial	0.54	0.50
SD R-biserial	0.08	0.08
Examinee Information		
ELA-Only Examinees		
Number of Examinees	30,053	19,776
Mean Scale Score	341	337
SD Scale Score	31	31
Median Scale Score	342	336
ELA and Mathematics Examinees		
Number of Examinees	358,051	19,390
Mean Scale Score	382	337
SD Scale Score	38	37
Median Scale Score	385	332

[1] Means and standard deviations for the item information section are computed on 72 MC items.

Table E.4.1: CAHSEE Summary Statistics—Mathematics (July and October 2011)

Administration	July 2011	October 2011
Scale Score Information		
Number of Examinees	6,910	35,680
Mean	339	345
SD	21	27
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	339	343
Raw Score Information		
Number of Examinees	6,910	35,680
Mean	37.08	38.81
Standard Deviation	10.57	12.98
Possible Range	0-80	0-80
Obtained Range	2-80	0-80
Median	37	38
Test Information		
Reliability	0.85	0.90
Raw Score Standard Error of Measurement	4.16	4.10
Mean Omits	0.45	0.59
SD Omits	3.42	4.19
Percentage Responding to:		
All Items	86	85
All Items - 1 Item	96	95
All Items - 2 Items	98	98
All Items - 3 Items	99	98
All Items - 4 Items	99	99
All Items - 5 Items	99	99
Item Information		
Number of Items	80	80
Mean Observed Average Item Score (AIS)	0.46	0.48
Equated Mean Rasch B-Value	-0.24	-0.19
Mean R-biserial	0.36	0.43
SD R-biserial	0.12	0.10
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	3,929	15,638
Mean Scale Score	341	345
SD Scale Score	17	20
Median Scale Score	341	344
ELA and Mathematics Examinees		
Number of Examinees	2,981	20,042
Mean Scale Score	337	345
SD Scale Score	25	31
Median Scale Score	334	339

Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2011, February 2012)

Administration	November 2011	December 2011	February 2012
Scale Score Information			
Number of Examinees	93,368	2,087	154,181
Mean	347	343	375
SD	27	21	40
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	344	341	370
Raw Score Information			
Number of Examinees	93,368	2,087	154,181
Mean	40.67	37.21	52.85
Standard Deviation	13.07	10.60	17.04
Possible Range	0-80	0-80	0-80
Obtained Range	0-80	2-79	0-80
Median	40	36	54
Test Information			
Reliability	0.90	0.85	0.95
Raw Score Standard Error of Measurement	4.10	4.14	3.67
Mean Omits	0.54	0.29	0.32
SD Omits	4.00	2.47	2.95
Percentage Responding to:			
All Items	85	87	90
All Items - 1 Item	96	97	97
All Items - 2 Items	98	99	99
All Items - 3 Items	99	99	99
All Items - 4 Items	99	99	99
All Items - 5 Items	99	100	99
Item Information			
Number of Items	80	80	80
Mean Observed Average Item Score (AIS)	0.51	0.46	0.66
Equated Mean Rasch B-Value	-0.24	-0.15	-0.24
Mean R-biserial	0.44	0.36	0.57
SD R-biserial	0.11	0.12	0.09
Examinee Information			
Mathematics-Only Examinees			
Number of Examinees	38,715	1,264	21,234
Mean Scale Score	347	344	344
SD Scale Score	21	19	23
Median Scale Score	348	344	343
ELA and Mathematics Examinees			
Number of Examinees	54,653	823	132,947
Mean Scale Score	346	341	380
SD Scale Score	31	24	41
Median Scale Score	341	337	379

Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2012)

Administration	March 2012	May 2012
Scale Score Information		
Number of Examinees	383,396	36,368
Mean	384	344
SD	39	30
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	383	340
Raw Score Information		
Number of Examinees	383,396	36,368
Mean	56.46	38.67
Standard Deviation	15.78	14.57
Possible Range	0-80	0-80
Obtained Range	0-80	0-80
Median	59	37
Test Information		
Reliability	0.95	0.92
Raw Score Standard Error of Measurement	3.56	4.05
Mean Omits	0.23	0.65
SD Omits	2.22	4.52
Percentage Responding to:		
All Items	91	86
All Items - 1 Item	98	96
All Items - 2 Items	99	98
All Items - 3 Items	99	98
All Items - 4 Items	100	99
All Items - 5 Items	100	99
Item Information		
Number of Items	80	80
Mean Observed Average Item Score (AIS)	0.71	0.48
Equated Mean Rasch B-Value	-0.20	-0.22
Mean R-biserial	0.55	0.48
SD R-biserial	0.09	0.10
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	25,345	16,978
Mean Scale Score	349	346
SD Scale Score	26	27
Median Scale Score	346	342
ELA and Mathematics Examinees		
Number of Examinees	358,051	19,390
Mean Scale Score	387	342
SD Scale Score	38	33
Median Scale Score	385	337

Chapter 1: Introduction

Background

The California Department of Education (CDE) initiated the development of the California High School Exit Examination (CAHSEE) to ensure that all students graduating from high school demonstrated competency with respect to the SBE content standards in reading, writing, and mathematics. The CAHSEE was first administered to ninth-graders on a voluntary basis in March and May 2001. Beginning October 1, 2001, the CDE awarded a contract to Educational Testing Service (ETS) and its subcontractors for the development and administration of the CAHSEE. ETS was awarded the new contract in July 2004 for three years and in August 2008 for four years. In 2011–12, the CAHSEE was administered seven times. Each administration took two days: English-language arts (ELA) on the first day and mathematics on the second. This report covers the seven administrations: July, October, November, and December 2011, and February, March, and May 2012.

Test Purpose

The primary purpose of the CAHSEE is to assess student achievement in public high schools and help ensure that students who graduate from public high schools can demonstrate competency in reading, writing, and mathematics. The CAHSEE's role in this process is to identify students who have not yet developed the academic competencies contained in the California content standards for ELA and mathematics, and to encourage districts to give these students the attention and resources needed to help them achieve these competencies during their high school years. All California public school students, except eligible students with disabilities (SWDs), must satisfy the CAHSEE requirement, as well as all other state and local requirements, to receive a high school diploma. The CAHSEE requirement can be satisfied by passing the examination, or for eligible SWDs, by meeting the exemption requirement pursuant to California *Education Code (EC)* Section 60852.3, or by receiving a local waiver pursuant to *EC* Section 60851(c). By definition, testing with a modification changes the construct of what is being tested. The results are reported as Modified rather than Passed or Not Passed. Under *EC* Section 60851(c), local school district governing boards may waive the requirement to pass the CAHSEE for SWDs who test with modifications and score 350 or above on one or both parts of the examination.

Content

The CAHSEE includes an examination in mathematics and an examination in ELA. Students may take either one or both examinations in a single administration. The multiple-choice (MC) portions of both examinations are scored such that one point is assigned for each correct answer. One test form was constructed for each operational administration consisting of available items from an item bank. In addition to the standard forms, Braille, large-print, and audio CD versions were also available. One

emergency form was also constructed to cover all administrations, although it was only necessary to use this form if test security was not maintained.

The ELA examination measures reading and writing skills as defined by the SBE standards through grade ten.[2] It includes both reading and writing components. The reading portion covers vocabulary and informational and literary reading. The writing portion covers writing strategies, applications, and conventions. The ELA examination consists of 72 MC questions and one constructed-response (CR) item. The CR item is a written response to a writing prompt.

The CAHSEE mathematics examination measures standards adopted by the SBE for grades six and seven mathematics and Algebra I.[3] It covers statistics, data analysis and probability, number sense, measurement and geometry, algebra, and mathematical reasoning. There are 80 operational questions in each mathematics form.

Target Population

The target population for the CAHSEE is students who are either enrolled in California public high schools in grade ten, eleven, or twelve, or are enrolled in adult schools. These students are working to attain a high school diploma and have not passed both the ELA and the mathematics portions of the CAHSEE.

Intended Use and Purpose of Test Scores

The results for the CAHSEE are used primarily to identify students who are not developing minimum competencies with respect to the standards in reading, writing, and mathematics that are contained in the California Content Standards. All California public school students, except eligible SWDs, must satisfy the CAHSEE requirement, as well as all other state and local requirements, to receive a high school diploma.

In addition, the state and federal governments use the CAHSEE results for grade ten as a measure of school and school district accountability. The state accountability program is the Public Schools Accountability Act; the federal accountability program is the Elementary and Secondary Education Act (ESEA). The use of the CAHSEE results for these accountability programs is intended to be independent of how the CAHSEE is used at the individual student level.

Schedule of Administrations and Participation Rules

The CAHSEE was administered seven times in the 2011–12 school year on dates that were determined by the State Superintendent of Public Instruction (SSPI). Schools must administer the CAHSEE on the designated dates, as shown in Table 1.1.

[2] The blueprints for the CAHSEE ELA examinations can be found on the CDE CAHSEE Program Resources Web page http://www.cde.ca.gov/ta/tg/hs/resources.asp.
[3] The blueprints for the CAHSEE Mathematics examinations can be found on the CDE CAHSEE Program Resources Web page http://www.cde.ca.gov/ta/tg/hs/resources.asp.

Table 1.1: Testing Date for Each Administration by Content: 2011–12

Administration	English-Language Arts (Tuesday)	Mathematics (Wednesday)
July 2011	July 26, 2011	July 27, 2011
October 2011	October 4, 2011	October 5, 2011
November 2011	November 1, 2011	November 2, 2011
December 2011	December 3, 2011[1]	December 10, 2011[1]
February 2012	February 7, 2012	February 8, 2012
March 2012	March 13, 2012	March 14, 2012
May 2012	May 8, 2012	May 9, 2012

[1]Saturday administration

Participation rules determine when and how many times a student may take the CAHSEE. The participation rules are:

- Grade ten students shall only be tested during the census administrations in February or March, or the make-up administration in May, as per state requirements.

- Grade eleven students may take the part(s) of the CAHSEE not previously passed up to two times per school year and may test in consecutive administrations (e.g., October and November). Grade 11 students cannot participate in the July administration.

- Grade twelve students may take the part(s) of the CAHSEE not previously passed at least three times and up to five times per school year and may test in consecutive administrations.

- Adult Education students may take the part(s) of the CAHSEE not previously passed up to three times per school year and may test in consecutive administrations.

- No student who has previously passed the CAHSEE is eligible to retake the exam.

Significant Developments in the 2011–12 School Year

There were no new or unusual changes in examination content and test administration in the 2011–12 school year. The test blueprint remained unchanged. As per federal guidelines, beginning with the July 2009 administration the demographic groups included a category called Two or More Races that was used for Adequate Yearly Progress (AYP) reporting. Equating and scoring methodologies were unchanged from the previous year.

Limitations of the Assessment

Score Interpretation

School districts use the CAHSEE results as part of the gateway to student graduation. However, it is important to remember that a single test can provide only limited information. Other relevant graduation requirements should be considered as well. It is also important to note that a student's CAHSEE score in a content area contains measurement error and could vary somewhat if the student were retested.

Groups and Organizations

State Board of Education

The State Board of Education (SBE) is the state education agency that sets education 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 2009, the SBE suspended the adoption of textbooks until 2013–2014.

The SBE is responsible for the maintenance of such programs as the ESEA for reporting results in terms of the Academic Performance Index (API), which measures the academic performance and growth of schools on a variety of academic measures.

California Department of Education

The CDE oversees California's public school system and is responsible for the education of more than seven million (7,000,000) children and young adults in more than 9,000 schools. The CDE's mission is to provide leadership, assistance, oversight, and resources so that every child in California has access to an educational system that meets world-class standards. As part of its mission to promote district and school accountability for improving student achievement as defined by the SBE, the CDE oversees the development and administration of the CAHSEE.

Test Contractors

Educational Testing Service

The CDE awarded a contract to ETS to develop and administer the CAHSEE program. As the prime contractor, ETS has overall responsibility to coordinate the work of its employees and its subcontractor, Pearson Educational Measurement, in order to fulfill all requirements of the contract. Activities conducted directly by ETS include:

- Overall management of the program.
- Development of all test items and test forms.
- Construction and production of test booklets and test-related materials.

- Scoring all responses, including performance scoring of the writing response.
- Production and distribution of all score reports, summary reports, and data files of test results.
- Support and training for all local education agencies (LEAs).
- Implementation and management of the CAHSEE Online System for ordering materials, pre-identification services, and data correction.
- Completion of all psychometric activities.
- Monitoring and managing the work of Pearson Educational Measurement, subcontractor for the CAHSEE program.

Pearson Educational Measurement

Pearson produces all scannable materials; packages, distributes, and retrieves test materials; accounts for all secure test materials; and scans all responses.

Overview of the Technical Report

This technical report describes the procedures applied to the CAHSEE for the four 2011 and the three 2012 administrations in the 2011–12 school year. This report also presents the results of statistical analyses based on the data from these administrations. The organization of the technical report demonstrates the process for accumulating evidence to support the validity of inferences made from the CAHSEE test scores, which is one of the most fundamental considerations in developing and evaluating tests. The process begins with the test design and continues throughout the entire assessment process, including content specifications, item development, psychometric quality, and inferences made from the results.

This technical report contains nine additional chapters as follows:

- Chapter 2 presents a conceptual overview of processes involved in a testing cycle for a CAHSEE administration. This includes test construction, test administration, generation of test scores, and dissemination of score reports.
- Chapter 3 describes the procedures followed in the development of valid CAHSEE items; the chapter explains the process of field testing new items and the review of items by contractors and content experts.
- Chapter 4 details the content and psychometric criteria applicable to the test assembly of the CAHSEE for the 2011–12 administrations.
- Chapter 5 presents the processes involved in the actual 2011–12 CAHSEE administrations, with an emphasis on efforts made to ensure standardization of

the tests. It also includes a detailed section that describes the procedures that were followed by ETS to ensure test security.

- Chapter 6 summarizes the results of the item-level analyses performed after each of the seven 2011–12 administrations. These include the classical item analyses, differential item functioning (DIF), item response theory (IRT), and model-fit analyses, as well as documentation of the equating along with CAHSEE score conversion tables. Also summarized in this chapter are the results of reliability analyses. These analyses include assessments of test reliability and the consistency and accuracy of the classifications, including the Pass/Not Pass classifications and the proficiency-level classifications. Finally, this chapter discusses the procedures designed to ensure the validity of the CAHSEE score use and interpretation.

- Chapter 7 describes the standard-setting process conducted for the CAHSEE when the SBE adopted new test blueprints in 2003. Also described is the procedure to establish the CAHSEE reporting scale. In addition, results describing students' proficiency classifications are also provided.

- Chapter 8 details the types of scores and score reports that are produced at the conclusion of each administration of the CAHSEE. Information about the distributions of scores, aggregated by subgroups based on demographics and the use of special services, is also included in this chapter.

- Chapter 9 highlights the importance of controlling and maintaining the quality of the CAHSEE.

- Chapter 10 presents historical comparisons of various item- and test-level results for the past three years and for the baseline year.

Each chapter contains summary tables in the body of the text. Extended appendices that give more detailed information are provided at the ends of the relevant chapters.

Chapter 2: Test Specifications

This chapter provides an overview of the processes involved in a typical test development and administration for the CAHSEE. The specifications maintained by ETS to conduct these processes are described here. This chapter is organized to provide a brief description of each major process followed by a summary of the associated specifications. More details about the specifications and the analyses associated with each process are described in the ensuing chapters that are referenced in the following sections.

Item Development

CAHSEE items are developed to measure California content standards and designed to conform to principles of item writing defined by ETS (ETS, 2002). ETS maintains item development specifications for the CAHSEE and has an Item Development Plan to guide the item writing for each content area. The emphasis in writing items is determined in consultation with the CDE.

The item development specifications describe the characteristics of the items written to measure each content standard. The item development specifications ensure that the items in any administration of the CAHSEE measure the content standards in the same way. This is achieved by providing detailed information to CAHSEE item writers.

The items selected for each CAHSEE administration undergo an extensive item review process that is designed to provide the best standards-based tests possible. Details about the item development specifications, the item development plan, the item review process, and field testing are presented in Chapter 3.

Item Formats

The CAHSEE tests contain four-option MC items. The ELA test also includes one CR item, which is a writing prompt that is polytomously scored.

Model for Generating Item Statistics

IRT is used to estimate item difficulty for the operational and field-test items. Items are calibrated using the Rasch model for the MC items and the Rasch partial-credit model for the ELA CR items. IRT expresses the probability that a student will achieve a certain score on an item (such as correct or incorrect) as a function of the item's statistical properties and the ability level of the student.

The fundamental equation of the Rasch or one-parameter logistic model (1PL) relates the probability that a person with ability θ will respond correctly to item i:

$$P_i(\theta) \equiv P_i(X_i = 1 \mid \theta) = \frac{1}{1 + \exp[-(\theta - b_i)]}$$

where:

X_i is the response to item *i*, 1 if correct and 0 if incorrect;

b_i is the threshold parameter of item *i*, characterizing its difficulty; and

θ is the ability level for an examinee.

The fundamental equation of the Rasch partial-credit model is the probability that a person with ability θ will obtain a score *of x* on item *i,* which is scored in score categories ordered from 0 to *m:*

$$P_{ix}(\theta) \equiv P(X_i = x \mid \theta) = \frac{\exp[\sum_{s=0}^{x}(\theta - b_i + d_{is})]}{\sum_{k=0}^{m_i}\exp[\sum_{s=0}^{k}(\theta - b_i + d_{is})]}$$

where:

X_i is the response to item *i*, with possible value 0, 1, …, m_i ;

b_i is the location parameter of item *i*, characterizing its difficulty;

d_{is} is the threshold parameter for score category *s*; and

θ is the ability level for an examinee.

A proprietary version of the PARSCALE computer program (Muraki & Bock, 1995) is used for all item calibration work. This program estimates parameters for both the three-parameter logistic model (3PL) and the generalized partial-credit model using procedures described by Muraki (1992). For CAHSEE forms, the PARSCALE program is constrained by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588) and by fixing the lower asymptote for all MC items to zero. The resulting estimation is equivalent to the Rasch model for MC items and the Rasch partial-credit model for CR items. Additional details of the item calibration procedures are found in Chapter 6.

Item Banking

The newly developed items are placed in the item bank along with the corresponding information obtained at the review sessions. Items that are accepted by the content experts are updated to a Field-test Ready status; items that are rejected are assigned to a Rejected Before Use status.

Items are field tested in census administrations (i.e., February and March) to obtain information concerning item performance and to obtain statistics that can be used to assemble operational forms. ETS identifies items flagged for high levels of DIF (C level) and the associated statistics for another round of review by content experts. Subsequent updates to items are based on the operational use of the items. The latest content of the item is retained in the bank at any time, along with the administration data from every administration that has included the item. The item statistics obtained from the latest census administration are used for test assembly.

ETS delivers the items to the CDE by means of a CAHSEE electronic item bank. Further details on item banking are presented in Chapter 3.

Item Refresh Rate and Released Test Questions

The Item Development Plan assumes that 70 percent of items on an operational ELA form and 70 percent of items on an operational mathematics form are refreshed each year; these items remain in the item bank for future use. Previously, a number of items were released to the public. The Released Test Questions (RTQs) show the content and types of questions that are included on the CAHSEE. Because RTQs are released and posted on the Internet for public viewing, they are precluded from future forms. Due to the state budget cuts, no ELA or mathematics RTQs were released in the 2011–12 school year.

Criteria for Selecting Released Test Questions

In selecting test questions for release, three criteria are used: (1) The questions adequately cover a selection of the content standards; (2) The questions demonstrate a range of difficulty; and (3) The questions represent a variety of ways in which students' achievements of the content standards can be assessed. These RTQs, however, do not reflect all the ways the standards may be assessed.

Test Assembly

The test assembly process adheres to rigorous standards for item selection and form construction based on alignment to the California content standards as well as psychometric criteria. This is a multi-faceted process that takes into account the factors described below.

Test Blueprint

ETS selects all CAHSEE test items to conform to the SBE-approved California content standards and test blueprints. The ELA examination measures reading and writing skills as defined by the SBE standards through grade ten. The CAHSEE mathematics examination measures standards adopted by the SBE through Algebra I. The content blueprints for the CAHSEE can be found on the CDE CAHSEE Program Resources Web page at http://www.cde.ca.gov/ta/tg/hs/resources.asp.

Test Length

The number of items on the CAHSEE varies by content area. The ELA test form consists of 73 operational items, including one writing prompt and seven field-test items for a total of 80 items. The mathematics test form consists of 80 operational items and 12 field-test items for a total of 92 items.

Number of Test Forms

In order to maintain the item pool for construction of future forms, multiple forms are administered, each containing a set of common operational items and unique MC field-test items. The ELA CR items are field tested separately from the operational CAHSEE administrations. Different forms contain the same operational items and different field-test items. Field-test data from only the February and March administrations are used to evaluate the quality of the field-test items. During the 2011-12 testing cycle, 104 ELA and 52 mathematics field-test forms were administered for the February administration and 155 ELA and 78 mathematics field-test forms were administered for the March administration. A summary of the items and forms included in the seven administrations is shown in Table 2.1.

Table 2.1: Items Included in the 2011–12 Administrations

Administration	Operational Items	Field-Test Forms	Number of Field-Test Items[1]
July 2011			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
October 2011			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
November 2011			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
December 2011			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
February 2012			
English–language arts	72 MC, 1 CR	104	694
Mathematics	80 MC	52	624
March 2012			
English–language arts	72 MC, 1 CR	155	1049
Mathematics	80 MC	78	936
May 2012			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12

[1] In some cases, the same field-test items may have appeared in more than one form.

Content and Process Categories

Although the test blueprints call for the number of items at the individual standard level, scores on the CAHSEE items are also grouped into sub-content areas, referred to as subscores or strands. For each strand, the number of questions correctly answered is reported on a student's score report. The ELA and mathematics strands reported for the CAHSEE are presented in Table 2.2.

Table 2.2: English-Language Arts and Mathematics Strands

ELA Content Strand	Number of Items	Mathematics Content Strand	Number of Items
Word Analysis (RW)	7	Probability & Statistics (PS)	13
Reading Comprehension (RC)	18	Number Sense (NS)	17
Literary Response & Analysis (RL)	20	Algebra & Functions (AF)	20
Writing Strategies (WS)	12	Measurement & Geometry (MG)	18
Writing Conventions (WC)	15	Algebra 1 (A1)	12
Writing Applications–CR (WA)	1		

Content Rules and Item Selection

When developing a new test for a given content area, test developers follow a number of rules. First and foremost, they select items that meet the blueprint for the content area. Using the electronic item bank, assessment specialists identify the two sets of linking items, which are chosen from the operational items of the census administrations in the previous year. The linking sets are used to equate the test forms for a given school year. Once the linking sets are approved, assessment specialists populate the rest of the test form. Another consideration is the difficulty of each item. Test developers strive to ensure that there are some easy and some hard items and that there are a large number of items in the middle range of difficulty. The detailed rules are presented in Chapter 4.

Psychometric Criteria

CAHSEE test developers and psychometricians strive to accomplish three goals while developing a test:

- The test must have the desired precision of measurement at all ability levels.
- The test score must be valid and reliable for the intended population and for the various subgroups of test takers.
- The test forms must be comparable across years of administrations to ensure that scores generalize over time.

In order to achieve these goals, a set of rules is developed that outlines the desired psychometric properties of the CAHSEE, referred to as statistical targets. Three types of assembly targets are developed for the test: the total test target, the linking set target, and content strand targets. These targets are provided to test developers before a test construction cycle begins. The test developers and psychometricians work together to design the tests according to these targets. The test targets used for the 2011–12 test development are presented in Tables 4.1 through 4.4 in Chapter 4.

Item Arrangement

The items in test forms are organized and sequenced differently depending on the requirements of the content area. Items are sequenced according to the reading passages in ELA, and items are sequenced according to strand in mathematics. Further details on the arrangement of items during test assembly are also described in Chapter 4.

Test Administration

It is of utmost priority to administer the CAHSEE in an appropriate, consistent, confidential, and standardized manner. The necessary measures are taken to ensure the standardization of the CAHSEE as described in this section.

Test Security and Confidentiality

All CAHSEE tests are secure documents. For the CAHSEE administration, every person having access to test materials maintains the security and confidentiality of the tests. ETS' Code of Ethics requires that all test information, including tangible materials (e.g., test booklets, test questions, test results), and confidential files, processes, and activities are kept secure. To ensure security for all the tests that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI). A description of the OTI and its mission is presented in Chapter 5.

In the pursuit of enforcing secure practices, ETS and the OTI strive to safeguard the various processes involved in a test development and administration cycle. The practices related to each process are discussed in detail in Chapter 5.

Procedures to Maintain Standardization

The CAHSEE processes are designed so that the tests are administered and scored in a standardized manner. The procedures implemented for the CAHSEE program are noted below.

Test Administrators

ETS employs personnel who facilitate various processes involved in the standardization of a CAHSEE administration cycle. Staff members at school districts who are central to the processes include the CAHSEE district coordinators, test examiners, proctors, and scribes. The responsibilities for each staff member are included in the *CAHSEE District and Test Site Coordinator's Manual* (*DTSCM*; CDE, 2011a), which is presented in more detail in Chapter 5.

Test Directions

ETS maintains a series of instructions, compiled in detailed manuals, that are available to the test administrators. These documents include the following test administration manuals:

Directions for Administration. The *Directions for Administration Manual* (*DFA*; CDE 2011b) includes directions to be read aloud to students as well as procedures the test administrator is to follow to assure that the validity of the test administration is maintained and the security of the test is protected.

Directions for Administration—Special Test Versions. The *DFA—Special Test Versions Manual* explains procedures concerning students using the Braille, large print, and audio CD test versions. The instructions mirror the standard administration, but reflect the appropriate changes needed to accommodate differences when a student uses a special test version.

District and Test Site Coordinator's Manual. The *DTSCM* provides specific information and forms appropriate to the respective roles carried out by the LEA and school coordinators.

Answer Documents

Pearson maintains strict security procedures when shipping, retrieving, and storing testing materials. The processing and warehouse facilities at Pearson, where answer documents are scanned, are secure and locked. All completed student answer documents are stored in the secure warehouse for a period of one year after the completion of each examination administration. ETS maintains an electronic file of all answer document images for one year following the review and approval process involving the CDE and ETS.

Demographic Distributions

CAHSEE results are presented by subgroup based on grade, gender, ethnicity, language fluency, economic disadvantage, and special education programs. Table 2.3 defines the demographic groups for which results are obtained. Students' economic statuses are determined by considering the education level of their parents and whether or not they are eligible to participate in the National School Lunch Program (NSLP).

Table 2.3: Subgroup Definitions

Subgroup	Definition
Grade	Tenth Eleventh Twelfth Adult Education
Gender	Male Female
Ethnicity	American Indian or Alaska Native Asian – Chinese – Japanese – Korean – Vietnamese – Asian Indian – Laotian – Cambodian – Hmong – Other Asian Pacific Islander – Native Hawaiian – Guamanian – Samoan – Tahitian – Other Pacific Islander Filipino Hispanic or Latino African American White (not Hispanic) Two or More Races
English Language Fluency	English-only (EO) Initially fluent English proficient (IFEP) English learner (EL) Reclassified fluent English proficient (RFEP)
Economic Status	Not economically disadvantaged Economically disadvantaged
Special Services	No special services Special services
Parent Education	Graduate school or post graduate training College graduate Some college High school graduate Not a high school graduate

Test Variations, Accommodations and Modifications

Grades ten, eleven, and twelve public high school students and adult education students who have not passed both the ELA and mathematics sections of the CAHSEE participate in the CAHSEE. Per the California *EC* Section 60852.3, SWDs are exempted from meeting the CAHSEE requirement until alternative means to the

CAHSEE are implemented. Many SWDs and English Learners (EL) take the CAHSEE under standard conditions; however, some of these students may need assistance when taking the CAHSEE. This assistance takes the form of test variations, accommodations, or modifications. All students may have test administration directions simplified or clarified. In addition, all eligible students may have test variations if these variations are regularly used in the classroom. They also must be allowed to use the accommodations and modifications that are specified in each student's individualized education program (IEP) or Section 504 plan. The accommodations and/or modifications must match the one(s) used for classroom work throughout the year.

The purpose of test variations, accommodations, and modifications is to enable the students to take the CAHSEE, not to give them an advantage over other students or to improve their scores. Accommodations change the way the test is administered but do not change what is tested. Test administration variations and accommodations do not result in changes to the students' scores for API or AYP calculations. Modifications fundamentally change what is being tested and the construct being measured. If a student takes one or both parts of the CAHSEE with a modification and has received the equivalent of a passing score, the student has not passed that part of the exam but is eligible to request a local waiver of the requirement to successfully pass that part of the exam. In addition, scores for students tested with modifications are counted as Not Pass for API calculations and Not Proficient for AYP calculations. The only exception is the calculator use in the mathematics exam. This modification is accepted and counted as tested. The Proficient cut point is a little higher than for those students who did not use a calculator.

Brief descriptions and abbreviations for the different modifications and accommodations are provided in Table 2.4. This table is also provided as a guide for district test site coordinators, who will mark all accommodations and modifications that are actually used during the testing in Box 25 of the students' answer documents. The shaded sections and the sections that cannot be marked for any portion of the CAHSEE are not applicable to that portion of the CAHSEE. For example, code Q (Calculators on the mathematics test) is considered a modification for mathematics and therefore cannot be marked for the ELA portion. The two-letter abbreviations included to the right of the description in Table 2.4 are used to reference the modifications and accommodations in Appendices 2.A to 2.G, Tables 3 to 6.

Modifications for mathematics include the use of a calculator, an arithmetic table, or a mathematics dictionary. Modifications for ELA include the use of a reader, an audio presentation, assisted devices, a scribe, a spell checker, a grammar checker, or an English dictionary.

Table 2.4: CAHSEE Modification/Accommodation Table: 2011–12

A. Student marks in test booklet (other than responses) -*Cannot be gridded for CAHSEE*		**N.** Dictionary. ELA Modification; Math Modification	DI (ELA) DM (M)
B. Student marks responses in test booklet and responses are transferred to a scorable answer document by an employee of the school, district, or nonpublic school. Accommodation	TS	**O.** Manually Coded English or American Sign Language to present test questions. ELA Modification; Math Accommodation	SL
C. Responses dictated [orally, or in Manually Coded English or American Sign Language] to a scribe for selected-response items (multiple-choice questions). Accommodation	OR	**P.** Test questions read aloud to student or audio presentation (CD). ELA Modification; Math Accommodation	OP
D. Word processing software with spell and grammar check tools turned off for the essay responses (writing portion of the test) - *Cannot be gridded on the math portion of box number 25, ELA side only.* ELA Accommodation	SO	**Q.** Calculators on mathematics test. -*Cannot be gridded for the ELA portion of box 25, math side only.* Math Modification	CA
E. Essay responses dictated orally or in Manually Coded English to a scribe, audio recorder, or speech-to-text converter and the student provides all spelling and language conventions -*Cannot be gridded on the math portion of box 25, ELA side only.* ELA Accommodation	EO	**R.** Arithmetic table on mathematics test - *Cannot be gridded for the ELA portion of Box 25, math side only.* Math Modification	AT
F. Assistive device that does not interfere with the independent work of the student on the .multiple-choice and/or essay responses (writing portion of the test) - *Cannot be gridded on the math portion of box number 25, ELA side only.* ELA Accommodation	AN	**S.** Math manipulatives on mathematics test -*Cannot be gridded for the ELA portion of Box 25, math side only.* Math Modification	MM
G. Braille transcriptions provided by the test contractor. Accommodation	BV	**T.** Word processing software with spell and grammar check tools enabled on the essay responses writing portion of the test - Cannot be gridded for the math portion of Box 25, *ELA side only.* ELA Modification	SC
H. Large print versions. Test items enlarged if font larger than required on large print version. Accommodation	LV	**U.** Essay responses dictated orally, in Manually Coded English, or in American Sign Language to a scribe, [audio recorder, or speech-to-text converter] and scribe provides spelling, grammar, and language conventions - *Cannot be gridded for the math portion of Box 25, ELA side only.* ELA Modification	ER
I. Extra time on test within a testing day -*Cannot be gridded for CAHSEE*		**V.** Assistive device that interferes with the independent work of the student on the multiple-choice and/or essay responses. Modification	AD
J. Test over more than one day for a test or test part to be administered in a single sitting. Accommodation	TD	**W.** Unlisted Modification. Modification	UM
K. Supervised breaks within a section of the test. Accommodation	SB	**X.** Unlisted Accommodation. Accommodation	UA
L. Administration of the test at the most beneficial time of day to the student. Accommodation	BT	**Y.** Writing task prompt ONLY was read aloud to the student or presented in Manually Coded English or American Sign Language – *Cannot be gridded for the math portion of box 25, ELA side only.* ELA Modification	WO

Table 2.4 (Continued)

M. Test administered at home or in hospital by test examiner. Accommodation	HH	**Z.** Leave Blank – *Cannot be gridded for CAHSEE*	

Note: The shaded sections are not applicable to the CAHSEE. The sections that cannot be marked for any portion of the CAHSEE are not applicable to that portion of the CAHSEE.

Scores for examinees who reported having a disability or who took the CAHSEE with an accommodation or modification were analyzed to investigate the relationships between CAHSEE scores and disability, accommodation, language fluency, and special program participation. Table 2.5 provides a listing of the tables summarizing student results based on disabilities and testing variations for the 2011–12 administrations. To simplify the presentation of these data, all tables for this section are located in Appendix 2.A to Appendix 2.G.

Tables 1 and 2 provide summary statistics in ELA and mathematics for each disability type and for all accommodated students, as well as separate statistics for accommodations versus modifications. Standard accommodations include students enrolled in an IEP/Section 504 plan who took either a Braille or audio CD form with no additional modifications. The largest disability group reported is Specific Learning Disability. The number of students in this group across the seven administrations ranged from 279 (December 2011) to 30,382 (March 2012) for ELA and 235 (December 2011) to 29,959 (March 2012) for mathematics. For students who tested with accommodations or modifications, the rates of achieving a score of 350 or higher ranged from 12 to 24 percent for ELA and from 13 to 26 percent for mathematics. Students who took the test with modifications are provided scale scores on their reports with Modified written beside the score. It is at the discretion of the local school boards whether each student using modifications and receiving a score of 350 or higher is granted a waiver.

Tables 3 and 4 present the percentage of students achieving less than 350 and 350 or higher, and summary statistics for each accommodation or modification used on the exam. Students enrolled in an IEP or Section 504 plan represent the largest number. The accommodation group with the largest percentage of students passing tended to vary across the seven administrations. Generally, students who tested at home or in the hospital, tested with the large print version, or had their responses transferred to an answer document achieved higher passing percentages. Tables 5 and 6 present scale scores at specific percentiles for each accommodation group.

Tables 7 and 8 present summary statistics for the breakdown of each testing variation by reported disability. Tables 9 and 10 summarize the comparison between language fluency categories within each testing variation group. The most commonly reported categories of accommodation/modification for students with limited English proficiency (i.e., initially fluent English proficient, reclassified fluent English proficient, EL) were directions read aloud or signed, additional breaks, and had access to a glossary or word list. English-only, reclassified fluent English proficient, and initially fluent English proficient students tended to outperform the EL students in most instances.

Table 2.5: Listing of Tables–Summary Statistics for Testing Variations and Disability

Table[1]	Content	Label
2.x.1	Scale Score Summary Statistics and Passing Rate Percentages for Testing Variations and Disability type—ELA	Summary Statistics by Testing Variations and Disability—ELA
2.x.2	Scale Score Summary Statistics and Passing Rate Percentages for Testing Variations and Disability type—Mathematics	Summary Statistics by Testing Variations and Disability—Mathematics
2.x.3	Demographic Summary and Passing Rate Percentages for All Examinees by Testing Variations—ELA	Demographic Summary for All Examinees by Testing Variations—ELA
2.x.4	Demographic Summary and Passing Rate Percentages for All Examinees by Testing Variations—Mathematics	Demographic Summary for All Examinees by Testing Variations—Mathematics
2.x.5	Scale Score Percentiles and Summary Statistics by Testing Variations—ELA	Percentiles of Scale Scores for Students with Testing Variations—ELA
2.x.6	Scale Score Percentiles and Summary Statistics by Testing Variations—Mathematics	Percentiles of Scale Scores for Students with Testing Variations—Mathematics
2.x.7	Scale Score Summary Statistics and Passing Rate Percentages by Disability and Testing Variations—ELA	Summary of Scale Scores and Passing Rates by Disability and Testing Variations —ELA
2.x.8	Scale Score Summary Statistics and Passing Rate Percentages by Disability and Testing Variations—Mathematics	Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics
2.x.9	Scale Score Summary Statistics and Passing Rates by Language Fluency and Testing Variations—ELA	Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA
2.x.10	Scale Score Summary Statistics and Passing Rates by Language Fluency and Testing Variations—Mathematics	Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics

[1] x = Administration, where tables A = July, B = October, C = November, D = December, E = February, F = March, G = May.

Scores

Student raw scores are transformed to three-digit scale scores using the equating process described in Chapter 6. The weighting of the total raw scores is described below. CAHSEE results are reported using scale scores that range from 275 to 450 for both tests. The passing status also is reported. If the scale score is 350 or higher, it will be marked as Passed. If the scale score is less than 350, it will be marked as Not Passed.

In addition to scale scores, student performance on various strands is reported. The strand score is obtained by adding an examinee's scores on the items in each strand. That information is reported in terms of a percent correct score. Detailed descriptions of CAHSEE scores are described in Chapter 8.

Weighting of Scores

The ELA section consists of 45 MC items measuring reading, 27 MC items measuring writing, and one CR prompt. Each correct MC item is worth one point, and the points are added to calculate the total MC score. The CR item is scored on a rubric ranging from 0 to 4 points. Each CR item is initially scored by two raters, and if the scores are the same or adjacent, the resulting item score is the average of the two ratings. Half-

point intervals are possible when the scores assigned by the raters are adjacent scores. When the raters assign non-adjacent scores, the scoring leader, who assigns the final score, provides resolution. The weighting of these components is described below. The total raw score points is 90.

MC Reading Items:	45 items times scoring weight of 1.0 = 45 points
MC Writing Items:	27 items times scoring weight of 1.0 = 27 points
CR Item:	4 point rubric times scoring weight of 4.5 = 18 points

Total points MC = 72 points (80 percent of the total points)
Total points CR = 18 points (20 percent of the total points)
Total points: Composite (MC + Weighted CR) = 90 points

There is no special weighting for the mathematics test. Each item is worth one score point for a total of 80 points.

Total raw scores on the CAHSEE are transformed to a reporting scale that ranges from 275 to 450, with the minimum passing score set at 350.

Aggregation Procedures

In order to provide meaningful results to the stakeholders, CAHSEE scores are aggregated at the school, independent charter school, district, county, and state levels. The aggregated scores are generated both for individual scores as well as group scores. The following section presents the types of aggregation performed on CAHSEE scores.

Individual Scores

Summary statistics of individual student scores expressed as raw scores and scale scores are provided in the Executive Summary. Tables E.3.1 to E.3.3 contain the information for the ELA tests, and Tables E.4.1 to E.4.3 contain the information for the mathematics tests. The tables include the means, standard deviations, range, and median. The percentages of students passing the CAHSEE are found in Table E.2 in the Executive Summary.

Group Scores

Results for the demographic groups may be found in Appendices 8.A to 8.G. Summary information is presented by demographic characteristic, including grade, gender, ethnicity, language fluency, economic status, and special education services for ELA and mathematics. Tables 5 and 6 provide summary statistics based on all students taking each of the administrations for the ELA and mathematics tests, respectively. The tables show the number of students with valid scores in each group, the number and percent of students in the Pass/Not Pass classifications, mean scale scores, mean percent correct for the strands, and the mean score for writing applications.

Additional subgroup information may be found in Tables 7 though 10 of Appendix 8. Tables 7 and 8 display the number and percent of students classified as Below

Proficient, Proficient, and Advanced according to the ESEA performance classifications. Selected percentiles, scale score means, and standard deviations for the subgroups are presented for all students in Tables 9 and 10 for ELA and mathematics, respectively.

Equating

Students taking the CAHSEE have multiple opportunities to take the examination until they pass both the ELA and mathematics portions. When administering multiple forms of a test, there is a need for a constant scale. This means that the passing score must represent the same level of achievement on all forms (versions) of the CAHSEE. To maintain comparability of scores across multiple test forms, the CAHSEE tests are equated to a reference form using a common-item nonequivalent groups design and methods based on IRT.

The procedure used for equating the CAHSEE involves three steps: item calibration, item parameter scaling, and true-score equating. The 2004 February administration is the baseline for equating all CAHSEE test forms. The 2011–12 items were calibrated and placed on the reference scale using a set of linking items selected from the previous year's forms and re-administered on the 2011–12 test forms. The number correct or raw score is converted to a scale score via true-score equating. The raw-score to scale-score conversion reflects the relationship between the difficulty of individual test items that make up each test form and the constant measure of achievement indicated by the reported scale scores. For different test forms, the expected number-correct score for a given level of achievement may vary somewhat due to (usually small) differences in the average difficulty of the items in one form compared to the average difficulty of items in other test forms. This is why the conversion tables for each test administration will differ slightly in relating raw scores to scale scores. Total scores on the CAHSEE are transformed to a reported scale that ranges from 275 to 450, with the minimum passing score set at 350. The equating specifications and procedures are described in detail in Chapter 6.

Appendix 2.A: Results of Testing Variations and Disability Analyses—July 2011

Table 2.A.1: Summary Statistics by Testing Variations and Disability—ELA, July 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	110	315	23	10
Modifications	331	329	30	21
All	441	325	29	18
Disability				
Autism	26	313	21	15
Deaf	26	315	28	8
Deaf-Blindness	-	-	-	-
Emotional Disturbance	80	325	31	23
Hard of Hearing	12	317	19	8
Mental Retardation	25	302	15	0
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	53	333	27	21
Specific Learning Disability	857	325	25	15
Speech or Language Impairment	58	328	20	7
Traumatic Brain Injury	-	-	-	-

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.2: Summary Statistics by Testing Variations and Disability—Mathematics, July 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	139	331	27	17
Modifications	433	328	19	11
All	572	329	21	13
Disability				
Autism	25	335	19	24
Deaf	14	333	37	21
Deaf-Blindness	-	-	-	-
Emotional Disturbance	78	326	22	15
Hard of Hearing	12	332	17	17
Mental Retardation	24	314	12	0
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	65	332	25	20
Specific Learning Disability	871	327	20	11
Speech or Language Impairment	51	332	19	18
Traumatic Brain Injury	-	-	-	-

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.3: Demographic Summary for All Examinees by Testing Variations—ELA, July 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	840	133	16	707	84	324	48	47	51	37	40	1.8
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-	-	-
Spell Checker/Grammar Checker Off (SO)	-	-	-	-	-	-	-	-	-	-	-	-
Essay Responses (EO)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	33	2	6	31	94	321	50	45	53	35	40	1.6
Supervised Breaks (SB)	145	18	12	127	88	319	42	44	48	36	38	1.8
Beneficial Time (BT)	17	4	24	13	76	316	45	46	46	36	34	1.4
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	115	20	17	95	83	329	55	49	54	40	41	1.9
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	252	58	23	194	77	330	53	49	55	41	44	1.9
Spell Checker or Grammar Checker (SC)	26	5	19	21	81	329	49	50	56	34	42	2.0
Essay Responses (ER)	26	5	19	21	81	325	52	46	51	32	39	2.0
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	23	5	22	18	78	321	41	43	50	34	45	1.7
Writing Only (WO)	21	16	76	5	24	379	82	75	79	75	80	2.1

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.A.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, July 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 plan	873	102	12	771	88	328	41	48	38	34	32
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	17	1	6	16	94	323	43	40	36	32	28
Supervised Breaks (SB)	139	10	7	129	93	325	39	45	37	32	31
Beneficial Time (BT)	21	5	24	16	76	326	37	48	40	30	36
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	17	4	24	13	76	334	45	54	44	36	34
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	174	28	16	146	84	331	43	51	39	37	35
Calculator (CA)	432	48	11	384	89	328	41	52	37	34	30
Arithmetic Table (AT)	41	2	5	39	95	324	38	49	33	33	29
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	26	1	4	25	96	321	40	35	37	30	29

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.A.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, July 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	286	304	322	340	372	407	324	27	840
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-
Spell Checker/Grammar Checker Off (SO)	-	-	-	-	-	-	-	-	-	-
Essay Responses (EO)	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	278	278	308	324	340	350	352	321	22	33
Supervised Breaks (SB)	275	283	301	320	334	360	384	319	25	145
Beneficial Time (BT)	275	275	289	299	342	369	369	316	33	17
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	280	287	308	330	342	372	392	329	27	115
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	280	289	308	324	346	395	411	330	31	252
Spell Checker or Grammar Checker (SC)	278	291	312	323	340	379	411	329	29	26
Essay Responses (ER)	280	295	306	319	346	367	367	325	24	26
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	276	283	302	310	336	376	392	321	31	23
Writing Only (WO)	302	306	372	395	407	411	419	379	38	21

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.A.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, July 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	284	299	313	327	339	364	398	328	21	873
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	304	304	311	317	339	350	350	323	16	17
Supervised Breaks (SB)	275	299	311	325	337	357	372	325	19	139
Beneficial Time (BT)	275	275	304	321	348	372	398	326	32	21
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	306	306	321	334	344	364	364	334	17	17
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	288	299	313	327	344	395	408	331	27	174
Calculator (CA)	294	301	315	327	341	360	390	328	19	432
Arithmetic Table (AT)	291	296	311	325	339	348	378	324	19	41
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	288	291	308	318	337	344	352	321	17	26

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.A.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, July 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	Emotional Disturbance	11	308	30	18
Dictionary	Specific Learning Disability	84	327	23	14
Essay Reponses (ER)	Specific Learning Disability	21	326	24	19
IEP or Section 504 Plan	Autism	14	321	21	21
	Emotional Disturbance	63	325	33	24
	Mental Retardation	18	302	15	0
	Other Health Impairment	33	335	31	24
	Specific Learning Disability	602	324	26	14
	Speech or Language Impairment	29	327	24	10
Oral Presentation	Other Health Impairment	12	331	39	25
	Specific Learning Disability	191	330	31	23
Spell Checker or Grammar Checker	Specific Learning Disability	22	327	25	14
Supervised Breaks	Emotional Disturbance	19	307	26	11
	Specific Learning Disability	107	321	23	11
Test Over More Than One Day	Specific Learning Disability	21	321	22	10
Unlisted Accommodation	Specific Learning Disability	19	321	31	21
Writing Only	Specific Learning Disability	13	394	26	92

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, July 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	31	324	19	3
Beneficial Time	Emotional Disturbance	11	318	20	9
Calculator	Autism	11	338	14	9
	Emotional Disturbance	20	323	20	10
	Other Health Impairment	25	334	27	24
	Specific Learning Disability	307	327	18	9
	Speech or Language Impairment	19	337	16	26
Dictionary for Math	Specific Learning Disability	13	334	15	23
IEP or Section 504 Plan	Autism	14	333	17	7
	Emotional Disturbance	61	325	23	16
	Mental Retardation	15	312	12	0
	Other Health Impairment	44	333	28	20
	Specific Learning Disability	626	327	20	10
	Speech or Language Impairment	31	332	20	23
Oral Presentation	Specific Learning Disability	132	331	27	14
	Speech or Language Impairment	11	333	20	18
Supervised Breaks	Emotional Disturbance	19	318	19	11
	Specific Learning Disability	97	326	16	6
Test Over More Than One Day	Specific Learning Disability	12	326	16	8
Unlisted Accommodation	Specific Learning Disability	22	322	18	5

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, July 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	English Only	14	308	29	14
Dictionary	English Only	57	330	26	21
	English Learner	52	328	22	12
Essay Reponses (ER)	English Only	12	324	26	25
	English Learner	13	325	24	15
IEP or Section 504 Plan	English Only	417	325	27	18
	Initially Fluent English Proficient	21	343	33	38
	Reclassified Fluency	35	327	27	20
	English Learner	354	322	25	12
	Unknown	13	320	29	8
Oral Presentation	English Only	117	331	29	23
	English Learner	113	329	33	22
Spell Checker or Grammar Checker	English Only	11	334	35	27
	English Learner	14	322	21	7
Supervised Breaks	English Only	58	318	28	14
	English Learner	78	319	21	9
Test Over More Than One Day	English Only	14	315	25	7
	English Learner	17	326	20	6
Writing Only	English Learner	12	389	31	83

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, July 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	24	330	18	8
	English Learner	16	316	17	0
Beneficial Time	English Only	16	321	32	19
Calculator	English Only	237	329	19	11
	Reclassified Fluency	18	324	15	11
	English Learner	161	327	19	11
Dictionary for Math	English Only	13	333	16	23
IEP or Section 504 Plan	English Only	469	328	21	12
	Initially Fluent English Proficient	20	332	18	15
	Reclassified Fluency	30	329	15	13
	English Learner	338	326	21	10
	Unknown	16	337	16	25
Oral Presentation	English Only	83	331	26	13
	English Learner	84	331	28	18
Supervised Breaks	English Only	63	322	20	6
	English Learner	69	326	16	6
Unlisted Accommodation	English Only	14	319	19	7

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.B: Results of Testing Variations and Disability Analyses—October 2011

Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,373	325	28	20
Modifications	2,447	329	25	22
All	3,820	327	26	21
Disability				
Unknown	-	-	-	-
Autism	250	324	28	18
Deaf	116	307	23	8
Deaf-Blindness	-	-	-	-
Emotional Disturbance	687	325	33	23
Hard of Hearing	106	320	22	11
Mental Retardation	267	309	24	6
Multiple Disability	24	308	22	4
Orthopedic Impairment	78	325	24	14
Other Health Impairment	628	331	26	26
Specific Learning Disability	7,268	326	24	18
Speech or Language Impairment	425	327	24	19
Traumatic Brain Injury	42	315	25	14
Visual Impairment	20	329	31	25

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.B.2: Summary Statistics by Testing Variations and Disability—Mathematics, October 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,133	328	22	17
Modifications	3,703	332	21	19
All	4,836	331	21	19
Disability				
Unknown	-	-	-	-
Autism	227	332	22	19
Deaf	91	328	18	16
Deaf-Blindness	-	-	-	-
Emotional Disturbance	786	328	26	19
Hard of Hearing	79	330	20	19
Mental Retardation	246	318	19	9
Multiple Disability	24	328	22	17
Orthopedic Impairment	81	332	20	21
Other Health Impairment	717	333	21	22
Specific Learning Disability	6,999	330	20	17
Speech or Language Impairment	368	333	20	21
Traumatic Brain Injury	40	329	21	23
Visual Impairment	23	332	19	22

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.B.3: Demographic Summary for All Examinees by Testing Variations—ELA, October 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	7,871	1,479	19	6,392	81	326	48	49	50	39	46	1.9
Transfer of T/B Responses to A/D (TS)	71	31	44	40	56	343	60	58	59	50	55	2.1
Oral Responses Dictated to a Scribe (OR)	16	2	13	14	88	320	48	41	48	40	41	1.6
Spell Checker/Grammar Checker Off (SO)	49	17	35	32	65	341	59	59	57	46	53	2.2
Essay Responses (EO)	41	10	24	31	76	327	52	47	51	39	43	2.0
Assistive Device No Interference (AN)	43	11	26	32	74	331	50	52	51	41	54	1.9
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	21	8	38	13	62	329	53	47	52	44	50	1.9
Test Over More Than One Day (TD)	415	105	25	310	75	329	51	51	52	41	46	2.0
Supervised Breaks (SB)	1,616	326	20	1,290	80	325	49	49	49	39	46	1.9
Beneficial Time (BT)	292	61	21	231	79	325	49	49	49	40	44	1.8
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	817	199	24	618	76	330	58	52	53	41	48	2.0
Sign Language (SL)	31	2	6	29	94	305	34	33	35	31	38	1.6
Oral Presentation (OP)	1,934	461	24	1,473	76	331	52	53	53	42	49	2.0
Spell Checker or Grammar Checker (SC)	173	42	24	131	76	331	52	51	53	42	50	2.1
Essay Responses (ER)	88	18	20	70	80	319	41	44	42	38	40	1.9
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	39	2	5	37	95	311	44	42	39	32	35	1.6
Unlisted Accommodation (UA)	290	70	24	220	76	330	51	52	53	44	49	1.9
Writing Only (WO)	109	18	17	91	83	320	47	42	43	37	45	1.9

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.B.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, October 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 plan	7,867	1,370	17	6,497	83	330	42	45	40	37	32
Transfer of T/B Responses to A/D (TS)	75	31	41	44	59	342	55	51	48	42	39
Oral Responses Dictated to a Scribe (OR)	14	2	14	12	86	333	49	49	43	37	27
Braille Version (BV)	13	3	23	10	77	336	50	56	41	42	26
Large Print Version (LV)	19	6	32	13	68	331	44	49	39	32	36
Test Over More Than One Day (TD)	294	74	25	220	75	332	43	47	41	38	32
Supervised Breaks (SB)	1,374	229	17	1,145	83	329	41	43	39	36	32
Beneficial Time (BT)	261	43	16	218	84	325	40	42	36	33	30
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	204	45	22	159	78	336	47	48	44	41	32
Sign Language (SL)	32	8	25	24	75	336	43	47	41	43	43
Oral Presentation (OP)	1,024	236	23	788	77	335	45	47	43	41	34
Calculator (CA)	3,627	687	19	2,940	81	332	43	47	41	39	33
Arithmetic Table (AT)	377	96	25	281	75	336	46	48	43	41	36
Math Manipulative (MM)	50	17	34	33	66	336	47	49	45	40	35
Assistive Device (AD)	14	8	57	6	43	347	53	64	53	44	34
Unlisted Modification (UM)	46	8	17	38	83	329	45	44	36	38	28
Unlisted Accommodation (UA)	287	53	18	234	82	331	43	43	40	38	34

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.B.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, October 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	285	307	326	344	369	389	326	26	7,871
Transfer of T/B Responses to A/D (TS)	275	285	320	342	356	401	450	343	37	71
Oral Responses Dictated to a Scribe (OR)	275	275	293	323	341	401	401	320	33	16
Spell Checker/Grammar Checker Off (SO)	275	293	322	338	356	402	450	341	33	49
Essay Reponses (EO)	275	287	307	319	346	389	392	327	30	41
Assistive Device No Interference (AN)	275	279	301	334	352	376	413	331	32	43
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	277	285	305	334	350	364	401	329	30	21
Test Over More Than One Day (TD)	275	279	309	328	350	373	395	329	28	415
Supervised Breaks (SB)	275	283	305	324	344	369	389	325	27	1,616
Beneficial Time (BT)	275	281	301	324	344	371	395	325	28	292
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	277	291	311	330	347	373	386	330	25	817
Sign Language (SL)	275	275	293	303	313	350	384	305	23	31
Oral Presentation (OP)	277	291	313	330	347	373	389	331	25	1,934
Spell Checker or Grammar Checker (SC)	285	295	315	330	347	373	395	331	24	173
Essay Reponses (ER)	275	275	293	320	342	371	386	319	30	88
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	275	275	291	315	332	350	352	311	23	39
Unlisted Accommodation (UA)	277	287	311	334	347	373	402	330	26	290
Writing Only (WO)	275	277	301	315	334	366	392	320	27	109

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.B.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, October 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	286	301	317	328	343	366	394	330	21	7,867
Transfer of T/B Responses to A/D (TS)	275	306	321	339	358	389	438	342	27	75
Oral Responses Dictated to a Scribe (OR)	275	275	314	332	343	403	403	333	31	14
Braille Version (BV)	304	304	313	327	344	403	403	336	31	13
Large Print Version (LV)	295	295	312	324	350	378	378	331	25	19
Test Over More Than One Day (TD)	275	295	317	328	350	370	422	332	26	294
Supervised Breaks (SB)	279	301	314	326	341	368	394	329	22	1,374
Beneficial Time (BT)	275	292	308	323	341	362	382	325	23	261
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	292	306	321	336	346	372	384	336	21	204
Sign Language (SL)	306	306	323	339	349	370	370	336	18	32
Oral Presentation (OP)	289	303	319	332	348	378	409	335	24	1,024
Calculator (CA)	292	303	319	330	344	368	397	332	21	3,627
Arithmetic Table (AT)	298	306	321	334	350	382	406	336	22	377
Math Manipulatives (MM)	298	303	317	335	353	370	378	336	22	50
Assistive Device (AD)	314	314	330	353	362	370	370	347	19	14
Unlisted Modification (UM)	275	298	310	328	343	358	392	329	22	46
Unlisted Accommodation (UA)	289	298	314	328	343	366	422	331	22	287

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.B.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, October 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	Specific Learning Disability	30	333	27	23
Beneficial Time	Emotional Disturbance	63	321	32	21
	Other Health Impairment	18	325	24	22
	Specific Learning Disability	165	326	26	20
Dictionary	Autism	14	328	28	14
	Emotional Disturbance	29	325	30	28
	Mental Retardation	28	324	27	21
	Other Health Impairment	32	337	20	25
	Specific Learning Disability	542	330	25	24
	Speech or Language Impairment	19	327	25	16
Essay Reponses (EO)	Specific Learning Disability	20	330	25	25
Essay Reponses (ER)	Deaf	36	300	21	6
	Specific Learning Disability	35	340	24	34
IEP or Section 504 Plan	Autism	180	325	27	17
	Deaf	99	306	21	6
	Emotional Disturbance	518	326	33	25
	Hard of Hearing	81	320	22	9
	Mental Retardation	204	309	23	6
	Multiple Disability	21	309	22	5
	Orthopedic Impairment	61	325	24	16
	Other Health Impairment	445	331	26	24
	Specific Learning Disability	5,180	326	24	18
	Speech or Language Impairment	264	325	24	17
	Traumatic Brain Injury	29	308	24	7
	Visual Impairment	16	328	33	31
Oral Presentation	Autism	47	324	23	17
	Emotional Disturbance	67	331	32	27
	Hard of Hearing	14	327	19	14
	Mental Retardation	78	318	27	13
	Orthopedic Impairment	17	327	28	29
	Other Health Impairment	92	334	23	28
	Specific Learning Disability	1,273	331	24	24
	Speech or Language Impairment	79	331	23	18
Sign Language	Deaf	23	302	16	4
Spell Checker or Grammar Checker	Deaf	14	302	11	0
	Other Health Impairment	13	340	20	23
	Specific Learning Disability	111	333	23	26
Spell Checker or Grammar Checker Off	Specific Learning Disability	18	342	21	39
Supervised Breaks	Autism	29	328	36	24
	Emotional Disturbance	151	323	31	21
	Hard of Hearing	15	310	23	7
	Mental Retardation	65	311	22	5
	Other Health Impairment	85	330	27	28
	Specific Learning Disability	1,025	326	25	19
	Speech or Language Impairment	45	329	25	22

Table 2.B.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Test Over More Than One Day	Autism	14	336	30	36
	Deaf	17	303	17	6
	Emotional Disturbance	59	328	34	25
	Other Health Impairment	19	328	26	21
	Specific Learning Disability	240	329	27	26
	Speech or Language Impairment	21	339	26	38
Transfer of Student T/B Responses to A/D	Specific Learning Disability	21	346	35	38
Unlisted Accommodation	Emotional Disturbance	13	327	29	23
	Other Health Impairment	21	326	24	19
	Specific Learning Disability	183	332	26	25
Unlisted Modification	Specific Learning Disability	19	318	23	11
Writing Only	Deaf	15	308	24	13
	Specific Learning Disability	70	318	24	13

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.B.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, October 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Emotional Disturbance	17	336	32	24
	Other Health Impairment	13	343	30	38
	Specific Learning Disability	204	334	20	22
Assistive Device	Specific Learning Disability	11	344	20	45
Beneficial Time	Emotional Disturbance	58	317	24	9
	Other Health Impairment	17	323	20	0
	Specific Learning Disability	153	328	21	20
Calculator	Autism	83	332	21	19
	Emotional Disturbance	208	331	26	22
	Hard of Hearing	19	335	19	26
	Mental Retardation	104	320	18	9
	Multiple Disability	11	334	20	18
	Orthopedic Impairment	28	331	17	21
	Other Health Impairment	220	333	20	20
	Specific Learning Disability	2,380	333	20	18
	Speech or Language Impairment	119	334	21	22
	Traumatic Brain Injury	14	330	24	36
Dictionary for Math	Emotional Disturbance	11	332	33	27
	Specific Learning Disability	124	336	20	23
IEP or Section 504 Plan	Autism	159	331	21	19
	Deaf	78	327	17	13
	Emotional Disturbance	605	329	27	20
	Hard of Hearing	59	330	20	19
	Mental Retardation	180	318	19	9
	Multiple Disability	20	326	20	15
	Orthopedic Impairment	59	331	20	19
	Other Health Impairment	507	332	20	20
	Specific Learning Disability	5,093	330	20	16
	Speech or Language Impairment	232	330	20	17
	Traumatic Brain Injury	28	327	21	21
	Visual Impairment	18	331	20	22
Math Manipulatives	Specific Learning Disability	32	339	21	38
Oral Presentation	Autism	22	326	18	9
	Emotional Disturbance	47	337	32	34
	Mental Retardation	33	322	24	12
	Other Health Impairment	48	334	26	25
	Specific Learning Disability	694	335	23	22
	Speech or Language Impairment	43	336	24	26
Sign Language	Deaf	21	335	17	24
Supervised Breaks	Autism	20	330	27	25
	Emotional Disturbance	142	323	23	13
	Mental Retardation	53	318	20	9

Table 2.B.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Other Health Impairment	84	333	21	20
	Specific Learning Disability	867	329	21	16
	Speech or Language Impairment	33	331	23	9
Test Over More Than One Day	Emotional Disturbance	55	324	24	16
	Other Health Impairment	12	346	29	42
	Specific Learning Disability	172	332	26	24
Transfer of Student T/B Responses to A/D	Specific Learning Disability	19	343	31	37
Unlisted Accommodation	Emotional Disturbance	15	341	31	33
	Other Health Impairment	17	320	22	18
	Specific Learning Disability	189	331	20	17
Unlisted Modification	Specific Learning Disability	26	330	19	15

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.B.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, October 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	27	334	35	33
Beneficial Time	English Only	173	325	31	24
	English Learner	103	325	25	17
Dictionary	English Only	429	331	26	27
	Initially Fluent English Proficient	24	328	26	29
	Reclassified Fluent English Proficient	31	333	27	23
	English Learner	325	329	24	21
Essay Reponses (EO)	English Only	32	327	31	25
Essay Reponses (ER)	English Only	74	316	29	18
	English Learner	12	337	29	33
IEP or Section 504 Plan	English Only	4,499	327	27	22
	Initially Fluent English Proficient	180	327	23	19
	Reclassified Fluent English Proficient	224	333	25	26
	English Learner	2,902	322	23	13
	Unknown	66	324	28	18
Large Print Version	English Only	13	327	28	46
Oral Presentation	English Only	1,066	332	26	27
	Initially Fluent English Proficient	40	335	26	35
	Reclassified Fluent English Proficient	82	334	25	30
	English Learner	727	327	23	18
	Unknown	19	326	20	16
Sign Language	English Only	14	308	30	7
	English Learner	15	304	17	7
Spell Checker or Grammar Checker	English Only	94	336	25	33
	Reclassified Fluent English Proficient	12	333	15	17
	English Learner	64	323	22	11
Spell Checker or Grammar Checker Off	English Only	36	343	36	33
	English Learner	12	334	24	33
Supervised Breaks	English Only	954	326	28	23
	Initially Fluent English Proficient	39	326	25	18
	Reclassified Fluent English Proficient	49	335	26	33
	English Learner	562	323	24	14
	Unknown	12	321	32	25
Test Over More Than One Day	English Only	266	328	29	24
	Initially Fluent English Proficient	13	331	30	38
	Reclassified Fluent English Proficient	12	335	19	33
	English Learner	122	329	26	25
Transfer of Student T/B Responses to A/D	English Only	53	346	39	51
	English Learner	14	336	27	21
Unlisted Accommodation	English Only	190	331	29	29
	Reclassified Fluent English Proficient	11	338	12	18
	English Learner	76	327	21	12
Unlisted Modification	English Only	20	303	24	0
	English Learner	15	320	21	7

Table 2.B.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Writing Only	English Only	62	324	30	24
	English Learner	38	311	19	3

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.B.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, October 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	231	337	22	25
	Initially Fluent English Proficient	11	330	23	27
	Reclassified Fluent English Proficient	11	345	20	27
	English Learner	118	336	21	25
Assistive Device	English Only	12	346	20	58
Beneficial Time	English Only	168	325	24	17
	Reclassified Fluent English Proficient	11	324	27	18
	English Learner	74	326	21	16
Calculator	English Only	2,192	333	21	20
	Initially Fluent English Proficient	95	333	21	25
	Reclassified Fluent English Proficient	130	336	20	25
	English Learner	1,168	331	20	16
	Unknown	42	333	16	19
Dictionary for Math	English Only	127	333	20	18
	English Learner	61	339	20	28
IEP or Section 504 Plan	English Only	4,779	331	22	18
	Initially Fluent English Proficient	205	330	22	19
	Reclassified Fluent English Proficient	242	337	20	27
	English Learner	2,572	329	20	16
	Unknown	69	329	18	13
Large Print Version	English Only	13	326	24	23
Math Manipulatives	English Only	32	333	20	25
	English Learner	12	346	18	50
Oral Presentation	English Only	613	336	25	25
	Initially Fluent English Proficient	21	330	24	14
	Reclassified Fluent English Proficient	37	339	24	38
	English Learner	341	333	21	19
	Unknown	12	328	15	17
Oral Responses Dictated to a Scribe	English Only	11	325	20	0
Sign Language	English Only	13	332	16	8
	English Learner	14	337	19	43
Supervised Breaks	English Only	872	329	22	16
	Initially Fluent English Proficient	35	323	16	6
	Reclassified Fluent English Proficient	50	337	25	32
	English Learner	406	330	21	17
	Unknown	11	326	16	9
Test Over More Than One Day	English Only	212	331	26	23
	Reclassified Fluent English Proficient	12	347	32	50
	English Learner	62	332	26	27
Transfer of Student T/B Responses to A/D	English Only	63	344	28	44
Unlisted Accommodation	English Only	198	332	24	20
	Reclassified Fluent English Proficient	11	336	20	18
	English Learner	68	327	19	15

Table 2.B.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Unlisted Modification	English Only	19	321	20	11
	English Learner	20	336	21	20

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.C: Results of Testing Variations and Disability Analyses—November 2011

Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	4,282	325	27	18
Modifications	6,118	327	25	17
All	10,400	326	26	18
Disability				
Autism	841	328	28	20
Deaf	275	311	21	6
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,678	327	35	25
Hard of Hearing	270	326	26	17
Mental Retardation	570	304	20	3
Multiple Disability	43	317	25	9
Orthopedic Impairment	222	327	26	20
Other Health Impairment	2,150	331	28	24
Specific Learning Disability	19,476	326	24	17
Speech or Language Impairment	1,268	330	25	20
Traumatic Brain Injury	117	324	26	16
Visual Impairment	70	333	28	29

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.C.2: Summary Statistics by Testing Variations and Disability—Mathematics, November 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,737	331	22	19
Modifications	10,060	333	20	21
All	13,797	333	20	20
Disability				
Autism	857	335	21	23
Deaf	223	330	20	18
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,975	331	27	23
Hard of Hearing	235	336	20	22
Mental Retardation	584	316	16	3
Multiple Disability	43	323	23	14
Orthopedic Impairment	246	332	21	18
Other Health Impairment	2,489	335	22	26
Specific Learning Disability	19,805	331	19	18
Speech or Language Impairment	1,156	335	20	24
Traumatic Brain Injury	102	333	22	22
Visual Impairment	82	337	26	32

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.C.3: Demographic Summary for All Examinees by Testing Variations—ELA, November 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	21,358	3,722	17	17,636	83	326	58	46	53	39	45	1.8
Transfer of T/B Responses to A/D (TS)	120	29	24	91	76	330	61	52	54	41	45	1.8
Oral Responses Dictated to a Scribe (OR)	45	10	22	35	78	324	59	43	49	39	45	1.6
Spell Checker/Grammar Checker Off (SO)	163	41	25	122	75	332	59	50	56	43	48	2.0
Essay Responses (EO)	70	13	19	57	81	329	62	50	52	43	45	1.9
Assistive Device No Interference (AN)	25	6	24	19	76	338	65	56	61	43	53	2.0
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	52	13	25	39	75	330	66	48	54	42	50	1.7
Test Over More Than One Day (TD)	1,307	268	21	1,039	79	326	57	47	53	40	45	1.8
Supervised Breaks (SB)	4,353	772	18	3,581	82	325	57	46	52	39	44	1.8
Beneficial Time (BT)	519	91	18	428	82	322	55	45	50	36	41	1.6
Tested at Home or Hospital (HH)	45	19	42	26	58	341	66	55	61	48	54	1.9
Dictionary (DI)	1,664	295	18	1,369	82	328	61	47	54	41	45	1.9
Sign Language (SL)	88	4	5	84	95	314	47	35	48	34	38	1.7
Oral Presentation (OP)	5,137	947	18	4,190	82	328	60	47	55	41	45	1.8
Spell Checker or Grammar Checker (SC)	450	93	21	357	79	330	60	48	55	41	46	2.0
Essay Responses (ER)	93	14	15	79	85	325	53	42	51	37	43	2.2
Assistive Device (AD)	31	11	35	20	65	341	64	54	60	42	50	2.5
Unlisted Modification (UM)	19	2	11	17	89	327	65	45	53	39	49	1.8
Unlisted Accommodation (UA)	609	98	16	511	84	326	57	46	52	39	45	1.8
Writing Only (WO)	371	49	13	322	87	320	48	42	50	37	43	1.7

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.C.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, November 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 plan	22,510	4,370	19	18,140	81	332	46	48	40	39	33
Transfer of T/B Responses to A/D (TS)	106	30	28	76	72	338	51	50	46	43	36
Oral Responses Dictated to a Scribe (OR)	46	12	26	34	74	334	45	51	43	38	36
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	54	16	30	38	70	339	53	52	45	44	36
Test Over More Than One Day (TD)	833	200	24	633	76	333	48	49	42	39	33
Supervised Breaks (SB)	4,008	759	19	3,249	81	330	46	47	39	38	32
Beneficial Time (BT)	477	96	20	381	80	328	44	46	38	37	31
Tested at Home or Hospital (HH)	39	8	21	31	79	335	51	46	43	42	34
Dictionary for Math (DM)	256	48	19	208	81	333	48	49	40	39	33
Sign Language (SL)	111	21	19	90	81	332	43	48	40	40	37
Oral Presentation (OP)	3,131	672	21	2,459	79	333	47	50	41	40	34
Calculator (CA)	9,920	2,042	21	7,878	79	333	47	51	41	39	33
Arithmetic Table (AT)	882	207	23	675	77	335	49	51	42	41	34
Math Manipulative (MM)	136	58	43	78	57	343	55	56	51	44	39
Assistive Device (AD)	25	13	52	12	48	346	60	58	49	47	41
Unlisted Modification (UM)	16	2	13	14	88	335	54	55	38	36	37
Unlisted Accommodation (UA)	635	130	20	505	80	333	47	48	41	39	34

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.C.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, November 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	286	306	324	343	369	393	326	26	21,358
Transfer of T/B Responses to A/D (TS)	275	284	310	331	349	371	384	330	27	120
Oral Responses Dictated to a Scribe (OR)	275	275	300	315	343	374	449	324	37	45
Spell Checker/Grammar Checker Off (SO)	282	294	311	334	351	369	393	332	25	163
Essay Responses (EO)	275	296	311	324	345	382	404	329	25	70
Assistive Device No Interference (AN)	306	311	332	339	349	367	374	338	18	25
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	275	278	313	330	349	371	376	330	26	52
Test Over More Than One Day (TD)	275	282	306	326	345	371	393	326	27	1,307
Supervised Breaks (SB)	275	284	306	324	343	369	393	325	26	4,353
Beneficial Time (BT)	275	275	298	320	343	374	396	322	30	519
Tested at Home or Hospital (HH)	275	282	313	341	369	393	450	341	37	45
Dictionary (DI)	275	290	309	328	345	369	387	328	24	1,664
Sign Language (SL)	282	286	298	313	326	347	382	314	20	88
Oral Presentation (OP)	275	290	309	328	345	369	390	328	25	5,137
Spell Checker or Grammar Checker (SC)	282	292	313	330	347	369	390	330	24	450
Essay Responses (ER)	275	292	307	320	339	367	390	325	23	93
Assistive Device (AD)	296	298	317	341	364	390	390	341	27	31
Unlisted Modification (UM)	284	284	313	330	339	369	369	327	22	19
Unlisted Accommodation (UA)	275	286	307	324	341	369	393	326	26	609
Writing Only (WO)	275	286	300	315	339	360	384	320	25	371

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.C.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, November 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	288	302	317	330	344	368	387	332	21	22,510
Transfer of T/B Responses to A/D (TS)	296	304	319	336	353	385	398	338	25	106
Oral Responses Dictated to a Scribe (OR)	296	304	313	330	350	370	387	334	23	46
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	302	306	325	334	353	385	395	339	23	54
Test Over More Than One Day (TD)	277	296	317	330	348	376	398	333	24	833
Supervised Breaks (SB)	285	299	315	328	344	366	385	330	21	4,008
Beneficial Time (BT)	275	291	313	325	343	374	390	328	25	477
Tested at Home or Hospital (HH)	281	296	317	330	344	398	401	335	26	39
Dictionary for Math (DM)	296	309	319	330	343	368	390	333	19	256
Sign Language (SL)	294	304	323	334	344	359	370	332	17	111
Oral Presentation (OP)	294	304	319	330	346	372	395	333	21	3,131
Calculator (CA)	294	304	319	332	346	368	387	333	20	9,920
Arithmetic Table (AT)	294	306	319	332	348	368	390	335	20	882
Math Manipulatives (MM)	296	309	328	344	358	378	390	343	22	136
Assistive Device (AD)	302	306	327	350	362	385	387	346	23	25
Unlisted Modification (UM)	313	313	325	336	343	372	372	335	14	16
Unlisted Accommodation (UA)	294	304	319	332	344	368	385	333	20	635

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.C.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, November 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	17	342	27	35
Assistive Device No Interference	Specific Learning Disability	12	342	19	33
Beneficial Time	Autism	27	334	27	30
	Emotional Disturbance	110	319	37	22
	Mental Retardation	27	299	20	4
	Other Health Impairment	40	322	31	13
	Specific Learning Disability	280	322	27	15
Dictionary	Autism	36	327	26	17
	Deaf	27	315	24	7
	Emotional Disturbance	52	336	33	29
	Hard of Hearing	23	331	24	17
	Mental Retardation	40	309	24	5
	Orthopedic Impairment	15	320	19	13
	Other Health Impairment	96	330	27	23
	Specific Learning Disability	1,255	329	23	18
	Speech or Language Impairment	62	325	29	21
Essay Reponses (EO)	Orthopedic Impairment	15	333	24	33
	Specific Learning Disability	30	330	25	13
Essay Reponses (ER)	Specific Learning Disability	48	326	22	17
IEP or Section 504 Plan	Autism	641	328	27	19
	Deaf	236	312	21	6
	Emotional Disturbance	1,332	326	35	24
	Hard of Hearing	184	326	26	17
	Mental Retardation	470	305	20	3
	Multiple Disability	35	317	26	11
	Orthopedic Impairment	176	327	26	22
	Other Health Impairment	1,580	330	27	22
	Specific Learning Disability	14,544	326	24	16
	Speech or Language Impairment	844	328	24	17
	Traumatic Brain Injury	81	324	26	17
	Visual Impairment	63	334	28	29
Large Print Version	Visual Impairment	23	334	21	22
Oral Presentation	Autism	169	326	25	14
	Emotional Disturbance	141	327	28	25
	Hard of Hearing	40	322	24	8
	Mental Retardation	157	308	19	4
	Orthopedic Impairment	56	326	28	25
	Other Health Impairment	311	331	26	22
	Specific Learning Disability	3,776	329	24	19
	Speech or Language Impairment	250	328	24	15
	Traumatic Brain Injury	25	328	22	20
	Visual Impairment	16	336	34	31

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Sign Language	Deaf	62	314	19	5
	Hard of Hearing	17	317	25	6
Spell Checker or Grammar Checker	Autism	13	343	17	31
	Deaf	14	302	12	0
	Emotional Disturbance	17	345	26	41
	Mental Retardation	21	308	17	0
	Other Health Impairment	45	330	26	16
	Specific Learning Disability	279	331	22	22
	Speech or Language Impairment	32	325	23	19
Spell Checker or Grammar Checker Off	Autism	13	331	20	23
	Other Health Impairment	14	333	27	36
	Specific Learning Disability	88	328	24	19
Supervised Breaks	Autism	144	328	29	19
	Deaf	25	323	22	8
	Emotional Disturbance	388	322	34	21
	Hard of Hearing	29	329	23	17
	Mental Retardation	134	304	19	3
	Orthopedic Impairment	43	323	26	16
	Other Health Impairment	335	331	27	25
	Specific Learning Disability	2,860	326	25	16
	Speech or Language Impairment	172	328	26	20
	Visual Impairment	12	314	23	8
Test Over More Than One Day	Autism	51	332	26	25
	Deaf	38	307	22	3
	Emotional Disturbance	116	321	34	26
	Hard of Hearing	14	311	27	7
	Mental Retardation	32	303	18	3
	Other Health Impairment	84	333	29	32
	Specific Learning Disability	861	327	25	19
	Speech or Language Impairment	56	332	30	29
Tested at Home or Hospital	Specific Learning Disability	14	327	27	29
Transfer of Student T/B Responses to A/D	Orthopedic Impairment	13	332	27	31
	Other Health Impairment	11	316	33	9
	Specific Learning Disability	46	330	23	24
	Visual Impairment	22	334	31	27
Unlisted Accommodation	Autism	19	319	27	11
	Emotional Disturbance	28	315	31	18
	Other Health Impairment	55	331	27	16
	Specific Learning Disability	426	326	23	16
	Speech or Language Impairment	16	318	26	13
Unlisted Modification	Specific Learning Disability	12	323	18	0
Writing Only	Deaf	111	308	21	5
	Other Health Impairment	14	328	22	21

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Specific Learning Disability	179	326	24	18
	Speech or Language Impairment	11	314	28	9

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.C.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, November 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Autism	21	335	19	29
	Emotional Disturbance	29	343	21	34
	Hard of Hearing	12	334	15	8
	Mental Retardation	13	321	25	8
	Orthopedic Impairment	13	323	15	0
	Other Health Impairment	88	337	21	27
	Specific Learning Disability	661	334	20	23
	Speech or Language Impairment	22	340	21	32
Assistive Device	Specific Learning Disability	12	339	20	25
Beneficial Time	Autism	20	338	26	40
	Emotional Disturbance	120	326	30	19
	Mental Retardation	23	315	18	4
	Other Health Impairment	38	328	25	13
	Specific Learning Disability	241	330	23	23
Calculator	Autism	319	336	20	23
	Deaf	48	332	15	17
	Emotional Disturbance	488	334	23	24
	Hard of Hearing	79	338	21	20
	Mental Retardation	211	320	17	5
	Multiple Disability	18	329	24	17
	Orthopedic Impairment	93	332	22	17
	Other Health Impairment	809	336	21	25
	Specific Learning Disability	6,857	333	19	20
	Speech or Language Impairment	415	335	20	24
	Traumatic Brain Injury	40	337	26	28
	Visual Impairment	28	335	24	25
Dictionary for Math	Other Health Impairment	11	340	19	36
	Specific Learning Disability	187	331	17	16
	Speech or Language Impairment	16	342	27	38
IEP or Section 504 Plan	Autism	684	335	21	24
	Deaf	182	330	20	18
	Emotional Disturbance	1,555	331	27	23
	Hard of Hearing	170	336	21	24
	Mental Retardation	482	317	16	4
	Multiple Disability	38	323	24	16
	Orthopedic Impairment	199	333	21	18
	Other Health Impairment	1,851	335	21	24
	Specific Learning Disability	15,060	331	19	18
	Speech or Language Impairment	780	334	19	22
	Traumatic Brain Injury	77	333	23	21
	Visual Impairment	71	336	25	28
Large Print Version	Visual Impairment	27	340	24	33
Math Manipulatives	Other Health Impairment	17	341	26	35
	Specific Learning Disability	80	346	21	51
	Speech or Language Impairment	12	333	21	25

Table 2.C.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Autism	113	337	22	27
	Emotional Disturbance	93	336	25	30
	Hard of Hearing	15	335	16	20
	Mental Retardation	99	318	18	6
	Orthopedic Impairment	42	332	26	19
	Other Health Impairment	197	337	21	27
	Specific Learning Disability	2,217	334	21	21
	Speech or Language Impairment	160	334	21	24
	Traumatic Brain Injury	12	339	30	25
	Visual Impairment	11	345	27	36
Oral Responses Dictated to a Scribe	Specific Learning Disability	17	335	22	29
Sign Language	Deaf	70	334	17	23
	Hard of Hearing	18	329	13	6
	Specific Learning Disability	15	331	22	20
Supervised Breaks	Autism	128	334	22	26
	Emotional Disturbance	377	326	26	19
	Hard of Hearing	21	338	20	33
	Mental Retardation	125	316	16	3
	Orthopedic Impairment	44	336	24	25
	Other Health Impairment	328	333	22	20
	Specific Learning Disability	2,628	331	20	19
	Speech or Language Impairment	124	331	19	18
	Visual Impairment	12	326	21	25
Test Over More Than One Day	Autism	35	340	20	40
	Emotional Disturbance	110	321	25	15
	Mental Retardation	19	322	23	11
	Other Health Impairment	60	340	31	38
	Specific Learning Disability	527	334	23	24
	Speech or Language Impairment	31	335	17	16
Tested at Home or Hospital	Specific Learning Disability	13	331	17	15
Transfer of Student T/B Responses to A/D	Orthopedic Impairment	11	339	20	18
	Other Health Impairment	12	342	20	42
	Specific Learning Disability	34	338	26	21
	Visual Impairment	23	340	27	39
Unlisted Accommodation	Autism	20	334	21	15
	Emotional Disturbance	21	324	19	5
	Other Health Impairment	71	333	20	23
	Specific Learning Disability	432	333	19	20
Unlisted Modification	Specific Learning Disability	14	326	13	0

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.C.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, November 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	19	341	27	37
Assistive Device No Interference	English Only	16	337	17	19
Beneficial Time	English Only	334	321	32	19
	Reclassified Fluent English Proficient	17	335	21	18
	English Learner	149	320	27	13
	Unknown	12	321	29	25
Dictionary	English Only	796	330	25	21
	Initially Fluent English Proficient	32	331	24	19
	Reclassified Fluent English Proficient	57	337	27	32
	English Learner	771	325	23	13
Essay Reponses (EO)	English Only	47	329	24	19
	English Learner	20	330	29	20
Essay Reponses (ER)	English Only	52	328	24	19
	English Learner	36	321	20	8
IEP or Section 504 Plan	English Only	11,820	328	27	20
	Initially Fluent English Proficient	543	329	25	19
	Reclassified Fluent English Proficient	781	337	26	32
	English Learner	8,004	322	23	12
	Unknown	210	324	29	16
Large Print Version	English Only	37	332	27	30
	English Learner	13	320	20	8
Oral Presentation	English Only	2,649	330	26	22
	Initially Fluent English Proficient	122	330	23	16
	Reclassified Fluent English Proficient	217	335	24	26
	English Learner	2,113	324	23	13
	Unknown	36	324	31	22
Oral Responses Dictated to a Scribe	English Only	30	328	38	20
Sign Language	English Only	49	321	19	6
	English Learner	35	305	18	3
Spell Checker or Grammar Checker	English Only	251	330	24	20
	Initially Fluent English Proficient	12	324	24	8
	Reclassified Fluent English Proficient	16	346	20	44
	English Learner	169	329	23	21
Spell Checker or Grammar Checker Off	English Only	82	336	28	34
	English Learner	66	329	22	17

Table 2.C.9 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Supervised Breaks	English Only	2,435	326	28	20
	Initially Fluent English Proficient	120	328	26	18
	Reclassified Fluent English Proficient	175	335	26	29
	English Learner	1,591	322	24	13
	Unknown	32	325	29	22
Test Over More Than One Day	English Only	669	326	29	22
	Initially Fluent English Proficient	38	330	25	18
	Reclassified Fluent English Proficient	35	339	32	34
	English Learner	557	326	25	19
Tested at Home or Hospital	English Only	29	344	36	45
	English Learner	12	321	33	17
Transfer of Student T/B Responses to A/D	English Only	86	332	27	29
	English Learner	27	327	27	15
Unlisted Accommodation	English Only	332	329	27	20
	Initially Fluent English Proficient	11	329	25	9
	Reclassified Fluent English Proficient	22	338	16	9
	English Learner	243	320	24	11
Unlisted Modification	English Only	14	330	22	14
Writing Only	English Only	241	321	25	12
	Initially Fluent English Proficient	11	317	20	9
	English Learner	102	315	24	12

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.C.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, November 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	540	335	21	26
	Initially Fluent English Proficient	17	338	20	35
	Reclassified Fluent English Proficient	37	344	20	38
	English Learner	282	332	18	16
Assistive Device	English Only	19	349	25	58
Beneficial Time	English Only	334	328	25	20
	English Learner	126	328	24	21
Calculator	English Only	5,864	334	21	23
	Initially Fluent English Proficient	254	334	20	24
	Reclassified Fluent English Proficient	424	339	19	29
	English Learner	3,311	330	18	15
	Unknown	67	336	23	25
Dictionary for Math	English Only	127	336	19	23
	English Learner	113	329	17	12
IEP or Section 504 Plan	English Only	13,347	333	22	21
	Initially Fluent English Proficient	586	333	20	21
	Reclassified Fluent English Proficient	867	339	20	30
	English Learner	7,505	329	19	15
	Unknown	205	331	23	22
Large Print Version	English Only	36	338	20	25
	English Learner	13	333	25	31
Math Manipulatives	English Only	99	345	23	47
	English Learner	26	336	20	23
Oral Presentation	English Only	1,753	335	22	24
	Initially Fluent English Proficient	70	333	20	23
	Reclassified Fluent English Proficient	109	336	18	28
	English Learner	1,172	332	20	17
	Unknown	27	326	22	19
Oral Responses Dictated to a Scribe	English Only	21	333	22	24
	English Learner	16	337	24	31
Sign Language	English Only	57	331	18	21
	Initially Fluent English Proficient	15	336	20	20
	English Learner	38	333	15	16
Supervised Breaks	English Only	2,376	331	22	20
	Initially Fluent English Proficient	119	330	18	15
	Reclassified Fluent English Proficient	175	337	20	30
	English Learner	1,317	329	19	15

Table 2.C.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Unknown	21	324	18	14
Test Over More Than One Day	English Only	439	334	26	28
	Initially Fluent English Proficient	25	331	17	8
	Reclassified Fluent English Proficient	18	350	18	50
	English Learner	347	332	22	19
Tested at Home or Hospital	English Only	26	335	24	19
Transfer of Student T/B Responses to A/D	English Only	76	337	23	25
	English Learner	22	338	29	32
Unlisted Accommodation	English Only	391	334	21	23
	Initially Fluent English Proficient	12	333	13	25
	Reclassified Fluent English Proficient	20	340	18	30
	English Learner	209	329	17	14
Unlisted Modification	English Only	11	333	12	9

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.D: Results of Testing Variations and Disability Analyses—December 2011

Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	40	329	21	15
Modifications	94	327	26	17
All	134	327	24	16
Disability				
Autism	-	-	-	-
Deaf	-	-	-	-
Emotional Disturbance	-	-	-	-
Hard of Hearing	-	-	-	-
Mental Retardation	-	-	-	-
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	19	332	27	16
Specific Learning Disability	279	325	23	14
Speech or Language Impairment	-	-	-	-

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.2: Summary Statistics by Testing Variations and Disability—Mathematics, December 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	11	341	36	27
Modifications	118	336	23	19
All	129	336	24	20
Disability				
Autism	-	-	-	-
Deaf	-	-	-	-
Emotional Disturbance	-	-	-	-
Hard of Hearing	-	-	-	-
Mental Retardation	-	-	-	-
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	24	339	21	25
Specific Learning Disability	235	332	22	17
Speech or Language Impairment	-	-	-	-

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.3: Demographic Summary for All Examinees by Testing Variations—ELA, December 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 Plan	251	39	16	212	84	326	48	48	51	38	49	1.9
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-	-	-
Spell Checker/Grammar Checker Off (SO)	-	-	-	-	-	-	-	-	-	-	-	-
Essay Reponses (EO)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	13	2	15	11	85	326	52	53	51	32	47	1.9
Supervised Breaks (SB)	22	4	18	18	82	330	49	48	53	37	54	2.0
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-	-	-
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	18	4	22	14	78	333	52	53	56	38	51	2.0
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	80	14	18	66	83	327	52	47	50	38	50	1.9
Spell Checker or Grammar Checker (SC)	36	2	6	34	94	318	49	39	43	33	47	1.9
Essay Reponses (ER)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-	-	-
Writing Only (WO)	-	-	-	-	-	-	-	-	-	-	-	-

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.D.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, December 2011

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 Plan	218	36	17	182	83	333	44	41	42	37	34
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	-	-	-	-	-	-	-	-	-	-	-
Supervised Breaks (SB)	-	-	-	-	-	-	-	-	-	-	-
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-	-
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	60	12	20	48	80	336	47	44	42	39	35
Calculator (CA)	118	23	19	95	81	336	46	45	43	40	34
Arithmetic Table (AT)	15	1	7	14	93	328	43	40	38	34	29
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-	-

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.D.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, December 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	286	311	325	343	362	384	326	23	251
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-
Spell Checker/Grammar Checker Off (SO)	-	-	-	-	-	-	-	-	-	-
Essay Reponses (EO)	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	275	275	317	323	335	384	384	326	26	13
Supervised Breaks (SB)	286	293	315	330	345	362	378	330	23	22
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	275	275	313	336	349	402	402	333	31	18
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	275	286	309	322	345	376	402	327	26	80
Spell Checker or Grammar Checker (SC)	275	275	299	319	334	359	378	318	23	36
Essay Reponses (ER)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-
Writing Only (WO)	-	-	-	-	-	-	-	-	-	-

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.D.6: Percentiles of Scale Scores for Students with Testing Variations— Mathematics, December 2011

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	296	302	318	330	344	376	399	333	22	218
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	-	-	-	-	-	-	-	-	-	-
Supervised Breaks (SB)	-	-	-	-	-	-	-	-	-	-
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-
Tested at Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	293	306	320	328	345	390	434	336	27	60
Calculator (CA)	299	307	320	330	344	382	396	336	23	118
Arithmetic Table (AT)	302	302	311	324	343	374	374	328	19	15
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.D.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, December 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Dictionary	Specific Learning Disability	14	323	23	7
IEP or Section 504 Plan	Other Health Impairment	15	335	28	20
	Specific Learning Disability	200	324	23	13
Oral Presentation	Specific Learning Disability	65	324	25	14
Spell Checker or Grammar Checker	Specific Learning Disability	29	318	24	7
Supervised Breaks	Specific Learning Disability	17	328	25	18
Test Over More Than One Day	Specific Learning Disability	12	326	27	17

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, December 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	13	324	16	0
Calculator	Other Health Impairment	13	335	21	23
	Specific Learning Disability	92	335	23	17
IEP or Section 504 Plan	Other Health Impairment	19	338	19	21
	Specific Learning Disability	169	331	23	15
Oral Presentation	Specific Learning Disability	50	337	28	22

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, December 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
IEP or Section 504 Plan	English Only	109	330	24	21
	English Learner	125	324	23	12
Oral Presentation	English Only	26	331	35	35
	English Learner	50	325	22	10
Spell Checker or Grammar Checker	English Learner	26	318	19	0
Supervised Breaks	English Learner	11	324	21	9

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, December 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Calculator	English Only	58	336	23	26
	English Learner	49	336	23	14
IEP or Section 504 Plan	English Only	107	333	20	19
	Reclassified Fluent English Proficient	14	346	29	36
	English Learner	89	331	22	12
Oral Presentation	English Only	26	334	22	23
	English Learner	28	335	27	14

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.E: Results of Testing Variations and Disability Analyses—February 2012

Table 2.E.1: Summary Statistics by Testing Variations and Disability—ELA, February 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,530	327	31	22
Modifications	3,710	326	25	18
All	7,240	326	28	20
Disability				
Unknown	-	-	-	-
Autism	807	343	40	39
Deaf	210	315	33	12
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,669	333	39	32
Hard of Hearing	257	331	35	25
Mental Retardation	539	303	20	3
Multiple Disability	51	312	29	14
Orthopedic Impairment	188	330	33	25
Other Health Impairment	1,950	338	34	37
Specific Learning Disability	16,387	328	27	21
Speech or Language Impairment	1,471	345	38	40
Traumatic Brain Injury	89	327	32	18
Visual Impairment	91	350	43	46

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.E.2: Summary Statistics by Testing Variations and Disability—Mathematics, February 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,046	333	26	23
Modifications	6,017	333	22	20
All	9,063	333	24	21
Disability				
Unknown	-	-	-	-
Autism	789	350	37	42
Deaf	162	337	31	25
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,837	335	31	27
Hard of Hearing	228	344	33	32
Mental Retardation	529	314	17	3
Multiple Disability	54	322	22	11
Orthopedic Impairment	198	333	27	22
Other Health Impairment	2,126	341	29	33
Specific Learning Disability	16,082	334	24	23
Speech or Language Impairment	1,396	354	39	46
Traumatic Brain Injury	73	336	28	21
Visual Impairment	95	351	41	45

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.E.3: Demographic Summary for All Examinees by Testing Variations—ELA, February 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 Plan	17,992	4,058	23	13,934	77	328	56	47	52	41	43	1.8
Transfer of T/B Responses to A/D (TS)	141	56	40	85	60	341	63	56	57	46	50	2.0
Oral Responses Dictated to a Scribe (OR)	50	14	28	36	72	327	55	47	50	41	41	1.8
Spell Checker/Grammar Checker Off (SO)	159	70	44	89	56	354	68	60	63	54	57	2.2
Essay Reponses (EO)	55	20	36	35	64	337	62	52	56	46	45	2.1
Assistive Device No Interference (AN)	22	11	50	11	50	354	76	59	65	57	61	2.2
Braille Version (BV)	12	3	25	9	75	323	58	40	47	41	51	1.7
Large Print Version (LV)	57	22	39	35	61	342	63	55	59	48	51	2.0
Test Over More Than One Day (TD)	887	152	17	735	83	323	54	43	49	39	40	1.8
Supervised Breaks (SB)	3,297	667	20	2,630	80	326	54	45	51	40	42	1.8
Beneficial Time (BT)	529	83	16	446	84	320	49	42	47	37	39	1.7
Tested at Home or Hospital (HH)	35	17	49	18	51	343	67	59	61	46	52	1.8
Dictionary (DI)	1,175	179	15	996	85	325	59	44	50	38	41	1.9
Sign Language (SL)	63	3	5	60	95	311	51	35	40	28	33	1.7
Oral Presentation (OP)	3,009	557	19	2,452	81	326	57	46	51	40	41	1.8
Spell Checker or Grammar Checker (SC)	208	39	19	169	81	327	58	45	51	39	41	2.0
Essay Reponses (ER)	77	14	18	63	82	323	54	43	46	36	37	2.1
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	57	14	25	43	75	322	57	44	49	39	37	1.6
Unlisted Accommodation (UA)	620	119	19	501	81	327	55	45	53	40	42	1.8
Writing Only (WO)	234	19	8	215	92	313	45	39	41	33	34	1.7

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.E.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, February 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 Plan	18,274	4,351	24	13,923	76	334	50	47	43	39	36
Transfer of T/B Responses to A/D (TS)	139	57	41	82	59	347	56	54	53	46	43
Oral Responses Dictated to a Scribe (OR)	28	10	36	18	64	339	50	51	48	46	34
Braille Version (BV)	21	4	19	17	81	329	40	48	38	36	28
Large Print Version (LV))	52	18	35	34	65	346	56	50	56	43	44
Test Over More Than One Day (TD)	577	119	21	458	79	331	47	45	42	37	33
Supervised Breaks (SB)	2,832	613	22	2,219	78	332	48	45	42	38	35
Beneficial Time (BT)	464	71	15	393	85	328	43	44	38	36	32
Tested at Home or Hospital (HH)	44	17	39	27	61	344	58	55	48	46	38
Dictionary for Math (DM)	251	52	21	199	79	334	49	50	41	39	33
Sign Language (SL)	42	9	21	33	79	336	46	53	44	38	37
Oral Presentation (OP)	1,956	422	22	1,534	78	334	48	49	42	38	36
Calculator (CA)	5,891	1,188	20	4,703	80	333	48	50	41	38	35
Arithmetic Table (AT)	615	150	24	465	76	335	50	50	42	40	36
Math Manipulative (MM)	59	18	31	41	69	336	51	48	43	40	38
Assistive Device (AD)	11	6	55	5	45	352	59	62	55	51	45
Unlisted Modification (UM)	69	13	19	56	81	332	43	48	40	38	36
Unlisted Accommodation (UA)	571	116	20	455	80	331	48	44	41	38	35

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.E.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, February 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	285	307	324	346	382	414	328	30	17,992
Transfer of T/B Responses to A/D (TS)	275	281	313	336	362	406	445	341	38	141
Oral Responses Dictated to a Scribe (OR)	275	275	305	320	354	390	414	327	36	50
Spell Checker/Grammar Checker Off (SO)	283	291	316	344	393	445	450	354	46	159
Essay Reponses (EO)	275	275	313	338	356	396	439	337	35	55
Assistive Device No Interference (AN)	275	316	334	351	372	410	439	354	34	22
Braille Version (BV)	275	275	291	314	355	389	389	323	39	12
Large Print Version (LV)	281	291	314	336	358	428	445	342	38	57
Test Over More Than One Day (TD)	275	279	303	320	342	369	396	323	27	887
Supervised Breaks (SB)	275	285	305	322	344	379	403	326	29	3,297
Beneficial Time (BT)	275	276	299	316	340	374	399	320	29	529
Tested at Home or Hospital (HH)	275	275	307	344	379	414	450	343	43	35
Dictionary (DI)	275	289	307	322	340	365	393	325	24	1,175
Sign Language (SL)	275	285	299	307	318	348	379	311	19	63
Oral Presentation (OP)	275	289	309	324	342	372	393	326	25	3,009
Spell Checker or Grammar Checker (SC)	283	295	309	322	341	377	414	327	26	208
Essay Reponses (ER)	276	287	301	318	342	377	433	323	29	77
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	275	275	297	316	344	377	393	322	32	57
Unlisted Accommodation (UA)	275	287	307	324	344	377	410	327	28	620
Writing Only (WO)	275	279	295	311	330	356	372	313	24	234

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.E.6: Percentiles of Scale Scores for Students with Testing Variations — Mathematics, February 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	287	301	317	330	348	383	421	334	26	18,274
Transfer of T/B Responses to A/D (TS)	280	301	323	343	364	421	438	347	34	139
Oral Responses Dictated to a Scribe (OR)	275	308	318	337	357	377	405	339	27	28
Braille Version (BV)	296	296	317	328	346	355	355	329	18	21
Large Print Version (LV)	296	301	328	342	356	431	450	346	33	52
Test Over More Than One Day (TD)	275	296	312	328	346	377	416	331	26	577
Supervised Breaks (SB)	280	298	315	328	346	377	412	332	25	2,832
Beneficial Time (BT)	275	296	312	325	341	372	402	328	24	464
Tested at Home or Hospital (HH)	275	280	320	344	363	402	431	344	37	44
Dictionary for Math (DM)	296	301	319	332	346	372	421	334	23	251
Sign Language (SL)	298	306	317	338	346	374	399	336	23	42
Oral Presentation (OP)	290	301	317	330	346	379	409	334	24	1,956
Calculator (CA)	290	301	319	332	345	370	402	333	22	5,891
Arithmetic Table (AT)	284	301	319	334	348	372	409	335	23	615
Math Manipulatives (MM)	280	293	319	330	353	377	446	336	27	59
Assistive Device (AD)	326	326	336	353	357	412	412	352	24	11
Unlisted Modification (UM)	275	306	315	325	339	388	409	332	26	69
Unlisted Accommodation (UA)	290	298	315	328	345	370	405	331	23	571

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported

Table 2.E.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, February 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	Autism	15	325	37	27
	Emotional Disturbance	129	315	32	16
	Mental Retardation	28	297	14	0
	Other Health Impairment	44	327	32	25
	Specific Learning Disability	266	323	26	14
Dictionary	Autism	28	327	24	14
	Emotional Disturbance	31	319	31	10
	Mental Retardation	32	313	22	6
	Orthopedic Impairment	13	319	22	8
	Other Health Impairment	71	322	23	14
	Specific Learning Disability	862	326	23	16
	Speech or Language Impairment	29	318	20	7
Essay Reponses (EO)	Specific Learning Disability	20	331	22	30
Essay Reponses (ER)	Deaf	16	317	33	13
	Specific Learning Disability	35	328	29	20
IEP or Section 504 Plan	Autism	546	339	39	35
	Deaf	143	311	29	8
	Emotional Disturbance	1,295	331	37	30
	Hard of Hearing	175	324	29	15
	Mental Retardation	403	303	19	3
	Multiple Disability	38	313	31	16
	Orthopedic Impairment	144	325	30	19
	Other Health Impairment	1,306	335	32	32
	Specific Learning Disability	11,753	326	26	19
	Speech or Language Impairment	723	329	27	21
	Traumatic Brain Injury	61	326	32	16
	Visual Impairment	69	343	39	41
Large Print Version	Visual Impairment	28	345	33	39
Oral Presentation	Autism	84	325	30	18
	Emotional Disturbance	85	325	31	21
	Hard of Hearing	31	327	21	19
	Mental Retardation	103	307	20	4
	Orthopedic Impairment	33	325	25	18
	Other Health Impairment	182	330	26	21
	Specific Learning Disability	2,191	327	24	19
	Speech or Language Impairment	140	326	22	16
	Traumatic Brain Injury	15	318	29	7
Oral Responses Dictated to a Scribe	Specific Learning Disability	14	312	30	14
Sign Language	Deaf	43	309	16	5
	Hard of Hearing	12	313	19	0
Spell Checker or Grammar Checker	Deaf	16	308	12	0
	Other Health Impairment	13	337	29	46
	Specific Learning Disability	126	327	24	17
	Speech or Language Impairment	15	321	20	7
Spell Checker or Grammar Checker Off	Autism	17	386	49	71

Table 2.E.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Other Health Impairment	13	371	45	77
	Specific Learning Disability	57	350	41	44
Supervised Breaks	Autism	111	329	38	25
	Deaf	20	322	37	20
	Emotional Disturbance	364	324	34	24
	Hard of Hearing	32	315	25	6
	Mental Retardation	86	307	21	5
	Orthopedic Impairment	32	326	34	25
	Other Health Impairment	260	331	31	29
	Specific Learning Disability	2,042	325	26	18
	Speech or Language Impairment	122	328	23	19
	Traumatic Brain Injury	13	319	36	8
	Visual Impairment	15	334	38	27
Test Over More Than One Day	Autism	23	339	43	35
	Deaf	16	309	13	6
	Emotional Disturbance	94	321	33	21
	Mental Retardation	21	310	25	5
	Other Health Impairment	76	328	29	20
	Specific Learning Disability	552	323	25	16
	Speech or Language Impairment	31	322	24	13
Transfer of Student T/B Responses to A/D	Other Health Impairment	15	333	28	27
	Specific Learning Disability	48	339	37	42
	Visual Impairment	30	344	35	33
Unlisted Accommodation	Autism	21	335	39	29
	Emotional Disturbance	49	329	37	29
	Mental Retardation	12	302	22	8
	Other Health Impairment	42	327	21	19
	Specific Learning Disability	423	325	23	16
	Speech or Language Impairment	15	326	22	13
Unlisted Modification	Emotional Disturbance	11	303	26	0
	Specific Learning Disability	26	331	33	35
Writing Only	Autism	12	304	21	0
	Deaf	62	303	22	2
	Specific Learning Disability	101	318	22	9

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.E.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, February 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Emotional Disturbance	16	327	24	19
	Mental Retardation	11	323	22	18
	Other Health Impairment	55	342	28	35
	Specific Learning Disability	468	334	22	24
	Speech or Language Impairment	12	334	21	25
Beneficial Time	Autism	12	336	33	25
	Emotional Disturbance	109	323	26	15
	Mental Retardation	23	311	14	0
	Other Health Impairment	40	328	24	15
	Specific Learning Disability	231	330	21	16
Calculator	Autism	153	339	25	28
	Deaf	24	334	21	17
	Emotional Disturbance	289	330	25	18
	Hard of Hearing	52	337	23	19
	Mental Retardation	141	318	16	4
	Multiple Disability	17	322	18	0
	Orthopedic Impairment	59	332	20	24
	Other Health Impairment	473	337	25	27
	Specific Learning Disability	4,073	333	21	19
	Speech or Language Impairment	232	336	23	20
	Traumatic Brain Injury	21	339	28	24
	Visual Impairment	15	329	22	27
Dictionary for Math	Other Health Impairment	11	319	17	9
	Specific Learning Disability	190	333	21	19
IEP or Section 504 Plan	Autism	544	347	35	38
	Deaf	107	334	27	20
	Emotional Disturbance	1,406	333	30	25
	Hard of Hearing	158	337	29	23
	Mental Retardation	394	314	16	3
	Multiple Disability	39	320	19	5
	Orthopedic Impairment	154	330	25	19
	Other Health Impairment	1,465	338	27	30
	Specific Learning Disability	11,770	333	23	21
	Speech or Language Impairment	695	339	28	29
	Traumatic Brain Injury	52	335	28	19
	Visual Impairment	75	343	35	37
Large Print Version	Visual Impairment	24	348	30	33
Math Manipulatives	Specific Learning Disability	35	333	22	29

Table 2.E.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Autism	63	340	27	30
	Emotional Disturbance	72	331	31	25
	Hard of Hearing	15	336	17	13
	Mental Retardation	73	320	16	3
	Orthopedic Impairment	26	330	21	19
	Other Health Impairment	142	335	26	24
	Specific Learning Disability	1,353	334	23	22
	Speech or Language Impairment	73	338	24	19
Sign Language	Deaf	24	333	22	17
Supervised Breaks	Autism	101	339	31	30
	Emotional Disturbance	306	327	27	21
	Hard of Hearing	19	340	35	37
	Mental Retardation	75	317	16	3
	Orthopedic Impairment	27	322	22	11
	Other Health Impairment	265	334	25	26
	Specific Learning Disability	1,703	332	23	20
	Speech or Language Impairment	104	338	27	26
	Visual Impairment	18	333	24	39
Test Over More Than One Day	Autism	18	355	43	39
	Emotional Disturbance	70	328	25	20
	Other Health Impairment	46	324	27	15
	Specific Learning Disability	366	332	22	20
	Speech or Language Impairment	16	327	30	13
Tested at Home or Hospital	Specific Learning Disability	14	334	28	29
Transfer of Student T/B Responses to A/D	Emotional Disturbance	11	344	34	36
	Other Health Impairment	15	352	32	53
	Specific Learning Disability	46	350	30	48
	Visual Impairment	33	345	32	39
Unlisted Accommodation	Autism	21	337	33	24
	Emotional Disturbance	47	322	21	15
	Other Health Impairment	44	331	23	23
	Specific Learning Disability	385	331	21	18
Unlisted Modification	Emotional Disturbance	12	315	12	0
	Specific Learning Disability	37	335	23	22

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.E.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, February 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	13	362	36	54
Beneficial Time	English Only	352	320	31	18
	Initially Fluent English Proficient	15	333	25	27
	English Learner	148	318	21	7
Dictionary	English Only	540	326	26	19
	Initially Fluent English Proficient	23	332	23	17
	Reclassified Fluent English Proficient	31	333	29	29
	English Learner	575	323	22	11
Essay Reponses (EO)	English Only	41	341	34	39
Essay Reponses (ER)	English Only	49	323	27	16
	English Learner	23	322	34	17
IEP or Section 504 Plan	English Only	10,981	331	32	27
	Initially Fluent English Proficient	403	331	31	26
	Reclassified Fluent English Proficient	617	340	29	39
	English Learner	5,826	320	23	11
	Unknown	165	331	32	30
Large Print Version	English Only	40	347	40	45
	English Learner	12	322	17	8
Oral Presentation	English Only	1,552	329	26	22
	Initially Fluent English Proficient	71	327	25	20
	Reclassified Fluent English Proficient	108	336	26	25
	English Learner	1,265	322	23	13
	Unknown	13	336	32	31
Oral Responses Dictated to a Scribe	English Only	37	327	38	30
	English Learner	12	323	28	17
Sign Language	English Only	46	312	19	4
	English Learner	14	303	9	0
Spell Checker or Grammar Checker	English Only	112	332	29	24
	English Learner	86	321	19	12
Spell Checker or Grammar Checker Off	English Only	108	365	46	56
	English Learner	39	326	26	10
Supervised Breaks	English Only	2,072	328	31	24
	Initially Fluent English Proficient	82	331	29	27
	Reclassified Fluent English Proficient	74	336	30	31
	English Learner	1,055	320	23	11
	Unknown	14	330	26	21

Table 2.E.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Test Over More Than One Day	English Only	525	323	29	19
	Initially Fluent English Proficient	33	327	26	24
	Reclassified Fluent English Proficient	14	329	32	36
	English Learner	315	322	23	13
Tested at Home or Hospital	English Only	24	349	47	58
Transfer of Student T/B Responses to A/D	English Only	104	341	39	41
	Reclassified Fluent English Proficient	11	366	48	64
	English Learner	21	320	21	10
Unlisted Accommodation	English Only	377	329	31	23
	Initially Fluent English Proficient	13	329	30	23
	English Learner	209	322	22	11
	Unknown	12	337	23	33
Unlisted Modification	English Only	37	318	33	27
	English Learner	17	323	24	12
Writing Only	English Only	168	313	24	6
	English Learner	53	313	24	13

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.E.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, February 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	366	335	23	25
	Initially Fluent English Proficient	11	332	24	36
	Reclassified Fluent English Proficient	26	350	25	42
	English Learner	211	333	21	21
Beneficial Time	English Only	327	327	25	15
	Initially Fluent English Proficient	15	347	22	33
	English Learner	110	326	18	10
Braille Version	English Only	14	332	21	29
Calculator	English Only	3,415	334	23	21
	Initially Fluent English Proficient	147	336	28	23
	Reclassified Fluent English Proficient	232	343	22	34
	English Learner	2,038	330	19	16
	Unknown	59	343	33	34
Dictionary for Math	English Only	119	337	27	28
	English Learner	119	331	19	15
IEP or Section 504 Plan	English Only	11,568	336	28	26
	Initially Fluent English Proficient	431	337	29	26
	Reclassified Fluent English Proficient	701	346	27	39
	English Learner	5,397	330	21	16
	Unknown	177	336	30	24
Large Print Version	English Only	39	348	36	33
Math Manipulatives	English Only	48	339	28	33
Oral Presentation	English Only	1,098	335	26	24
	Initially Fluent English Proficient	58	337	30	26
	Reclassified Fluent English Proficient	60	341	19	42
	English Learner	730	330	21	16
Oral Responses Dictated to a Scribe	English Only	19	339	30	32
Sign Language	English Only	34	337	24	24
Supervised Breaks	English Only	1,854	333	27	23
	Initially Fluent English Proficient	82	337	27	23
	Reclassified Fluent English Proficient	78	342	28	35
	English Learner	802	329	21	16
	Unknown	16	339	35	25
Test Over More Than One Day	English Only	335	331	28	20
	Reclassified Fluent English Proficient	25	336	28	20
	English Learner	210	331	21	20
Tested at Home or Hospital	English Only	30	343	37	40

Table 2.E.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Transfer of Student T/B Responses to A/D	English Only	106	345	35	36
	Reclassified Fluent English Proficient	13	364	28	77
	English Learner	15	342	21	40
Unlisted Accommodation	English Only	381	332	25	24
	Initially Fluent English Proficient	13	333	20	15
	English Learner	161	328	19	13
Unlisted Modification	English Only	45	329	27	16
	English Learner	20	333	21	20

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.F: Results of Testing Variations and Disability Analyses—March 2012

Table 2.F.1: Summary Statistics by Testing Variations and Disability—ELA, March 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	6,824	331	34	30
Modifications	5,157	324	26	17
All	11,981	328	31	24
Disability				
Unknown	-	-	-	-
Autism	2,042	349	42	47
Deaf	377	309	33	10
Deaf-Blindness	-	-	-	-
Emotional Disturbance	2,829	333	40	35
Hard of Hearing	512	344	40	43
Mental Retardation	731	299	21	2
Multiple Disability	48	308	24	4
Orthopedic Impairment	432	343	39	44
Other Health Impairment	4,520	344	35	45
Specific Learning Disability	30,382	330	29	26
Speech or Language Impairment	2,459	348	41	44
Traumatic Brain Injury	160	331	35	27
Visual Impairment	185	359	44	59

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.F.2: Summary Statistics by Testing Variations and Disability—Mathematics, March 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	5,771	340	30	31
Modifications	8,808	336	22	23
All	14,579	337	26	26
Disability				
Unknown	-	-	-	-
Autism	2,013	361	41	52
Deaf	311	337	31	25
Deaf-Blindness	-	-	-	-
Emotional Disturbance	2,910	338	32	30
Hard of Hearing	491	355	36	49
Mental Retardation	704	316	17	3
Multiple Disability	48	322	26	8
Orthopedic Impairment	431	349	35	43
Other Health Impairment	4,572	349	32	43
Specific Learning Disability	29,959	339	26	29
Speech or Language Impairment	2,388	359	39	51
Traumatic Brain Injury	141	343	31	29
Visual Impairment	191	363	40	60

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.F.3: Demographic Summary for All Examinees by Testing Variations—ELA, March 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 Plan	32,184	9,312	29	22,872	71	332	59	50	59	44	50	1.9
Transfer of T/B Responses to A/D (TS)	219	106	48	113	52	351	70	59	70	54	58	2.1
Oral Responses Dictated to a Scribe (OR)	79	28	35	51	65	335	61	51	62	46	49	1.8
Spell Checker/Grammar Checker Off (SO)	251	140	56	111	44	359	74	65	73	55	62	2.3
Essay Reponses (EO)	90	24	27	66	73	328	55	48	57	43	44	1.9
Assistive Device No Interference (AN)	144	64	44	80	56	341	64	56	61	48	55	2.0
Braille Version (BV)	23	13	57	10	43	359	74	62	68	54	66	2.3
Large Print Version (LV)	132	66	50	66	50	349	69	58	69	52	59	2.0
Test Over More Than One Day (TD)	1,943	614	32	1,329	68	333	60	51	59	45	51	1.9
Supervised Breaks (SB)	5,792	1,397	24	4,395	76	327	56	47	56	41	48	1.9
Beneficial Time (BT)	1,066	233	22	833	78	323	55	44	54	39	45	1.8
Tested at Home or Hospital (HH)	74	41	55	33	45	355	69	61	71	59	63	2.0
Dictionary (DI)	1,715	256	15	1,459	85	323	57	44	54	38	45	1.9
Sign Language (SL)	86	4	5	82	95	307	49	38	37	32	37	1.6
Oral Presentation (OP)	4,080	699	17	3,381	83	324	56	45	55	40	46	1.9
Spell Checker or Grammar Checker (SC)	381	67	18	314	82	326	55	45	55	41	48	2.0
Essay Reponses (ER)	90	19	21	71	79	327	58	46	54	42	44	2.1
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	45	19	42	26	58	336	60	54	63	49	50	1.8
Unlisted Accommodation (UA)	741	223	30	518	70	332	60	50	60	44	51	1.9
Writing Only (WO)	479	83	17	396	83	320	51	43	51	37	44	1.8

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.F.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, March 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 Plan	32,084	9,710	30	22,374	70	341	50	48	49	43	39
Transfer of T/B Responses to A/D (TS)	208	94	45	114	55	355	58	55	56	52	48
Oral Responses Dictated to a Scribe (OR)	65	27	42	38	58	347	52	52	53	48	42
Braille Version (BV)	24	7	29	17	71	341	50	50	46	45	37
Large Print Version (LV)	122	58	48	64	52	355	57	56	58	53	47
Test Over More Than One Day (TD)	1,393	505	36	888	64	343	52	49	51	45	42
Supervised Breaks (SB)	5,200	1,361	26	3,839	74	337	48	46	46	41	37
Beneficial Time (BT)	956	223	23	733	77	333	45	44	43	39	34
Tested at Home or Hospital (HH)	62	30	48	32	52	356	56	59	57	53	44
Dictionary for Math (DM)	254	54	21	200	79	334	44	48	43	39	34
Sign Language (SL)	181	32	18	149	82	331	42	41	40	39	38
Oral Presentation (OP)	2,677	597	22	2,080	78	336	47	46	45	41	36
Calculator (CA)	8,711	1,958	22	6,753	78	335	46	47	44	41	35
Arithmetic Table (AT)	584	110	19	474	81	334	45	47	43	40	33
Math Manipulatives (MM)	82	19	23	63	77	338	45	52	45	43	36
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	20	13	65	7	35	358	62	56	56	60	51
Unlisted Accommodation (UA)	694	235	34	459	66	342	52	49	50	44	39

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.F.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, March 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	283	307	329	352	391	423	332	33	32,184
Transfer of T/B Responses to A/D (TS)	275	291	321	346	379	423	450	351	41	219
Oral Responses Dictated to a Scribe (OR)	275	275	297	333	369	409	428	335	42	79
Spell Checker/Grammar Checker Off (SO)	275	293	329	355	388	433	450	359	41	251
Essay Reponses (EO)	275	275	299	322	350	409	428	328	38	90
Assistive Device No Interference (AN)	275	285	312	342	366	405	433	341	38	144
Braille Version (BV)	291	295	309	371	379	450	450	359	48	23
Large Print Version (LV)	275	283	317	349	378	418	450	349	42	132
Test Over More Than One Day (TD)	275	283	309	331	355	388	413	333	32	1,943
Supervised Breaks (SB)	275	279	303	325	348	382	413	327	32	5,792
Beneficial Time (BT)	275	275	297	319	346	379	418	323	34	1,066
Tested at Home or Hospital (HH)	275	276	319	357	398	433	450	355	48	74
Dictionary (DI)	275	283	303	321	339	366	388	323	26	1,715
Sign Language (SL)	275	275	291	303	319	346	409	307	24	86
Oral Presentation (OP)	275	285	305	323	342	369	388	324	26	4,080
Spell Checker or Grammar Checker (SC)	275	287	311	325	342	369	401	326	26	381
Essay Reponses (ER)	275	289	305	323	342	379	418	327	29	90
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	275	279	303	339	369	388	409	336	39	45
Unlisted Accommodation (UA)	275	283	307	331	355	388	413	332	33	741
Writing Only (WO)	275	281	295	315	339	377	418	320	30	479

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.F.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, March 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	291	304	320	335	355	398	440	341	29	32,084
Transfer of T/B Responses to A/D (TS)	291	304	327	346	374	433	450	355	38	208
Oral Responses Dictated to a Scribe (OR)	291	296	318	342	368	407	450	347	37	65
Braille Version (BV)	291	304	317	327	365	418	423	341	35	24
Large Print Version (LV)	294	309	327	345	374	440	450	355	37	122
Test Over More Than One Day (TD)	284	302	322	339	362	401	433	343	31	1,393
Supervised Breaks (SB)	284	299	318	333	351	390	423	337	28	5,200
Beneficial Time (BT)	275	294	313	327	348	390	410	333	28	956
Tested at Home or Hospital (HH)	281	302	318	342	387	448	450	356	47	62
Dictionary for Math (DM)	288	299	318	333	346	372	414	334	24	254
Sign Language (SL)	294	304	315	326	344	374	423	331	23	181
Oral Presentation (OP)	291	304	320	333	348	380	410	336	24	2,677
Calculator (CA)	291	304	320	333	349	374	398	335	22	8,711
Arithmetic Table (AT)	288	304	320	331	346	368	398	334	21	584
Math Manipulative (MM)	296	307	318	334	349	392	450	338	28	82
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	318	319	329	358	376	423	450	358	33	20
Unlisted Accommodation (UA)	288	304	320	337	360	395	423	342	28	694

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.F.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, March 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	Deaf	11	305	28	9
	Specific Learning Disability	81	343	32	48
	Visual Impairment	12	364	45	58
Beneficial Time	Autism	43	327	40	26
	Deaf	12	307	23	8
	Emotional Disturbance	233	321	38	21
	Mental Retardation	32	292	14	0
	Other Health Impairment	82	330	33	29
	Specific Learning Disability	521	322	29	18
	Speech or Language Impairment	24	318	22	8
Braille Version	Visual Impairment	21	356	46	57
Dictionary	Autism	34	327	28	18
	Deaf	13	314	23	8
	Emotional Disturbance	58	319	28	12
	Hard of Hearing	22	315	27	5
	Mental Retardation	44	301	18	2
	Orthopedic Impairment	20	319	17	5
	Other Health Impairment	99	328	27	22
	Specific Learning Disability	1,249	324	25	15
	Speech or Language Impairment	57	324	26	18
Essay Reponses (EO)	Orthopedic Impairment	24	337	42	33
	Other Health Impairment	11	321	36	18
	Specific Learning Disability	25	319	28	20
Essay Reponses (ER)	Orthopedic Impairment	11	340	46	27
	Specific Learning Disability	49	326	23	18
IEP or Section 504 Plan	Autism	1,369	343	41	42
	Deaf	261	309	30	10
	Emotional Disturbance	2,177	330	39	31
	Hard of Hearing	312	334	34	29
	Mental Retardation	543	299	19	2
	Multiple Disability	36	308	25	6
	Orthopedic Impairment	296	338	39	37
	Other Health Impairment	2,995	341	35	41
	Specific Learning Disability	20,340	328	29	24
	Speech or Language Impairment	1,082	331	31	25
	Traumatic Brain Injury	115	327	33	23
	Visual Impairment	152	357	43	57
Large Print Version	Orthopedic Impairment	11	342	34	45
	Specific Learning Disability	19	327	29	32
	Visual Impairment	72	357	42	56
Oral Presentation	Autism	150	324	29	18
	Emotional Disturbance	139	320	30	19
	Hard of Hearing	49	322	24	14
	Mental Retardation	126	307	24	7

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Orthopedic Impairment	44	318	24	11
	Other Health Impairment	267	327	27	20
	Specific Learning Disability	2,902	325	25	17
	Speech or Language Impairment	170	325	24	17
	Traumatic Brain Injury	22	319	24	14
	Visual Impairment	20	340	46	45
Oral Responses Dictated to a Scribe	Orthopedic Impairment	32	344	43	44
	Specific Learning Disability	12	345	31	33
Sign Language	Deaf	56	308	25	5
	Hard of Hearing	24	301	20	0
Spell Checker or Grammar Checker	Autism	13	327	21	8
	Deaf	13	302	13	0
	Emotional Disturbance	11	319	25	9
	Other Health Impairment	33	331	34	24
	Specific Learning Disability	247	328	26	18
	Speech or Language Impairment	26	328	17	23
Spell Checker or Grammar Checker Off	Autism	24	363	48	63
	Emotional Disturbance	21	355	44	48
	Orthopedic Impairment	13	361	37	69
	Other Health Impairment	26	359	32	54
	Specific Learning Disability	84	353	32	54
	Speech or Language Impairment	12	349	21	50
	Visual Impairment	15	361	47	53
Supervised Breaks	Autism	245	334	40	33
	Deaf	46	304	23	7
	Emotional Disturbance	559	325	40	27
	Hard of Hearing	65	323	35	20
	Mental Retardation	138	300	22	4
	Multiple Disability	12	311	23	8
	Orthopedic Impairment	50	333	39	30
	Other Health Impairment	455	336	33	38
	Specific Learning Disability	3,629	327	28	21
	Speech or Language Impairment	203	331	30	25
	Traumatic Brain Injury	31	326	29	16
	Visual Impairment	26	348	39	50
Test Over More Than One Day	Autism	85	341	36	44
	Deaf	48	300	21	4
	Emotional Disturbance	126	326	36	28
	Hard of Hearing	31	323	33	19
	Mental Retardation	43	301	18	2
	Orthopedic Impairment	21	355	41	62
	Other Health Impairment	164	344	32	43
	Specific Learning Disability	1,237	333	30	31

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Speech or Language Impairment	87	338	29	34
	Traumatic Brain Injury	12	336	27	33
	Visual Impairment	12	356	40	50
Tested at Home or Hospital	Specific Learning Disability	17	332	41	24
Transfer of Student T/B Responses to A/D	Autism	23	322	41	26
	Orthopedic Impairment	22	353	32	55
	Other Health Impairment	20	352	44	50
	Specific Learning Disability	58	340	27	40
	Visual Impairment	52	371	45	63
Unlisted Accommodation	Autism	31	332	44	35
	Emotional Disturbance	44	331	36	34
	Mental Retardation	17	306	34	12
	Other Health Impairment	67	338	31	39
	Specific Learning Disability	433	330	29	24
	Speech or Language Impairment	27	320	27	15
Unlisted Modification	Specific Learning Disability	18	337	35	39
Writing Only	Autism	15	331	34	27
	Deaf	69	296	19	3
	Emotional Disturbance	11	332	33	45
	Mental Retardation	12	295	19	0
	Other Health Impairment	21	329	31	29
	Specific Learning Disability	278	321	26	16

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.F.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, March 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Autism	19	335	24	32
	Emotional Disturbance	21	328	21	24
	Mental Retardation	11	318	14	0
	Orthopedic Impairment	12	336	14	25
	Other Health Impairment	47	332	22	15
	Specific Learning Disability	424	333	20	17
	Speech or Language Impairment	13	342	22	38
Beneficial Time	Autism	39	327	30	21
	Deaf	11	331	23	18
	Emotional Disturbance	220	326	29	15
	Mental Retardation	26	313	14	4
	Other Health Impairment	74	337	26	26
	Specific Learning Disability	452	334	26	25
	Speech or Language Impairment	23	330	26	22
Braille Version	Visual Impairment	20	343	38	35
Calculator	Autism	316	341	26	31
	Deaf	38	333	19	18
	Emotional Disturbance	454	331	24	20
	Hard of Hearing	69	338	23	26
	Mental Retardation	193	321	16	4
	Multiple Disability	14	318	17	0
	Orthopedic Impairment	76	331	20	16
	Other Health Impairment	782	338	24	28
	Specific Learning Disability	5,897	336	21	22
	Speech or Language Impairment	325	335	20	21
	Traumatic Brain Injury	39	334	26	21
	Visual Impairment	24	340	26	29
Dictionary for Math	Emotional Disturbance	17	324	26	24
	Other Health Impairment	15	343	19	40
	Specific Learning Disability	165	334	23	20
IEP or Section 504 Plan	Autism	1,354	356	39	47
	Deaf	223	338	31	27
	Emotional Disturbance	2,242	336	31	27
	Hard of Hearing	300	345	30	37
	Mental Retardation	522	316	16	3
	Multiple Disability	35	321	28	6
	Orthopedic Impairment	298	346	33	39
	Other Health Impairment	3,039	346	31	39
	Specific Learning Disability	20,158	338	25	26
	Speech or Language Impairment	1,046	342	29	32
	Traumatic Brain Injury	104	340	28	24
	Visual Impairment	156	362	40	58

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Large Print Version	Orthopedic Impairment	14	347	34	43
	Specific Learning Disability	17	340	28	24
	Visual Impairment	66	362	39	58
Math Manipulatives	Specific Learning Disability	47	327	16	6
Oral Presentation	Autism	113	337	25	20
	Emotional Disturbance	86	335	27	26
	Hard of Hearing	31	337	26	16
	Mental Retardation	100	321	17	6
	Orthopedic Impairment	31	333	25	23
	Other Health Impairment	187	337	25	23
	Specific Learning Disability	1,858	336	23	23
	Speech or Language Impairment	106	339	23	26
	Traumatic Brain Injury	16	330	19	6
	Visual Impairment	19	347	43	42
Oral Responses Dictated to a Scribe	Orthopedic Impairment	27	359	39	52
Sign Language	Deaf	134	333	23	21
	Hard of Hearing	29	322	17	7
Supervised Breaks	Autism	227	344	35	33
	Deaf	33	334	23	21
	Emotional Disturbance	517	331	31	23
	Hard of Hearing	52	335	26	23
	Mental Retardation	122	316	16	4
	Multiple Disability	11	320	18	9
	Orthopedic Impairment	41	342	32	37
	Other Health Impairment	414	342	29	35
	Specific Learning Disability	3,249	336	25	25
	Speech or Language Impairment	168	342	29	32
	Traumatic Brain Injury	27	338	27	26
	Visual Impairment	24	350	35	46
Test Over More Than One Day	Autism	63	351	35	46
	Emotional Disturbance	102	331	29	20
	Hard of Hearing	17	361	33	53
	Mental Retardation	27	315	15	4
	Orthopedic Impairment	19	353	30	68
	Other Health Impairment	136	351	35	46
	Specific Learning Disability	883	343	29	36
	Speech or Language Impairment	60	348	27	43
	Visual Impairment	12	357	55	50
Tested at Home or Hospital	Emotional Disturbance	11	341	37	18
	Specific Learning Disability	12	322	34	17
Transfer of Student T/B Responses to A/D	Autism	18	327	29	22
	Orthopedic Impairment	21	348	29	48

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Other Health Impairment	19	349	35	37
	Specific Learning Disability	55	345	31	35
	Visual Impairment	60	369	41	60
Unlisted Accommodation	Autism	28	345	39	39
	Emotional Disturbance	47	335	27	32
	Mental Retardation	16	323	29	19
	Other Health Impairment	59	345	26	37
	Specific Learning Disability	397	339	26	30
	Speech or Language Impairment	24	329	18	13

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.F.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, March 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	88	347	34	53
	Reclassified Fluent English Proficient	11	370	43	73
	English Learner	35	311	25	11
Beneficial Time	English Only	670	325	36	25
	Initially Fluent English Proficient	12	348	30	50
	Reclassified Fluent English Proficient	33	345	30	45
	English Learner	332	316	26	12
	Unknown	19	311	37	16
Dictionary	English Only	782	323	27	15
	Initially Fluent English Proficient	34	337	31	29
	Reclassified Fluent English Proficient	51	347	32	53
	English Learner	835	321	22	11
	Unknown	13	326	30	31
Essay Reponses(EO)	English Only	52	335	42	35
	English Learner	23	317	20	4
Essay Reponses(ER)	English Only	54	327	28	22
	English Learner	32	320	25	13
IEP or Section 504 Plan	English Only	18,449	336	36	35
	Initially Fluent English Proficient	797	340	34	38
	Reclassified Fluent English Proficient	1,807	351	31	56
	English Learner	10,865	321	25	14
	Unknown	266	338	41	40
Large Print Version	English Only	68	354	43	54
	Reclassified Fluent English Proficient	22	364	30	73
	English Learner	29	315	26	14
Oral Presentation	English Only	2,011	326	27	20
	Initially Fluent English Proficient	91	326	23	16
	Reclassified Fluent English Proficient	172	334	26	28
	English Learner	1,777	321	23	12
	Unknown	29	337	31	41
Oral Responses Dictated to a Scribe	English Only	43	339	44	37
	English Learner	19	326	36	21
	Unknown	11	312	40	18
Sign Language	English Only	33	310	18	0
	English Learner	49	301	23	6
Spell Checker or Grammar Checker	English Only	205	331	28	22
	Initially Fluent English Proficient	11	320	15	0
	English Learner	157	321	22	12
Spell Checker or Grammar Checker Off	English Only	187	363	40	62

Table 2.F.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Reclassified Fluent English Proficient	20	383	38	70
	English Learner	36	327	28	19
Supervised Breaks	English Only	3,185	330	34	29
	Initially Fluent English Proficient	143	334	32	31
	Reclassified Fluent English Proficient	352	344	30	45
	English Learner	2,048	319	25	13
	Unknown	64	328	35	28
Test Over More Than One Day	English Only	984	334	35	33
	Initially Fluent English Proficient	54	344	31	50
	Reclassified Fluent English Proficient	136	361	26	71
	English Learner	760	326	27	21
Tested at Home or Hospital	English Only	51	364	47	63
	English Learner	13	317	31	23
Transfer of Student T/B Responses to A/D	English Only	149	350	40	47
	Reclassified Fluent English Proficient	25	371	46	68
	English Learner	33	334	30	39
Unlisted Accommodation	English Only	484	335	35	36
	Initially Fluent English Proficient	15	328	32	27
	Reclassified Fluent English Proficient	36	345	27	39
	English Learner	194	321	23	12
	Unknown	12	370	35	83
Unlisted Modification	English Only	30	341	40	47
	English Learner	12	319	35	25
Writing Only	English Only	240	323	32	20
	Initially Fluent English Proficient	22	325	43	27
	Reclassified Fluent English Proficient	19	342	32	47
	English Learner	196	313	24	10

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.F.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, March 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	323	334	23	19
	Initially Fluent English Proficient	16	335	17	6
	Reclassified Fluent English Proficient	18	341	18	28
	English Learner	220	332	19	18
Beneficial Time	English Only	621	335	30	26
	Initially Fluent English Proficient	12	345	33	50
	Reclassified Fluent English Proficient	29	342	23	38
	English Learner	273	327	24	16
	Unknown	21	322	21	14
Calculator	English Only	4,956	336	23	24
	Initially Fluent English Proficient	220	339	22	23
	Reclassified Fluent English Proficient	385	347	24	42
	English Learner	3,069	332	19	17
	Unknown	81	337	29	31
Dictionary for Math	English Only	131	331	24	18
	English Learner	98	335	21	19
IEP or Section 504 Plan	English Only	18,861	343	31	34
	Initially Fluent English Proficient	827	347	31	37
	Reclassified Fluent English Proficient	1,864	357	32	53
	English Learner	10,239	333	22	19
	Unknown	293	345	37	39
Large Print Version	English Only	65	356	36	48
	Reclassified Fluent English Proficient	22	373	35	73
	English Learner	24	330	22	21
Math Manipulatives	English Only	53	340	31	26
	English Learner	18	328	14	6
Oral Presentation	English Only	1,385	337	24	24
	Initially Fluent English Proficient	52	336	18	25
	Reclassified Fluent English Proficient	110	345	27	37
	English Learner	1,100	333	22	18
	Unknown	30	347	31	47
Oral Responses Dictated to a Scribe	English Only	33	349	38	45
	English Learner	14	334	21	21
	Unknown	12	335	31	33
Sign Language	English Only	96	333	23	19
	English Learner	76	328	21	14
Supervised Breaks	English Only	2,950	338	29	28
	Initially Fluent English Proficient	136	345	28	38
	Reclassified Fluent English Proficient	331	350	29	45

Table 2.F.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	1,730	332	23	18
	Unknown	53	336	32	30
Test Over More Than One Day	English Only	681	342	31	36
	Initially Fluent English Proficient	41	349	33	41
	Reclassified Fluent English Proficient	104	373	32	73
	English Learner	562	338	26	30
Tested at Home or Hospital	English Only	46	361	47	54
Transfer of Student T/B Responses to A/D	English Only	142	353	40	44
	Reclassified Fluent English Proficient	23	376	39	65
	English Learner	30	346	23	33
Unlisted Accommodation	English Only	461	342	30	36
	Initially Fluent English Proficient	14	347	25	43
	Reclassified Fluent English Proficient	36	352	25	50
	English Learner	170	335	22	22
	Unknown	13	362	27	62
Unlisted Modification	English Only	16	358	36	63

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Appendix 2.G: Results of Testing Variations and Disability Analyses—May 2012

Table 2.G.1: Summary Statistics by Testing Variations and Disability—ELA, May 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,481	316	27	11
Modifications	2,478	322	24	13
All	3,959	320	25	12
Disability				
Unknown	-	-	-	-
Autism	322	321	28	13
Deaf	143	306	20	2
Deaf-Blindness	-	-	-	-
Emotional Disturbance	890	319	33	17
Hard of Hearing	110	318	24	11
Mental Retardation	294	301	18	1
Multiple Disability	27	315	29	19
Orthopedic Impairment	88	317	28	11
Other Health Impairment	866	324	29	15
Specific Learning Disability	8,048	318	23	10
Speech or Language Impairment	499	323	25	12
Traumatic Brain Injury	59	326	27	14
Visual Impairment	27	347	40	41

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.G.2: Summary Statistics by Testing Variations and Disability—Mathematics, May 2012

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,170	325	23	11
Modifications	3,574	330	21	15
All	4,744	329	22	14
Disability				
Unknown	-	-	-	-
Autism	291	330	23	14
Deaf	102	323	19	7
Deaf-Blindness	-	-	-	-
Emotional Disturbance	951	327	26	18
Hard of Hearing	88	329	22	11
Mental Retardation	283	312	16	2
Multiple Disability	24	322	21	4
Orthopedic Impairment	92	326	22	12
Other Health Impairment	935	331	23	18
Specific Learning Disability	7,489	327	20	12
Speech or Language Impairment	420	330	23	16
Traumatic Brain Injury	41	329	27	10
Visual Impairment	30	325	26	13

Note: Students who tested with the Testing Variations of Accommodations and Modifications are counted in both rows.

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.G.3: Demographic Summary for All Examinees by Testing Variations—ELA, May 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Reading[2] Mean Percent Correct			Writing[2] Mean Percent Correct		Writing Applications Mean Score[3]
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 Plan	8,756	954	11	7,802	89	319	52	43	47	38	40	1.8
Transfer of T/B Responses to A/D (TS)	47	13	28	34	72	332	63	51	53	43	46	2.1
Oral Responses Dictated to a Scribe (OR)	22	6	27	16	73	329	62	48	48	50	41	2.0
Spell Checker/Grammar Checker Off (SO)	70	19	27	51	73	335	61	54	54	48	49	2.2
Essay Reponses (EO)	55	11	20	44	80	322	53	43	49	40	39	2.0
Assistive Device No Interference (AN)	13	3	23	10	77	335	66	47	53	48	45	2.3
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	19	7	37	12	63	336	62	54	56	49	48	2.1
Test Over More Than One Day (TD)	358	67	19	291	81	325	55	47	51	43	42	1.9
Supervised Breaks (SB)	1,496	164	11	1,332	89	318	50	42	47	37	39	1.8
Beneficial Time (BT)	352	31	9	321	91	312	47	38	44	33	36	1.7
Tested at Home or Hospital (HH)	28	13	46	15	54	349	67	60	63	55	58	2.1
Dictionary (DI)	709	91	13	618	87	323	59	46	49	40	42	1.9
Sign Language (SL)	59	5	8	54	92	313	49	42	41	37	38	1.6
Oral Presentation (OP)	1,964	279	14	1,685	86	323	57	47	50	42	42	1.9
Spell Checker or Grammar Checker (SC)	205	30	15	175	85	324	55	46	48	40	41	2.1
Essay Reponses (ER)	47	8	17	39	83	321	56	42	44	41	37	2.2
Assistive Device (AD)	16	3	19	13	81	320	50	41	43	40	41	2.3
Unlisted Modification (UM)	40	7	18	33	83	320	54	43	51	41	37	1.8
Unlisted Accommodation (UA)	216	27	13	189	88	321	51	45	49	40	42	1.9
Writing Only (WO)	148	7	5	141	95	307	40	36	37	33	35	1.6

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Response & Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] Writing Applications Mean Score is based on the unweighted score.

Table 2.G.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, May 2012

Testing Variations	N Tested[1]	N (≥350)	Percent (≥350)	N (<350)	Percent (<350)	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
IEP or Section 504 Plan	8,478	1,124	13	7,354	87	327	42	42	38	36	28
Transfer of T/B Responses to A/D (TS)	55	19	35	36	65	340	53	54	47	44	29
Oral Responses Dictated to a Scribe (OR)	16	1	6	15	94	322	40	38	34	34	27
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	18	3	17	15	83	330	44	46	40	37	27
Test Over More Than One Day (TD)	228	49	21	179	79	333	48	46	42	40	30
Supervised Breaks (SB)	1,226	141	12	1,085	88	326	42	41	38	35	27
Beneficial Time (BT)	327	36	11	291	89	320	39	39	33	32	24
Tested at Home or Hospital (HH)	27	15	56	12	44	361	63	61	60	57	46
Dictionary for Math (DM)	85	21	25	64	75	335	47	50	43	41	32
Sign Language (SL)	37	6	16	31	84	328	36	40	44	38	31
Oral Presentation (OP)	1,078	181	17	897	83	331	44	46	40	39	29
Calculator (CA)	3,523	544	15	2,979	85	330	43	46	40	38	28
Arithmetic Table (AT)	271	33	12	238	88	327	42	44	38	37	26
Math Manipulatives (MM)	27	7	26	20	74	339	50	51	47	44	32
Assistive Device (AD)	13	2	15	11	85	330	37	50	43	37	26
Unlisted Modification (UM)	57	14	25	43	75	334	49	47	43	41	31
Unlisted Accommodation (UA)	225	35	16	190	84	328	44	42	38	36	28

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 2.G.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, May 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	281	300	315	336	364	392	319	26	8,756
Transfer of T/B Responses to A/D (TS)	275	289	304	330	351	389	450	332	36	47
Oral Responses Dictated to a Scribe (OR)	275	277	293	315	355	419	419	329	42	22
Spell Checker/Grammar Checker Off (SO)	275	293	311	332	351	419	450	335	34	70
Essay Reponses (EO)	275	285	300	311	337	385	419	322	32	55
Assistive Device No Interference (AN)	277	277	310	319	341	419	419	335	46	13
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	291	291	308	332	353	450	450	336	37	19
Test Over More Than One Day (TD)	275	279	304	324	345	373	395	325	29	358
Supervised Breaks (SB)	275	277	298	315	334	364	389	318	26	1,496
Beneficial Time (BT)	275	275	293	308	328	364	389	312	27	352
Tested at Home or Hospital (HH)	275	279	310	346	386	419	419	349	46	28
Dictionary (DI)	277	287	306	321	337	364	386	323	24	709
Sign Language (SL)	275	281	296	310	323	355	371	313	22	59
Oral Presentation (OP)	275	287	306	323	339	366	389	323	24	1,964
Spell Checker or Grammar Checker (SC)	275	285	306	323	339	366	375	324	24	205
Essay Reponses (ER)	283	287	300	311	341	368	392	321	27	47
Assistive Device (AD)	283	283	297	315	346	366	366	320	28	16
Unlisted Modification (UM)	279	281	298	315	334	369	419	320	30	40
Unlisted Accommodation (UA)	275	287	304	319	334	366	381	321	24	216
Writing Only (WO)	275	275	293	302	316	349	373	307	22	148

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.G.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, May 2012

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	282	297	312	324	340	365	396	327	22	8,478
Transfer of T/B Responses to A/D (TS)	275	294	318	338	365	405	412	340	33	55
Oral Responses Dictated to a Scribe (OR)	278	278	308	323	339	354	354	322	20	16
Braille Version (BV))	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV))	275	275	312	324	345	420	420	330	30	18
Test Over More Than One Day (TD)	275	297	314	329	347	386	412	333	27	228
Supervised Breaks (SB))	275	297	312	322	338	367	402	326	23	1,226
Beneficial Time (BT)	275	285	307	318	331	359	377	320	22	327
Tested at Home or Hospital (HH)	307	312	324	361	386	445	450	361	42	27
Dictionary for Math (DM)	275	302	320	333	349	379	388	335	22	85
Sign Language (SL)	297	297	314	324	342	375	377	328	21	37
Oral Presentation (OP)	285	300	314	328	342	375	408	331	23	1,078
Calculator (CA)	289	300	316	328	342	365	393	330	21	3,523
Arithmetic Table (AT)	285	300	314	326	340	359	377	327	19	271
Math Manipulatives (MM)	300	302	329	338	351	363	393	339	21	27
Assistive Device (AD)	307	307	320	328	338	358	358	330	16	13
Unlisted Modification (UM)	275	292	312	333	349	393	425	334	29	57
Unlisted Accommodation (UA)	289	300	314	322	338	367	416	328	23	225

[1] SD — Standard Deviation

[2] Results for groups with fewer than 11 students are not reported.

Table 2.G.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, May 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	Autism	18	325	38	11
	Emotional Disturbance	104	308	28	8
	Mental Retardation	16	294	15	0
	Other Health Impairment	18	315	30	11
	Specific Learning Disability	168	314	23	8
Dictionary	Autism	12	325	21	8
	Emotional Disturbance	30	318	32	20
	Mental Retardation	22	314	29	14
	Other Health Impairment	36	323	23	11
	Specific Learning Disability	495	324	23	12
	Speech or Language Impairment	29	311	19	3
Essay Reponses(EO)	Specific Learning Disability	27	318	27	11
Essay Reponses(ER)	Specific Learning Disability	19	314	19	5
IEP or Section 504 Plan	Autism	240	323	27	14
	Deaf	130	306	20	2
	Emotional Disturbance	686	319	33	17
	Hard of Hearing	83	318	25	12
	Mental Retardation	225	302	19	2
	Multiple Disability	24	316	28	17
	Orthopedic Impairment	64	315	24	9
	Other Health Impairment	643	323	28	14
	Specific Learning Disability	5,785	318	23	9
	Speech or Language Impairment	311	321	23	10
	Traumatic Brain Injury	29	317	25	7
	Visual Impairment	22	351	42	45
Oral Presentation	Autism	68	325	28	18
	Emotional Disturbance	59	323	31	22
	Hard of Hearing	16	326	25	25
	Mental Retardation	71	306	19	6
	Orthopedic Impairment	21	320	30	19
	Other Health Impairment	127	328	25	17
	Specific Learning Disability	1,387	323	23	13
	Speech or Language Impairment	98	322	24	14
Sign Language	Deaf	42	312	22	7
	Hard of Hearing	11	314	21	0
Spell Checker or Grammar Checker	Deaf	15	303	19	0
	Other Health Impairment	14	330	21	14
	Specific Learning Disability	132	325	23	14
	Speech or Language Impairment	15	316	21	7
Spell Checker or Grammar Checker Off	Specific Learning Disability	33	328	24	18
Supervised Breaks	Autism	49	323	30	10
	Deaf	13	302	20	0
	Emotional Disturbance	179	313	32	15

Table 2.G.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Hard of Hearing	11	319	26	9
	Mental Retardation	55	298	18	2
	Other Health Impairment	87	328	27	18
	Specific Learning Disability	972	318	24	10
	Speech or Language Impairment	40	317	20	10
Test Over More Than One Day	Autism	20	339	41	25
	Deaf	14	300	15	0
	Emotional Disturbance	51	313	31	12
	Other Health Impairment	20	324	29	20
	Specific Learning Disability	194	327	26	20
Transfer of Student T/B Responses to A/D	Visual Impairment	13	358	43	54
Unlisted Accommodation	Emotional Disturbance	26	324	31	27
	Other Health Impairment	19	338	24	26
	Specific Learning Disability	144	318	22	8
Unlisted Modification	Specific Learning Disability	27	317	25	15
Writing Only	Deaf	76	301	17	1
	Specific Learning Disability	41	313	23	5

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.G.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, May 2012

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Other Health Impairment	16	334	22	19
	Specific Learning Disability	201	327	18	11
	Speech or Language Impairment	11	333	22	27
Beneficial Time	Autism	13	332	25	31
	Emotional Disturbance	101	317	22	10
	Mental Retardation	13	310	22	8
	Other Health Impairment	23	323	19	9
	Specific Learning Disability	151	321	21	11
Calculator	Autism	110	331	20	15
	Deaf	18	329	20	11
	Emotional Disturbance	197	333	27	23
	Hard of Hearing	27	334	22	15
	Mental Retardation	86	315	16	5
	Multiple Disability	11	329	25	9
	Orthopedic Impairment	31	327	24	13
	Other Health Impairment	301	333	22	15
	Specific Learning Disability	2,400	329	20	15
	Speech or Language Impairment	133	329	22	17
	Traumatic Brain Injury	12	329	16	8
Dictionary for Math	Specific Learning Disability	58	335	20	22
IEP or Section 504 Plan	Autism	222	330	20	13
	Deaf	91	323	20	8
	Emotional Disturbance	750	328	26	18
	Hard of Hearing	65	328	22	12
	Mental Retardation	213	314	16	3
	Multiple Disability	24	322	21	4
	Orthopedic Impairment	63	325	20	10
	Other Health Impairment	724	331	23	17
	Specific Learning Disability	5,480	326	20	12
	Speech or Language Impairment	264	329	20	16
	Traumatic Brain Injury	25	324	16	4
	Visual Impairment	24	326	27	13
Math Manipulatives	Specific Learning Disability	22	336	18	23
Oral Presentation	Autism	32	326	16	6
	Emotional Disturbance	37	338	36	35
	Mental Retardation	41	317	18	5
	Orthopedic Impairment	12	314	18	0
	Other Health Impairment	82	334	25	21
	Specific Learning Disability	755	330	22	16
	Speech or Language Impairment	42	334	23	21
Sign Language	Deaf	26	328	20	15
Supervised Breaks	Autism	41	331	21	20
	Deaf	12	325	18	8
	Emotional Disturbance	167	324	29	15

Table 2.G.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Mental Retardation	46	314	22	9
	Other Health Impairment	92	332	24	18
	Specific Learning Disability	762	326	21	10
	Speech or Language Impairment	30	326	19	10
Test Over More Than One Day	Autism	15	333	21	20
	Emotional Disturbance	40	329	26	15
	Other Health Impairment	13	353	32	38
	Specific Learning Disability	132	331	27	21
Transfer of Student T/B Responses to A/D	Specific Learning Disability	24	347	35	46
Unlisted Accommodation	Emotional Disturbance	29	328	29	17
	Other Health Impairment	24	340	27	38
	Specific Learning Disability	143	325	21	10
Unlisted Modification	Specific Learning Disability	40	336	29	23

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.G.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, May 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	English Only	232	311	28	9
	English Learner	98	314	24	10
Dictionary	English Only	363	325	27	18
	Initially Fluent English Proficient	13	323	20	8
	Reclassified Fluent English Proficient	24	324	23	13
	English Learner	304	320	20	7
Essay Reponses(EO)	English Only	31	328	36	26
	English Learner	19	314	26	11
Essay Reponses(ER)	English Only	26	334	27	27
	English Learner	19	307	17	5
IEP or Section 504 Plan	English Only	4,988	321	28	14
	Initially Fluent English Proficient	180	322	25	9
	Reclassified Fluent English Proficient	266	324	26	12
	English Learner	3,193	315	21	6
	Unknown	129	322	25	11
Large Print Version	English Only	11	340	21	45
Oral Presentation	English Only	1,033	327	26	19
	Initially Fluent English Proficient	41	327	22	7
	Reclassified Fluent English Proficient	61	326	23	11
	English Learner	820	319	22	9
Oral Responses Dictated to a Scribe	English Only	14	330	42	36
Sign Language	English Only	35	317	24	14
	English Learner	20	304	15	0
Spell Checker or Grammar Checker	English Only	103	331	25	23
	English Learner	93	316	21	4
Spell Checker or Grammar Checker Off	English Only	38	342	35	39
	English Learner	27	319	18	4
Supervised Breaks	English Only	891	319	28	14
	Initially Fluent English Proficient	35	322	27	3
	Reclassified Fluent English Proficient	50	321	23	6
	English Learner	504	315	22	7
	Unknown	16	324	28	13
Test Over More Than One Day	English Only	220	329	32	26
	Reclassified Fluent English Proficient	19	321	25	11
	English Learner	97	318	23	7
	Unknown	20	313	16	0

Table 2.G.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Tested at Home or Hospital	English Only	23	351	48	48
Transfer of Student T/B Responses to A/D	English Only	32	330	30	28
Unlisted Accommodation	English Only	134	322	25	14
	English Learner	72	317	22	11
Unlisted Modification	English Only	23	319	21	13
	English Learner	14	314	34	21
Writing Only	English Only	105	307	24	7
	English Learner	32	308	16	0

[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.G.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, May 2012

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	167	330	20	16
	English Learner	94	322	17	4
Beneficial Time	English Only	225	320	21	11
	English Learner	80	317	20	6
Calculator	English Only	2,152	331	22	18
	Initially Fluent English Proficient	87	333	25	22
	Reclassified Fluent English Proficient	132	335	22	19
	English Learner	1,125	326	19	10
	Unknown	27	333	29	37
Dictionary for Math	English Only	63	339	22	30
	English Learner	20	324	23	10
IEP or Section 504 Plan	English Only	5,203	328	23	15
	Initially Fluent English Proficient	194	330	23	16
	Reclassified Fluent English Proficient	293	332	22	17
	English Learner	2,677	324	19	9
	Unknown	111	326	26	16
Large Print Version	English Only	12	337	32	25
Math Manipulatives	English Only	22	337	22	23
Oral Presentation	English Only	578	334	25	21
	Initially Fluent English Proficient	21	331	23	24
	Reclassified Fluent English Proficient	36	335	23	19
	English Learner	438	326	20	10
Oral Responses Dictated to a Scribe	English Only	12	323	22	8
Sign Language	English Only	25	326	21	16
Supervised Breaks	English Only	781	327	24	13
	Initially Fluent English Proficient	33	331	23	21
	Reclassified Fluent English Proficient	42	329	23	17
	English Learner	356	322	19	6
	Unknown	14	321	37	21
Test Over More Than One Day	English Only	154	336	29	27
	English Learner	52	322	18	6
	Unknown	12	328	31	17
Tested at Home or Hospital	English Only	20	364	45	60
Transfer of Student T/B Responses to A/D	English Only	40	343	28	38
	English Learner	11	326	45	18
Unlisted Accommodation	English Only	153	329	25	18
	English Learner	58	324	18	10
Unlisted Modification	English Only	43	337	31	30

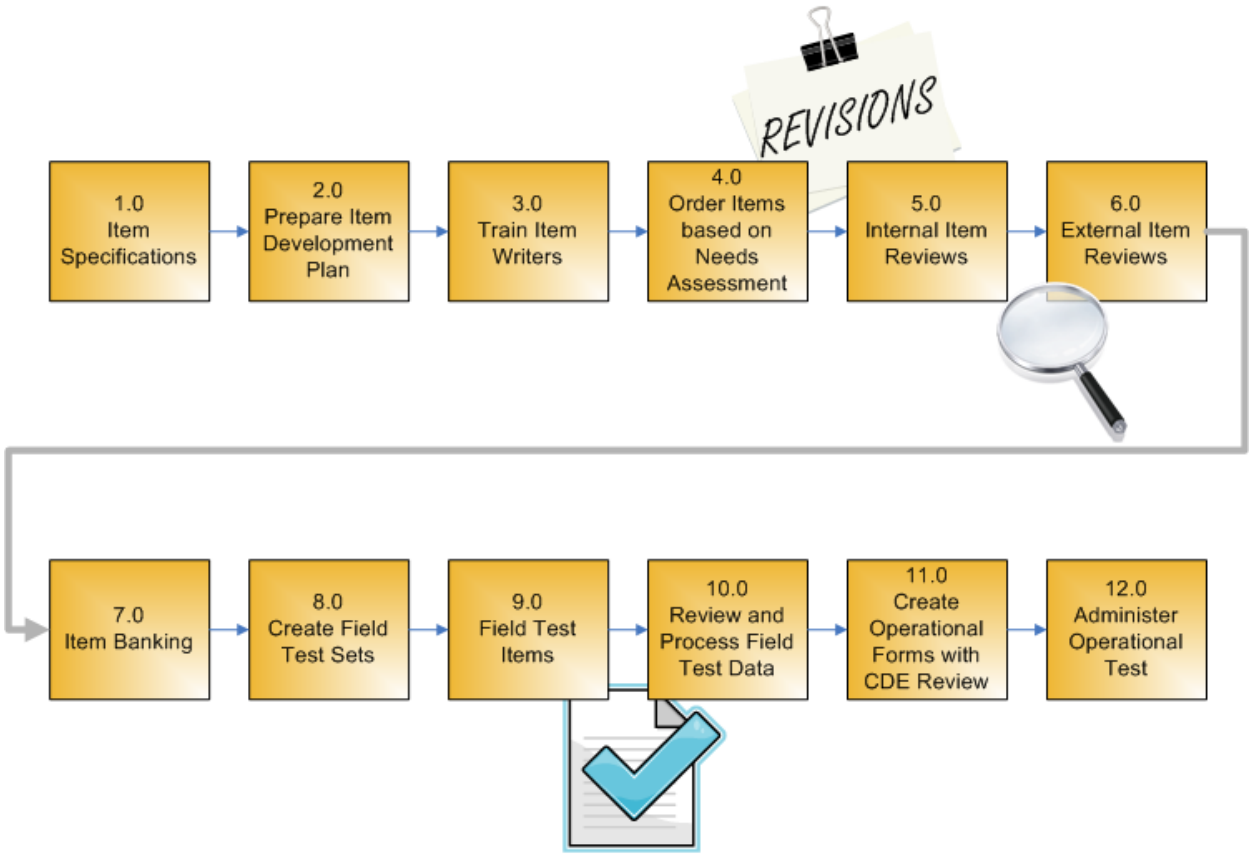
[1] Results for groups with fewer than 11 students are not reported.

[2] SD — Standard Deviation

Chapter 3: Item Development

The CAHSEE items are developed to measure California's content standards and are designed to conform to the principles of item writing defined by ETS (ETS, 2002). Each CAHSEE item goes through a comprehensive development cycle as described in Figure 3.1 below.

Figure 3.1 The ETS Item Development Process for the CAHSEE



1.0 Item Specifications

The first step in the item development process is to develop item specifications for ELA and mathematics. The item specifications describe the characteristics of the items that are needed to measure each content standard. The item specifications help ensure that the items on the CAHSEE measure the content standards as intended. To accomplish this, the item specifications provide detailed information to item writers who are developing items for the CAHSEE. The specifications include the following:

- A full statement of each academic content standard, as defined by the SBE (CDE, 2011c).
- A description of each content strand.
- The expected depth of knowledge (DOK), coded as 1, 2, 3, or 4.

  - DOK level 1 (low) – recall or simple reproduction of information
  - DOK level 2 (medium) – skills and concepts; comprehension and processing of text
  - DOK level 3 (high) – strategic thinking, prediction, elaboration
  - DOK level 4 (CR) – extended reasoning, complex analyses

- The homogeneity of the construct measured by each standard.

- A description of the kinds of item stems appropriate for MC items used to assess each standard.

- Sample items to serve as a guide for the types of MC items that are appropriate for assessing each standard.

- A description of appropriate stimuli (such as charts, tables, graphs, or other illustrations) for mathematics items.

- When applicable, the content limits for the standard (such as one or two variables, maximum place values of numbers) for mathematics items.

- When applicable, a description of appropriate reading passages for ELA items.

- When applicable, a description of specific types of items to be avoided (e.g., negatives should not appear in both the stem and the options for an item).

In addition, the ELA item specifications contain guidelines for passages used to assess reading comprehension and writing. These guidelines include the following:

- The acceptable ranges for passage length.

- The expected distribution of passages by genre.

- Guidelines for readability and cognitive load, using standards agreed to by the CDE and ETS.

- Expected use of illustrations.

- The target number of items that should follow each reading passage.

- Writing passages and reading passages to have a readability level appropriate to the tested grade level.

- A list of topics to be avoided.

- Diversity and Fairness of passages.

- Passages and items developed using Universal Design Principles.

2.0 Prepare Item Development Plan

Next, the test blueprint is used to determine item needs and to create an Item Development Plan. The Item Development Plan includes strategies for developing items that will permit coverage of all appropriate standards for both the ELA and mathematics tests. ETS test development staff uses this plan to determine the number of items to develop for each content area. Item writing emphasis for particular standards/content is determined in consultation with the CDE.

The Item Development Plan assumes that 70 percent of items on an operational ELA form and 70 percent of items on an operational mathematics form are refreshed each year; these items remain in the item bank for future use. Previously, a certain number of items were released to the public. However, due to the state budget cuts, no ELA or mathematics RTQs were released in the 2011–12 testing cycle.

The field-test percentages and item counts for the 2012 field tests are shown in Table 3.1. The percentages are based on the ratio of the number of items field tested to the number of unique operational items on all the test forms, plus the number of items in the linking sets. The number of ELA items field tested was 4.74 times or 474 percent of the number of unique items used to score students across the administrations. The number of mathematics items field tested was 3.74 times or 374 percent of the number of unique items used to score students across the administrations.

Table 3.1 Field-Test Percentages for the CAHSEE

Content Area	Percentage of Operational Form Field Tested	Number of Items Field Tested
English–Language Arts	≈ 474	1,743
Mathematics	≈ 374	1,560

The plan calls for larger numbers of items to be field tested for ELA than for mathematics because most ELA items are based on passages. An entire set of items associated with a passage may be eliminated if a few items in the set perform poorly, so the potential rate of item loss is greater in ELA than in mathematics.

3.0 Train Item Writers

The items selected for each operational form are written by special panels of item writers with expertise in the California content standards. Applicants for item writing training are screened by senior ETS content staff and approved by the CDE staff. Only those with strong content and teaching backgrounds are approved for inclusion in the training. Thus, the participants are particularly experienced in writing test questions to the standards measured by the CAHSEE. However, due to state budget cuts, the item writer training was not conducted in the 2011–12 school year.

All item writers meet the following minimum qualifications:

- Bachelor’s or Master’s degree in a specified content area being tested.

- Three or more years of general education teaching experience in the content areas of ELA or mathematics.
- Knowledge about the capabilities of the students taking these tests.
- Knowledge of and experience with California content standards in ELA and mathematics.

Before items can be written for the CAHSEE ELA test, passages must be written for both reading and language. Applicants for passage writing are screened by senior ETS content staff, and only those with strong writing skills are approved to write passages. It is only after the passages have gone through an extensive internal review process that the passages can be sent out to item writers. This internal review process includes: content review, fairness review, fact checking review, and editorial review.

4.0 Order Items Based on Needs Assessment

Content Leads use the Item Bank Inventory and the Item Development Plan to determine how many items are needed for each standard and order items from item writers based on these numbers. The item writers are given, at a minimum, the following materials to guide them in their work:

- *CAHSEE Guidelines for Item Writers*.
- ETS' *Guidelines for Bias and Sensitivity*.
- A copy of the test item specifications document for the CAHSEE, which includes the California content standards assessed on the CAHSEE, sample stems, sample items, and a checklist for item writers.
- An explanation of the DOK ratings.
- The Internet Web link to the previous RTQs.

Throughout the item writing process, ETS assessment specialists provide feedback to item writers.

5.0 Internal Item Reviews (Educational Testing Service)

Purpose

The items selected for the CAHSEE undergo an extensive item review process that is designed to provide the CDE with the best standards-based tests possible. This section summarizes the various reviews performed that ensure the quality of the CAHSEE items and test forms.

Once the items have been written by external item writers, a series of ETS internal reviews are conducted. The reviews help establish the criteria used to judge the quality of the item content and are designed to ensure that each item is measuring

what it is intended to measure. The internal reviews also examine the overall quality of the test items before they are prepared for presentation to the content review committees and the CDE. Because of the complexities involved in producing defensible items for high-stakes programs such as the CAHSEE, it is essential that many experienced individuals review each item before it is presented to the content review committees, the CDE, and the Statewide Pupil Assessment Review (SPAR) panels.

The Internal Review process for the CAHSEE includes the following stages:

1. Internal Content Review
2. Internal Bias and Sensitivity Review
3. Internal Editorial Review

Throughout this multi-step item review process, the lead content-area assessment specialists and development team members continually evaluate adherence to the rules for item development.

Internal Content Review

Test items and materials undergo three reviews by the content-area assessment specialists. These assessment specialists make sure that the test items and related materials are in compliance with ETS' written guidelines for clarity, style, accuracy, and appropriateness for California students, as well as in compliance with the approved item specifications. Assessment specialists review each item in terms of the following characteristics:

- Relevance of each item to the purpose of the test
- Match of each item to the item specifications, including DOK
- Adherence 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
- Calculator sensitivity in mathematics items

Each item is classified with a code for the standard it is intended to measure. The assessment specialists check all items against these classification codes, both to evaluate the correctness of the classification and to ensure that the task posed by the item is relevant to the outcome it is intended to measure. The internal content reviewers may accept the item and classification as written, suggest revisions, or recommend that the item be rejected.

Internal Bias and Sensitivity Review

ETS assessment specialists, who are specially trained to identify and eliminate questions that contain content or wording that could be construed as offensive to or

biased against members of specific ethnic, racial, learning-disabled, or gender groups, conduct the next level of review.

The review process promotes 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.
- 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.
- Ensuring the language of items is appropriate for EL students.

Internal Editorial Review

After the items are reviewed by the content-area assessment specialists and the bias and sensitivity reviewers, a group of specially trained, project-specific editors review each item in preparation for review by the CAHSEE content review committees and the CDE. The editors check 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.

6.0 External Item Reviews (California Educators and California Department of Education)

Purpose

The CAHSEE Content and Bias and Sensitivity Review committees review newly developed items for accuracy of item content, clarity of phrasing, item quality, and fairness. The review panels are provided with the opportunity to review newly developed items and to make recommendations for the use of items in embedded field tests. For the 2011–12 development cycle, Content and Bias and Sensitivity Review committees were convened.

The Content and Bias and Sensitivity Review committees are composed of current and former teachers, resource specialists, administrators, curriculum experts, and other education professionals. Current school staff members must meet minimum qualifications to serve on the panels, including:

- Three or more years of general teaching experience in middle school and high school and in the content areas of English–language arts and mathematics.

- Bachelor's degree or higher in a grade or subject area related to English–language arts or mathematics.

- Knowledge and experience with the California content standards in English–language arts or mathematics.

Every effort is made to ensure that Content and Bias and Sensitivity Review committees include representation of gender and of the geographic regions and ethnic groups in California. Efforts are also made to ensure representation by members with experience serving California's diverse special education population.

Committee members are recruited through an application process. Recommendations are solicited from districts, charter schools, and county offices of education in addition to the CDE staff. Applications are received and reviewed throughout the year. Committee applications are reviewed by ETS Assessment Directors, who confirm that the applicants' qualifications meet the specified criteria. Applications that meet the criteria are forwarded to the CDE for review and final approval. Upon approval, the applicant is notified that he or she has been selected to serve on the Content Review committee or the Bias and Sensitivity Review committee.

Bias and Sensitivity Review Meetings for Review of the CAHSEE Items

The ETS test development manager facilitates the Bias and Sensitivity Review meetings. The CDE consultants also attend. Each meeting begins with a brief training session on how to review items for bias and sensitivity. The training consists of the following steps:

- An overview of the purpose and scope of the CAHSEE.

- An overview of the CAHSEE test development process.

- An overview of the Universal Design Principles.

- An overview of guidelines for reviewing CAHSEE items for potential issues of bias and sensitivity.

Guidelines for reviewing items are provided by ETS and approved by the CDE. The guidelines include the following:

- A definition of item bias
- A discussion of common forms of bias and stereotypes in test items
- A definition of potential issues of sensitivity in test items
- A discussion of methods for detecting item bias
- The purpose of bias and sensitivity review
- Guiding questions
- Sample items with embedded bias and sensitivity issues

The committee is also trained on how to make recommendations for revising items.

Content Review Meetings for Review of CAHSEE Items

Content review committees serve in an advisory capacity and review the newly developed items for alignment to the California content standards. They determine if the items are:

- Measuring the California standards as appropriate for the CAHSEE testing population.
- Interesting and appropriate to students tested at a particular grade/course level.

The ETS content-area assessment specialists facilitate the CAHSEE content review meetings. The CDE consultants also attend. Each meeting begins with a brief training session conducted by ETS on how to review items. The training consists of the following topics:

- An overview of the purpose and scope of the CAHSEE
- An overview of the CAHSEE test development process
- An overview of guidelines for reviewing CAHSEE items
- An overview of Universal Design practices
- An analysis of CAHSEE item specifications

Guidelines for reviewing items are provided by ETS and approved by the CDE. The guidelines for reviewing items are summarized below.

Item Guidelines

Does the item:

- Measure the content standard?
- Match the test item specifications?
- Align with the construct being measured?
- Test worthwhile concepts or information?
- Include administrative directions in the stem or question that give the student a full sense of what the item is asking?
- Avoid unnecessary wordiness?
- Reflect content that is free from bias against any person or group?
- Match the DOK rating?

Stimulus Guidelines

Is the stimulus (if any) for the item:

- Required in order to answer the item?
- Likely to be interesting to students?
- Clearly and correctly labeled?
- Providing all the information needed to answer the item?

As the first step of the item review process, panel members review a set of items independently and record their individual comments. Next, the group discusses each item. The content-area assessment specialists facilitate the discussion and record all recommendations in a master item review binder.

ETS staff maintains the master binders showing committee recommendations. Any discrepancies are reconciled with the CDE consultants. The committee is also trained on how to make recommendations for revising items.

Statewide Pupil Assessment Review

The SPAR is responsible for reviewing and approving test items before they are used operationally or as field-test items. The SPAR examines the items for intrusiveness into students' personal lives such as student and family beliefs, morality, religion, or sexuality. The SPAR panel representatives ensure that the test items conform to the requirements of *EC* Section 60614. The CR writing tasks are also presented for review. At the SPAR panel meetings, items are presented in binders for review. If the SPAR panel rejects specific items and/or CR writing tasks, the items and/or tasks are replaced. For the SPAR panel meeting, the item development coordinator or an ETS content specialist, requested in advance by the CDE, attends the opening session and remains accessible to respond to any questions during the course of the meeting.

7.0 Item Banking

Once the Content Review and Bias and Sensitivity Review committees are finished, the items are placed in the item bank along with their corresponding review information. Items that are accepted are updated to a Field-Test Ready status; items that are rejected are updated to a Rejected Before Use status. ETS then delivers the items to the CDE by means of the CAHSEE electronic item bank. Subsequent updates to items are based on field-test and operational use of the items. However, only the latest version of the item is in the bank at any one time. Data from every administration in which the item was used are included. Item statistics from the census administrations (i.e., February and March) are used for test assembly.

After field-test or operational use, items may be rejected that do not meet statistical specifications; such items are updated with a status of Rejected for Statistical Reasons and remain unavailable in the bank. The research group at ETS evaluates each item for difficulty, discrimination, and conformance to the IRT Rasch model. Researchers also determine if each item functions similarly for various subgroups of interest by performing DIF analyses. Field-tested items are temporarily unavailable if any subgroup has C+/- DIF; these items are updated with a status of Operational

Ready – Needs DIF Review. Once items are reviewed by California educators at a Data Review meeting (refer to section 10.0) and are accepted as valid measures of the content standards, they will be made available and updated in the item bank with a status of Operational Ready.

Items that are released or pending release and awaiting the CDE's approval are updated with a Released or Scheduled for Release status and remain unavailable in the bank. CAHSEE items used operationally Rest or are unusable for two years plus one administration. As items appear on forms, they go into a Resting status and are unavailable until their Wake-up date.

All unavailable items are marked with an availability indicator of Unavailable, and reason for rejection as described above. Statuses and availability are updated programmatically as items are presented for review, accepted or rejected, placed on a form for field testing, presented for statistical review, used operationally, and released. All rejection and release indications are monitored and controlled through ETS' assessment development processes.

ETS currently provides and maintains the electronic item banks for several of the California assessments, including the California Standards Tests (CST), California Modified Assessment (CMA), California Alternate Performance Assessment (CAPA), Standards-based Tests in Spanish (STS), and CAHSEE. The CST, CMA, CAPA, STS, and CAHSEE item banks are currently consolidated in the California Item Banking system. ETS works with the CDE to obtain the data for those assessments under contract with other vendors for inclusion in the item bank, using the tools previously developed. ETS provides the item banking application using the LAN architecture and the relational database management system, SQL 2000, already deployed. ETS provides updated versions of the item bank to the CDE on an ongoing basis and works with the CDE to determine the optimum process if a change in databases is desired.

8.0 Create Field-Test Sets

The primary purpose of field testing is to obtain information about item performance and to obtain statistics that can be used to assemble operational forms. Two types of field tests are conducted for the CAHSEE. A stand-alone field test is used for the ELA writing prompts. An embedded field test is used for both the ELA and mathematics MC items.

Stand-Alone Field Testing

Continual development and field testing of writing prompt items are essential to maintaining a robust item bank. Due to the time required to complete the writing prompt items, separate testing is conducted. The purpose of the ELA field test is to try out a large number of writing prompts to augment the CAHSEE item bank. The results of the field-test analyses are not provided to students.

Samples of grade 10 students were selected to ensure that all geographic areas of the state were represented and that districts of various sizes, including charter schools, were included in the sample. Twenty-eight forms were spiraled at the student

level to ensure that each form was administered to randomly equivalent groups. Every field-test form contained one unique writing prompt and 13 MC items, which were used as anchor items. Each form was designed to be administered within a single classroom period. A concurrent calibration was performed to place all items onto a common scale of measurement. The MC items were then used to link the writing prompt items onto the operational scale.

Embedded Field-Test Items

Although a stand-alone field test is useful for developing a new test because it can produce a large pool of quality items, embedded field testing is generally preferred because the field-test items are dispersed throughout the operational test. Variables such as test-taker motivation and test security are the same for embedded field testing as when the items are later administered operationally. Such field testing involves distributing the field-test items within each operational test form. Different forms contain the same operational items and different field-test items. Only field-test data from the February and March administrations are used to evaluate the item performance.

Allocation of Students to Field-Test Items

The operational test forms for the CAHSEE are spiraled among students in the state so that a large representative sample of test takers responds to the field-test items embedded in these forms. The spiraling design ensures that a diverse sample of students takes each field-test item. The students do not know which items are field-test items and which items are operational items; therefore, their motivation is not expected to vary between the two types of items (Patrick & Way, 2008).

9.0 Field-Test Items

Newly field-tested items are analyzed to obtain statistical information about item performance. A set of seven field-test items is administered on the ELA forms and a set of 12 field-test items is administered on the mathematics forms. The sets of field-test items differ across forms, and the number of forms varies across administrations.

Table 3.2 shows the number of forms administered during the CAHSEE census administrations in 2012 and the numbers of examinees included in samples used for the field test or Final Item Analyses (FIA) of these forms. The field-test samples constitute the entire population tested except students who did not have a valid form number, did not complete the test, or did not attempt at least the first five questions of the test. One hundred and four ELA field-test forms were administered in February and 155 ELA field-test forms were administered in March. Fifty-two mathematics field-test forms were administered in February and 78 field-test forms were administered in March.

Table 3.2 Summary of Items and Forms Presented in 2012 CAHSEE

		Operational			Field Test			
Subject	**Administration**	**No. of Items**	**No. of Examinees FIA Sample**	**No. of Examinees Total[1]**	**No. of Forms**	**No. of Items**	**No. of Examinees FIA Sample[2]**	**No. of Examinees Total**
ELA	**February**	73	157,508	156,446	104	694	1,437-4,470	156,446
	March	73	388,059	386,472	155	1049	2,145-6,570	386,472
Mathematics	**February**	80	154,011	153,429	52	624	2,835-3,705	153,429
	March	80	382,980	381,777	78	936	4,771-5,962	381,777

[1] The Number of Examinees Total is the total number of examinees for score reporting purposes, and due to different inclusion criteria for score reporting than for FIA, it is smaller than the number of examinees in the FIA sample. See note under Table E.1 for criteria.

[2] The number of examinees for the ELA March FIA sample varied per form.

10.0 Review and Process Field-Test Data

Internal Review

Newly field-tested items are analyzed to obtain statistical information on student performance. This information is used to evaluate items that are candidates for operational test forms. These criteria ensure that the item (1) has an appropriate level of difficulty for the target population; (2) discriminates well between examinees that differ in ability; and (3) contains distractors that are functioning properly. Items with statistics that fall outside of the target range are flagged. These items undergo review by the ETS content specialists to verify that there is one correct key, all distractors are incorrect, and the item appeared in the test booklet as expected.

Differential Item Functioning Review

A Data Review meeting is held to review the field-test items that were flagged for DIF. The items and their associated statistics are prepared and presented to the Data Review committee, which is composed of members from the Content Review and Bias and Sensitivity Review committees and those who volunteered for one or both committees. ETS assessment specialists facilitate the Data Review sessions with qualified psychometric staff for technical assistance.

Committee members review and discuss the items that have been flagged for C-level DIF. Some of the items have also been flagged for poor statistics and do not meet the psychometric criteria for high item quality. The CDE has defined the criteria for acceptable or unacceptable item statistics. These criteria ensure that the item (1) has

an appropriate level of difficulty for the target population, (2) discriminates well between examinees that differ in ability, and (3) conforms well to the statistical model underlying the measurement of the intended constructs. The panel members also use the DIF results to make judgments about the appropriateness of items for various subgroups. The panelists respond to questions such as:

- Is there a content problem within the item?
- Are there any instructional issues that have negatively affected the performance of students on the item?

The panelists recommend whether to accept or reject each item for inclusion in the CAHSEE item bank. At the completion of the meeting, ETS provides the CDE with summaries of the recommendations based on the field-test analyses and the content reviews that are relevant to future form construction of the CAHSEE. All final decisions on acceptance of items rest with the CDE. Typically, a Data Review meeting is held once a year, although no meeting was held during the 2011–12 school year.

11.0 Create Operational Forms with California Department of Education Review

Forms are created by ETS using the California Item Bank, based on the CAHSEE blueprints and statistical requirements. Forms are reviewed internally by the content specialist, the psychometrician, and editorial staff before being posted to the CDE for their review. These forms are posted to and reviewed by the CDE three times and include additional ETS content, psychometric, and editorial checks during each step.

12.0 Administer Operational Test

The CAHSEE test is administered seven times per year. The months of administration are July, October, November, December, February, March, and May. The census administrations, consisting primarily of grade ten students, are February and March. The other administrations permit testing opportunities for those students who have not yet passed the examination. The July administration is only for grade twelve and adult education students.

Chapter 4: Test Development

The CAHSEE test forms are constructed to measure student performance relative to California's content standards approved by the SBE. They are also constructed to meet professional standards for validity and reliability. Each CAHSEE test form consists of operational and field-test items. Operational items are used to produce student scores. The field-test items are scored along with the operational items, but students' scores on these items are not included in the computation of a total test score. Instead, student performance on the field-test items from the census administrations is analyzed, and the calibrated items are placed in the item bank. The test development process described here refers to the process used to assemble items into operational test forms.

For each CAHSEE test, the content standards and psychometric attributes are used as the basis for assembling the test forms. The match of proposed forms to the specified psychometric criteria is evaluated using estimates based on the most recent item statistics obtained from embedded field testing conducted during the census administrations or previous operational administrations. The test construction process is completed using the CAHSEE item bank.

Test Length

The selection of items in each CAHSEE test form is decided by considering the construct that the test is intended to measure and the level of psychometric quality desired. Test length is closely related to the complexity of content to be measured by each test; this content is defined by California's content standards for each content area. Also considered is the goal that the tests be short enough that most of the students complete the test in a reasonable amount of time.

Each ELA form consists of 80 items, which include 72 operational MC items, one operational CR item, and seven embedded field-test MC items. Each mathematics form consists of 92 items, which include 80 operational MC items and 12 field-test MC items. For more details on the distribution of items across seven 2011–12 CAHSEE administrations, see Table 2.1 in Chapter 2.

Rules for Item Selection

Test Blueprint

ETS developed all CAHSEE test items to conform to the SBE-approved content standards and CAHSEE blueprints. The test blueprints for the CAHSEE can be found on the CDE CAHSEE Program Resources Web page at http://www.cde.ca.gov/ta/tg/hs/resources.asp.

Although the CAHSEE blueprints indicate the number of items at the individual standard level, scores on the CAHSEE items are grouped into subscore reporting categories referred to as content strands. For each CAHSEE content strand, the number of questions correctly answered and the percentage of the total number of

items in the strand are reported on a student's score report. For ELA, a CR item score is also given. A summary of the strand scores reported for the CAHSEE is provided in Table 2.2 in Chapter 2.

Content Rules and Item Selection

When developing a new test for a given examination, test developers follow a number of rules. First and foremost, they select items that meet the CAHSEE blueprint for that content area. Using an electronic item bank, assessment specialists begin by identifying a number of linking items. These are operational items that appeared in the previous year's census administrations and are used to equate the test forms administered in the subsequent year. Approximately one-half of the anchor items are selected from each of the February and March administrations in order to minimize item exposure for students retaking the test in another testing cycle. Linking items are selected to proportionally represent the full blueprint. For example, if 25 percent of all of the items in a test are in the Algebra and Functions strand, then approximately 25 percent of the linking items are targeted for inclusion from that strand. The linking items are selected for their content quality and are reviewed to ensure that they meet specified psychometric criteria as described below.

After the linking items are approved by the psychometricians and the CDE, assessment specialists select the remainder of the test form. Their first consideration is the strength of the content and the match of each item to the standard. In selecting items, test developers ensure that a variety of formats and content is included. Some items should include graphics that are visually interesting.

One psychometric consideration is the difficulty of each item. Test developers strive to ensure that there are a variety of easy, medium, and difficult items on the test. If items do not meet all content and psychometric criteria, test developers review alternate selections that could improve the match of the test to the requirements. If such a match is not attainable, the test developers work in conjunction with psychometricians and the CDE to determine which combination of items will be the best match possible within the constraints of the available item pool.

Psychometric Criteria

The CAHSEE test developers and psychometricians strive to accomplish three goals while developing a test:

1. The test must have the desired precision of measurement at all ability levels.

2. The test scores must be valid and reliable for the intended population and for the various subgroups of test takers.

3. The test forms must be comparable across administrations to ensure that scores generalize over time.

In order to achieve these goals, a set of rules has been developed that outlines the desired psychometric properties of each content area. Such rules are referred to as

statistical targets, which are provided to test developers before a test construction cycle begins. Tables 4.1 to 4.4 list the recommended statistical specifications for the CAHSEE test assembly, articulated in terms of equated item Rasch difficulty values (b-values) and item biserial correlations (r-bis) for the total test. The item b-values are based on the IRT Rasch model. When using the IRT Rasch model, the b-value targets make it possible to choose items to produce a test that has the desired precision of measurement at all ability levels, and test forms are comparable across administrations. The biserial correlation is a measure of how well the items discriminate among test takers that differ in their ability, and it is related to the overall reliability of the test.

In general, test developers are asked to match the statistical characteristics as closely as possible to the statistical targets. Tables 4.1 and 4.3 provide the target distributions of MC items in each of the intervals. They are used as guidelines by the test developers to match the mean equated Rasch difficulty for each form. Tables 4.2 and 4.4 provide the statistical guidelines for content strands in each of the test forms.

Figure 4.1 is an example of a test characteristic curve (TCC) which falls within the specifications. The TCC depicts the relationship between students' abilities and their expected true scores, expressed in the raw score metric. The dotted line shows that the TCC for the new test being constructed matches the average difficulty of the base form (depicted by the solid line) throughout the range of ability.

In addition to item difficulty and discrimination specifications, information about model-data fit and DIF is taken into account during test assembly (see Chapter 6 for a description of the procedures used for evaluating model-data fit). Test developers are instructed to avoid items with fit classifications of F, as well as items that have been flagged for severe (C-/C+) DIF unless they are approved by the CDE content specialists and DIF review committees for use.

Once constructed, the forms are reviewed and approved first by ETS psychometricians and then by the CDE. If any items are replaced, test developers resend the forms to the psychometricians for approval.

Figure 4.1: Example of Test Characteristic Curves

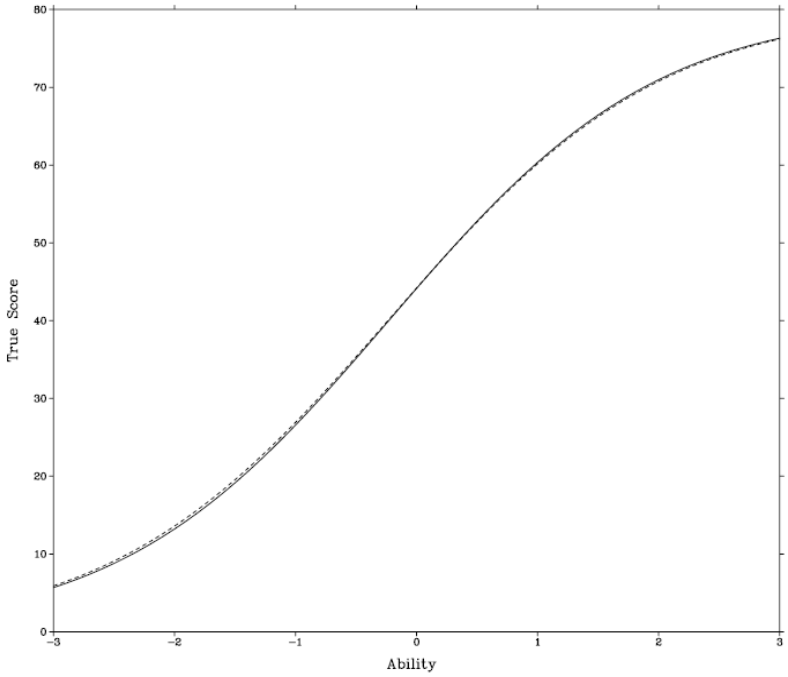


Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items

Low B	High B	Specifications
-1.75	-1.50	0 - 1
-1.50	-1.25	1 - 2
-1.25	-1.00	2 - 4
-1.00	-0.75	4 - 6
-0.75	-0.50	7 - 9
-0.50	-0.25	9 - 13
-0.25	0.00	10 - 14
0.00	0.25	9 - 13
0.25	0.50	7 - 12
0.50	0.75	7 - 10
0.75	1.00	2 - 5
1.00	1.25	2 - 5
1.25	1.50	1 - 3
1.50	1.75	0 - 2
	No. MC Items	**72**
	MC Mean	**-0.10 - 0.10**
	MC SD	**0.55 - 0.70**
Low R-bis	**High R-bis**	**Specifications**
0.0	0.1	0 - 0
0.1	0.2	0 - 0
0.2	0.3	4 - 10
0.3	0.4	16 - 24
0.4	0.5	16 - 24
0.5	0.6	16 - 24
0.6	0.7	7 - 10
0.7	0.8	1 - 3
	No. MC Items	**72**
	MC Mean	**0.44 - 0.54**
	MC SD	**0.10 - 0.15**

Table 4.2: Number of Items and Mean B-Value Ranges by Content Strand for ELA

Content Class	No. Items	Mean B Range
Reading Comprehension	18	-0.05 - 0.20
Literary Response & Analysis	20	-0.20 - 0.00
Word Analysis	7	-0.40 - 0.15
Writing Conventions	15	-0.03 - 0.25
Writing Strategies	12	0.07 - 0.50
Total	**72**	

Table 4.3: Difficulty (B) and Discrimination (R-bis) Specifications for Mathematics

Low B	High B	Specifications
-1.75	-1.50	0 - 1
-1.50	-1.25	1 - 2
-1.25	-1.00	2 - 4
-1.00	-0.75	7 - 10
-0.75	-0.50	7 - 10
-0.50	-0.25	7 - 10
-0.25	0.00	9 - 13
0.00	0.25	9 - 13
0.25	0.50	7 - 10
0.50	0.75	7 - 10
0.75	1.00	2 - 5
1.00	1.25	2 - 4
1.25	1.50	0 - 1
	No. MC Items	**80**
	MC Mean	**-0.30 - -0.20**
	MC SD	**0.65 - 0.80**
Low R-bis	**High R-bis**	**Specifications**
0.0	0.1	0 - 0
0.1	0.2	0 - 0
0.2	0.3	0 - 1
0.3	0.4	10 - 13
0.4	0.5	13 - 16
0.5	0.6	25 - 30
0.6	0.7	20 - 24
0.7	0.8	2 - 4
	No. MC Items	**80**
	MC Mean	**0.44 - 0.54**
	MC SD	**0.10 - 0.15**

Table 4.4: Number of Items and Mean B-Value Ranges by Content Strand for Mathematics

Content Class	No. Items	Mean B Range
Probability and Statistics	13	-0.8 - -0.4
Number Sense	17	-0.7 - -0.3
Algebra and Functions	20	-0.7 - -0.3
Measurement & Geometry	18	-0.4 - 0.0
Algebra I	12	0.0 - 0.4
Total	**80**	

Rules for Item Sequence and Layout

The items on test forms are organized and sequenced differently according to the requirements of the content area.

- **English-Language Arts**: Since the ELA test is primarily passage-dependent, items are sequenced according to their associated reading passages. Passages are sequenced according to genre and interest level. Test developers alternate potentially higher-interest pieces (typically narrative selections) with lower-interest pieces (typically functional or technical writing) to help alleviate reader fatigue. ELA items are divided into two sessions in the following order: first session—reading passages with their associated items followed by the writing CR item; second session—reading passages and writing passages with their associated items and writing stand-alone MC items.
- **Mathematics**: The mathematics test is sequenced according to reporting categories; that is, all items from a single reporting category are presented together, then all of the items from the next reporting category are presented. This ordering permits students to concentrate on one reporting category at a time. The reporting categories are organized in the following order: Number Sense, Probability and Statistics, Algebra and Functions, Measurement and Geometry, and Algebra I. Mathematical reasoning items are interspersed among the Number Sense, Probability and Statistics, Algebra and Functions, and Measurement and Geometry sections of the test. Mathematical reasoning is part of the *Mathematics Framework for California Public Schools* (CDE, 2006) and therefore, the blueprints. “It characterizes the thinking skills that students can carry from doing mathematics into other disciplines. Constructing valid arguments and criticizing invalid ones are inherent in doing mathematics” (CDE, 2006, p. xvi). Mathematical reasoning items are not scored as a separate reporting cluster but are scored under the associated content reporting cluster.

Chapter 5: Test Administration

Test Security and Confidentiality

All tests within the CAHSEE program are secure documents. For the CAHSEE administrations, every person having access to test materials maintains the security and confidentiality of the tests. ETS' Code of Ethics requires that all test information, including tangible materials (such as test booklets), confidential files, processes, and activities are kept secure. ETS has systems in place that maintain tight security for test questions and test results, as well as for student data. To ensure security for all the tests that ETS develops or handles, ETS maintains the OTI, which is described in the next section.

Educational Testing Service Office of Testing Integrity

The OTI is part of the ETS legal department and is a division that oversees test security standards for all testing programs administered by ETS. The Quality Assurance division, also within the legal department, publishes and maintains *ETS Standards for Quality and Fairness* and supports the OTI's goals and activities. The purposes of the *ETS Standards for Quality and Fairness* are to help ETS design, develop, and deliver technically sound, fair, and useful products and services and to help the public and auditors evaluate those products and services.

The OTI's mission is to

- Minimize any testing security violations that can impact the fairness of testing
- Minimize and investigate any security breach
- Report on security activities

The OTI helps prevent misconduct on the part of test takers and administrators, detects potential misconduct through empirically established indicators, and resolves situations in a fair and balanced way that reflects the laws and professional standards governing the integrity of testing. In their pursuit of enforcing secure practices, ETS and the OTI strive to safeguard the various processes involved in a test development and administration cycle. These practices are discussed in detail in the following sections.

Test Development

During the test development process, ETS staff members consistently adhere to the following established security procedures:

- Only authorized individuals have access to test content at any step in the test development, item review, and data analysis processes.

- Test developers keep all hard copy test content, computer disk copies, art, film, proofs, and plates in locked storage when not in use.

- ETS shreds working copies of secure content as soon as they are no longer needed in the test development process.

- Test developers take further security measures when test materials are to be shared outside of ETS; this is achieved by using registered and/or secure mail, using express delivery methods, and actively tracking records of dispatch and receipt of the materials.

Item Review

ETS enforces security measures at item review panel meetings to protect the integrity of meeting materials using the following protocols:

- Individuals who participate in the review panels must sign a confidentiality agreement.

- Meeting materials are strictly managed before, during, and after the review meetings.

- Meeting participants are supervised at all times during the meetings.

- Meeting participants are required to use cell phones outside of the room and not to have cell phones at the meeting tables.

Item Bank

Once the item review panel completes its review, the items are placed in the item bank along with their statistics. ETS then delivers the items to the CDE through the CAHSEE electronic item bank. Subsequent updates to content and statistics associated with items are based on data collected from field testing and the operational use of the items. The latest version of the item is retained in the bank along with the data from every administration that has included the item.

Security of the electronic item banking system is of critical importance. The measures for assuring the security of electronic files include the following:

- Electronic forms of test content, documentation, and item banks are backed up, and the backups are kept offsite.

- The offsite backup files are kept in secure storage with access limited to only authorized personnel.

- To prevent unauthorized electronic access to the item bank, state-of-the-art network security measures are used.

ETS routinely maintains many secure electronic systems for both internal and external access. The current electronic item banking application includes a login/password system to authorize access to the database or designated portions of the database. In addition, only users authorized to access the specific SQL database are able to use

the electronic item banking system. Designated administrators at the CDE and at ETS are authorized users.

Transfer of Forms and Items to the California Department of Education

ETS shares a secure file transfer protocol (SFTP) site with the CDE. SFTP is a method for reliable and exclusive routing of files. Files reside on a password-protected server that only authorized users may access. On that site, ETS posts Microsoft Word and Excel documents, Adobe Acrobat PDFs, or other document files for the CDE to review. ETS sends a notification e-mail to the CDE to announce that files are posted. Item data are always transmitted in an encrypted format to the SFTP site; test data are never sent via e-mail. The SFTP server is used as a conduit for the transfer of files; secure test data are not stored permanently on the shared SFTP server.

Printing

After items and test forms are approved, the files are sent for printing on a CD using a secure courier system. According to established procedures, the OTI pre-approves all printing vendors before they can work on secured confidential and proprietary testing material. The printing vendor must submit a completed ETS Printing Plan and a Typesetting Facility Security Plan; both plans document security procedures, access to testing materials, a log of work in progress, personnel procedures, and access to the facilities by the employees and visitors. After reviewing the completed plans, representatives of the OTI visit the printing vendor to conduct an onsite inspection. The printing vendor ships printed test booklets to Pearson Educational Measurement (Pearson) in Iowa City.

Test Administration

Pearson receives testing materials from printers, prints a unique barcode identifier on each test booklet, packages them, and ships them to school districts in triple-walled boxes. Pearson ships secure and non-secure test materials in separate shipments, both via secure, expedited delivery. Materials must be signed for and inventoried when they arrive at the district. After testing, the school districts return materials to Pearson for scanning. During these processes, Pearson takes extraordinary measures to protect the testing materials. Pearson's customized Oracle business applications verify that inventory controls are in place from materials receipt to packaging. The reputable carriers used by Pearson provide a specialized handling and delivery service that maintains test security and meets the CAHSEE program schedule. The carriers provide inside delivery directly to the CAHSEE district coordinators or authorized recipients of the assessment materials.

Test Delivery

Test security requires accounting for all secure materials before, during, and after each test administration. All booklets (including those for stand-alone field tests) are printed with unique readable bar code numbers. Packing lists in each box of secure materials contain the bar code range of booklets within that shipment. Any missing

materials are to be reported immediately to ETS. The same process is required when materials are distributed from the district's main location to the individual test sites. The CAHSEE district coordinators are required to keep all test materials in central locked storage except during actual test administration times. After personnel inventory all test booklets at the district and local sites, the boxes in which the test booklets were shipped are resealed and not opened until the day of the examination. Test site coordinators are responsible for accounting for and returning all secure materials to the CAHSEE district coordinator, who is responsible for returning them to the CAHSEE Processing Center at Pearson. All test booklets, upon return receipt at Pearson's receiving facility, are scanned and accounted for. If booklets are missing, district test coordinators are contacted and asked to recover and return the booklets. A final report of missing booklets is forwarded to the CAHSEE Office at the CDE.

The following measures are in place to ensure security of the CAHSEE testing materials:

- The CAHSEE district coordinators are required to sign and submit a *CAHSEE Test Security Agreement* form to the CAHSEE Technical Assistance Center before ETS can ship any testing materials to the district.

- Test site coordinators have to sign and submit a *CAHSEE Test Security Agreement* form to the CAHSEE district coordinator before any testing materials can be delivered to the school/test site.

- Anyone having access to the test materials must sign and submit a *CAHSEE Test Security Affidavit* form to the test site coordinator before receiving access to any testing materials.

- All testing is conducted on specific dates as determined by the SSPI. To ensure security throughout the state, test sessions must begin between the hours of 8 and 10 a.m. Test sites that must test outside this window must advise the CDE of their planned testing schedule.

- Test security during testing is managed by the CAHSEE district coordinator as well as the CAHSEE test site coordinator. All students are seated facing the same direction and with spacing at least four feet from center of desk to center of desk. Any information on bulletin boards, chalkboards, whiteboards, or charts that could be used by students to help answer questions on the test is removed or covered.

- If a security breach occurs at a test site before, during, or after a CAHSEE administration and is determined to be a violation of the *Test Security Agreement*, the district coordinator records all pertinent information on the *Test Security Breach Report* supplied in the *CAHSEE DTSCM* (CDE, 2011a) and faxes the form immediately to ETS. ETS, in consultation with the CDE CAHSEE offices, then launches an investigation.

- Any irregularities in test security may result in the invalidation of student test results.

- Procedures for test site security are documented in the *DTSCM* (CDE, 2011a) and are reviewed annually with the CDE CAHSEE Office.

Processing Security

Pearson has established the following security safeguards at their sites:

- Access to the facility is controlled.
- No test materials may leave the facility during the project.
- All staff must wear Pearson identification badges at all times in Pearson facilities.
- No recording or photographic equipment is allowed in the processing area.

The scanned answer documents are stored in secure warehouses. After they are stored they will not be handled again unless questions arise about a student's score. For example, a school district may request that a student's test responses be rescored. In such a case, the answer document is removed from storage, copied, and sent securely to ETS for hand scoring. Afterwards, the copy is destroyed. School and district personnel are not allowed to look at a completed answer document unless required for transcription or to investigate irregular cases.

All answer documents are securely destroyed one year after each examination administration, and test booklets are destroyed 90 days after each administration.

Data Management

Data, electronic files, test files, programs (source and object), and all associated tables and parameters are maintained in secure network libraries for all systems developed and maintained in a client-server environment. Only authorized software development employees are given access as needed for development, testing, and implementation in a strictly controlled environment.

Transfer of Files via Secure Data Exchange

ETS maintains a SFTP site to transmit secure data (test items, test forms, detail files, aggregate files, and other data files as needed) to the CDE and/or other third parties such as the independent evaluator. ETS also uses an SFTP site to share files between ETS and Pearson. Files posted to SFTP sites are protected using 128-bit encryption. Access to the CAHSEE data files is limited to appropriate personnel with direct project responsibilities.

Scoring and Analysis

After quality assurance procedures have been completed, Pearson transmits electronic files containing all information captured from the answer documents to ETS for scoring and analysis. MC items are scored in the Score Key Management (SKM)

system. Images of student essays are uploaded to the *Online Scoring Network*™ (OSN) where they are scored and results downloaded to the SKM system and then merged with the students' MC scores in the CAHSEE Management System.

Extracts are generated from the database, and the Data Quality Services (DQS) area performs quality control procedures before passing files to the Statistical Analysis group. The Statistical Analysis group maintains the files on secure servers and adheres to the ETS Code of Ethics to prevent any unauthorized access.

Reporting and Posting Results

After statistical analysis and quality control have been completed on student data, the data are reported in several ways. Student demographic detail files, which include student exam results, and district and school rosters and summary reports are posted in secure, password-protected district folders on the CAHSEE Web site. Individual student score reports are printed on security paper and shipped by secure delivery to the LEA.[4] Encrypted files of summary results are sent to the CDE by means of the SFTP. Additionally, CDs with student detail and summary results are created and delivered to the CAHSEE Office at the CDE by ETS staff located in Sacramento. Any summary results that have fewer than ten students are not reported. The statistics based on the results are also entered into the item bank.

Student Confidentiality

To meet ESEA and state requirements, school districts must collect demographic data about students. This includes information about student ethnicity, parent education, disabilities, whether the student qualifies for the NSLP, and so forth. In addition, students may reveal other information about themselves through the essays they write. ETS takes precautions to prevent any of this information from becoming public or being used for anything other than testing purposes. These procedures are applied to all documents where this demographic information may appear, including the following:

- Pre-ID files
- Reports
- Essays

Data Security

ETS is committed to safeguarding the information in its possession from unauthorized access, disclosure, modification, or destruction and adheres to strict information security policies in order to protect the confidentiality of client data. ETS staff's 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.

[4] LEA includes public school districts, statewide benefit charter schools, state board-authorized charter schools, county of education programs, and charter schools testing independently from their home district.

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 to unauthorized access or denial of service. Routers, switches, firewalls, and gateways may possess little in the way of logical access.

Facilities and procedures that protect computer files such as firewalls, intrusion detection, and virus control are in place at ETS to provide for physical security, data security, and disaster recovery. Comprehensive disaster recovery facilities are available and tested regularly at the SunGard installation in Philadelphia, Pennsylvania. ETS routinely sends backup data cartridges and files for critical software, applications, and documentation to a secure offsite storage facility for safekeeping.

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

- Scoring
- Transfer of scores by means of secure data exchange
- Statistical Analysis
- Reporting
- Internet postings
- Storage

In addition to protecting the confidentiality of testing materials, the ETS Code of Ethics further prohibits 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, such as the CAHSEE examination. The OTI verifies that these standards are followed throughout ETS. It does this, in part, by conducting periodic onsite security audits of departments with follow-up reports containing recommendations for improvement.

Procedures to Maintain Standardization

The CAHSEE processes are designed so that the tests are administered and scored in a standardized manner. ETS takes all necessary measures to ensure the standardization of the CAHSEE tests, as described in this section.

Test Administrators

The CAHSEE tests are administered seven times in a school year. To that end, ETS employs personnel who facilitate various processes involved in the standardization of an administration cycle.

The responsibilities for district and test site staff members are included in the *CAHSEE DTSCM* (CDE, 2011a). The manual is described in a later section.

The LEA staff members involved in the test administration are as follows:

CAHSEE District Coordinator

Each LEA designates a CAHSEE district coordinator who is responsible for ensuring the proper and consistent administration of the CAHSEE tests. The district coordinator also is responsible for securing testing materials upon receipt, distributing testing materials to schools, tracking the materials, training and answering questions from district staff and test site coordinators, receiving scorable and nonscorable materials from schools after an administration, and returning the materials to the CAHSEE contractor for processing.

Test Examiner

The CAHSEE is administered by test examiners who may be assisted by test proctors and scribes. A test examiner is an employee of a school district or an employee of a nonpublic, nonsectarian school (NPS). The test examiner has been trained to administer the tests and has signed a *CAHSEE Test Security Affidavit*. Test examiners must follow the directions in the *CAHSEE DFA* (CDE, 2011b) exactly.

Test Proctor

A test proctor is either an employee of the school district or a person assigned by an NPS to implement the IEP of a student. The test proctor is trained to assist the test examiner in the administration of the CAHSEE. Test proctors must sign the *CAHSEE Test Security Affidavit*.

Scribe

A scribe is either an employee of the school district or a person assigned by an NPS to implement the IEP of a student. The scribe is required to transcribe a student's responses to the format required by the test. A student's parent or guardian is not eligible to serve as a scribe. Scribes must sign the *CAHSEE Test Security Affidavit*.

CAHSEE Directions for Administration

The CAHSEE *DFA Manual* is used by test examiners to administer the CAHSEE to students. The test examiners must follow all directions and guidelines in this manual and read, word-for-word, the instructions to students in the SAY boxes to ensure test standardization.

CAHSEE District and Test Site Coordinator's Manual

Test administration procedures found in the *CAHSEE DTSCM* must be followed so all students have an equal opportunity to demonstrate their academic achievements. The *CAHSEE DTSCM* (CDE, 2011a) contributes to this goal by providing information about the responsibilities of district and test site coordinators, as well as those of the other staff involved in the administration cycle. However, the manual is not intended as a substitute for the *California Code of Regulations*, Title 5, *EC*, or to detail all of the coordinator's responsibilities.

CAHSEE Online

CAHSEE Online is a secure, Web-based application that allows CAHSEE district coordinators to order materials, submit student Pre-ID data, maintain district contact and shipping information, and correct student demographic data. Access to the online system is managed by unique passwords assigned to each CAHSEE district coordinator.

Test Booklets

For each ELA and mathematics test, multiple versions of test booklets are administered in the census administrations (February and March). The versions differ only in terms of the field-test items. These versions are spiraled, or packaged, consecutively and are distributed at the student level; that is, each classroom or group of test takers receives at least one of each version of the test. Only one version is administered in non-census administrations (May, July, October, November, and December).

The answer documents are packaged by the school and may be sorted by group(s) within the school, depending on whether the CAHSEE district coordinator provided a sorting order on the Pre-ID file. For example, if a school wanted all of a teacher's answer documents together, they would enter their own district code for the teacher and the documents would come packaged that way. All materials are sent to the CAHSEE district coordinator for proper distribution within the LEA. Special formats of test booklets are also available for test takers who require test variations to participate in testing. These special formats include audio CDs, large-print test materials, and Braille test materials.

Students with Disabilities

All students participate in the CAHSEE Program, including SWDs. Per the California *EC* Section 60852.3, SWDs are exempted from meeting the CAHSEE requirement until alternative means to the CAHSEE are implemented. During the test, students may use testing variations that are regularly used in the classroom and any accommodations or modifications specified in their IEPs or Section 504 plans. Examples of test variations only available to students who regularly use them in the classroom include special or adaptive furniture, special lighting, or being tested individually in a separate room. An accommodation is any variation in the assessment environment or process that does not fundamentally alter what the test measures or affect the comparability of test scores. Examples of accommodations for the CAHSEE include using a Braille transcription, having the mathematics section presented orally via audio presentation on a CD, or having extra time beyond that day to complete the test. A modification is any variation in the assessment environment or process that fundamentally alters what the test measures or affects the comparability of test scores. Examples of modifications for the CAHSEE include using a calculator on the mathematics section of the test; having the MC portion of the ELA section presented orally to the student, or using Manually Coded English or American Sign Language to present the MC questions of the ELA section to the student.

Identification of Students with Test Variations, Accommodations or Modifications

Most SWDs and EL take the CAHSEE without test variations. However, some of these students may need assistance when taking the CAHSEE; the assistance takes the form of test variations, accommodations, or modifications (see Table 2.4 in Chapter 2 for details). Test site coordinators are responsible for providing test examiners with any information about students who require accommodations or modifications as specified in their IEPs or Section 504 plans. If a student uses accommodations or modifications for the CAHSEE, the CAHSEE Test Site Coordinator or designated person completes a section of the answer document to indicate which accommodation(s) and modification(s) the student used on each test – ELA and mathematics. The coordinator also indicates, by signature, that the student had access to the modifications and/or accommodations specified in the student's IEP or Section 504 plan.

Scoring

The purpose of test variations and accommodations is to enable students to take the CAHSEE but not to give them an advantage over other students or to improve their scores. However, testing with a modification (e.g., calculators on the mathematics test; test questions read aloud on the ELA test) impacts the construct being measured and affects the comparability of test scores. If a student takes one or both parts of the CAHSEE with a modification and has received the equivalent of a passing score, the student has not passed that part of the exam but is eligible to request a local waiver of the requirement to meet the high school graduation requirements. In addition, scores for students tested with modifications are counted as Far Below Basic for API calculations and Not Proficient for AYP calculations, and these students are not counted in the cohort of students tested. Test administration variations and accommodations do not result in changes to students' scores for API or AYP calculations, and these students are counted as Tested. The only exception is the calculator use in the mathematics exam. This modification is accepted and counted as tested. The Proficient cut point is a little higher than that for those students who did not use a calculator.

Demographic Data Corrections

After reviewing student data, some school districts may discover demographic data that are incorrect or incomplete. The Demographics Data Corrections function of the CAHSEE Online exam gives school districts the ability to correct these data within a specified availability window.

Testing Irregularities

Testing irregularities are circumstances that may compromise the reliability and validity of test results. If more than five percent of the students tested are involved, these irregularities could affect a school's API and AYP. These scores are invalidated.

In the event of an administration irregularity, the test site coordinator completes a *Test Administration Incident Form* that details all pertinent information and immediately notifies ETS and the CDE. This form is kept at the school or district office for one year following the test. If the school and/or district determines that a student's answer document should be invalidated due to cheating, the Score Code field on the answer document is coded I. This invalidates the student's score for that portion of the test, and the student's testing status is reported as Score Invalidated. The information and procedures to assist in identifying irregularities and notifying the CDE are provided in the *CAHSEE DTSCM* (CDE, 2011a).

Test Administration Incidents

A test administration incident is any event that occurs before, during, or after test administration that does not conform to the instructions stated in the *DFA* and the *CAHSEE DTSCM* (CDE, 2011a). These events include test administration errors, disruptions, and student cheating. Except in the case of cheating, test administration incidents generally do not affect test results. These administration incidents are not reported to the CDE or the CAHSEE Program testing contractor. The CAHSEE test site coordinator should immediately notify the CAHSEE district coordinator of any test administration incidents that occur. It is recommended by the CDE that districts and schools maintain records of these incidents.

Chapter 6: Analyses

This chapter summarizes item- and test-level statistics obtained for the CAHSEE administered in 2011–12. The statistics presented in this chapter are divided into five sections and are presented in the following order:

1. Classical Item Analyses
2. DIF Analyses
3. IRT Analyses (Calibration, Scaling, and Equating)
4. Reliability Analyses
5. Validity Evidence

Each of those sets of analyses is presented in the text and in the appendices listed below.

1. Appendices 6.B through 6.H present item-level statistics for operational items. Included are the following summary tables:

    a. Item characteristics, including IRT Rasch item difficulty statistics (b-values) and point-biserial correlations for items in each operational test.

    b. The distribution of items based on their fit to the Rasch model.

    c. DIF analyses that list items flagged for significant DIF and the distributions of items across DIF categories.

    d. Summary of the ELA CR item that incorporates the polyserial correlation, DIF results, and the IRT b-value and step parameters.

2. Appendices 6.I and 6.J present item-level statistics for MC field-test items for February and March administrations, respectively. Field-test data from February and March administrations only are used to evaluate the item performance. Therefore, summary tables are provided only for the February and March administrations. Tables in Appendices 6.I and 6.J follow the same numbering system as Appendices 6.B through 6.H.

3. Appendices 6.K through 6.N present results of the reliability analyses. Appendix 6.K presents intercorrelations, reliability, and standard errors of measurement for total test scores and strand scores for the population as a whole and for selected subgroups. Appendix 6.L presents inter-rater reliability results and Appendix 6.M presents generalizability analyses for the ELA writing prompts. The results of the classifications consistency and accuracy of the Pass/Not Pass designations and ESEA performance levels are shown in Appendix 6.N.

4. Appendix 6.O presents the scoring tables obtained from the IRT equating process.

To simplify the presentation of these data, Table 6.1 provides a description of the tables located in Appendices 6.B through 6.J. Information pertaining to the operational items is found in Appendices 6.B through 6.H; information pertaining to the field-test items is found in Appendices 6.I and 6.J.

Table 6.1: Listing of Summary Tables for Items

Table[1,2]	Content	Label
6.x.1	Statistics for ELA items: IRT b-values and point-biserial correlations for all items and summarized by strand	Summary of Item Statistics – ELA
6.x.2	Statistics for Mathematics items: IRT b-value and point-biserial correlations for all items and summarized by strand	Summary of Item Statistics – Mathematics
6.x.3	IRT model fit statistics for ELA items	IRT Model Data Fit Distribution of Items – ELA
6.x.4	IRT model fit statistics for Mathematics items	IRT Model Data Fit Distribution of Items – Mathematics
6.x.5	Items flagged for significant DIF	Items Containing Significant DIF
6.x.6	DIF classifications – ELA	Distribution of DIF Classifications – ELA
6.x.7	DIF classifications – Mathematics	Distribution of DIF Classifications – Mathematics
6.x.8	Summary of ELA CR item, including the IRT b-value and step parameters, polyserial correlations, and DIF results (operational items only)	Listing of CR item statistics – ELA (operational items only)

[1] For operational items: x = Administration, where B = July, C = October, D = November, E = December, F = February, G = March, H = May.
[2] For field-test items: x = Administration, where I = February, J = March.

Samples Used for the Analyses

CAHSEE analyses were conducted at different times in the testing process and involved varying proportions of the full CAHSEE population. The reliability statistics and the correlations were calculated using the aggregate data file, which contains test results of the overall population. Following the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999, Standard 6.4), the results of the summary analyses are presented for specific populations in addition to the overall test-taking population. For the two census administrations (February and March), classical item-analysis, item-level DIF results, and IRT results were based on a sample of students used in the FIA which was close to 100 percent of students. For the five non-census administrations (May, July, October, November, December), classical item-analysis and item-level DIF results were based on a sample with a minimum of 70 percent of students. The IRT analyses for the operational items were based on a sample, also used in the equating, which comprised a minimum of 90 percent of students.

Classical Item Analyses

For each administration, classical item analyses are completed prior to DIF and item calibration, scaling, and equating. These analyses involve computing a set of statistics based on classical test theory for every item in each form. Each statistic is designed to provide some essential empirical information about the quality of each item. The statistics estimated for the CAHSEE are described below.

Classical item difficulty ("p-value"):

This statistic indicates the percentage of examinees in the sample that answered the item correctly. Desired p-values generally fall within the range of 0.25 to 0.90.

Item discrimination:

An item is considered discriminating if high-ability students tend to answer it correctly and low-ability students tend to answer it incorrectly. Item discrimination is generally assessed by comparing how the performance on an item is related to the performance on the criterion score, which is usually the total score on the test. For ELA, the test criterion score is the raw weighted composite score, which includes the total number-correct score on the MC items plus the weighted CR item score. For mathematics, the test criterion score is the total number-correct score.

Item discrimination indices used for the CAHSEE include the biserial correlation, the point-biserial correlation, and the polyserial correlation. Both the biserial correlation, or r-biserial, and the point-biserial correlation measure the relationship between a dichotomous item and the criterion score. A dichotomous item is an item that is scored as either correct or incorrect, such as the MC items in ELA or mathematics. Biserial and point-biserial correlations differ in their assumptions and in how they are computed. In biserial correlations, scores on the dichotomous item are treated as an indicator of a theoretical, underlying, normally-distributed proficiency; the biserial correlation is the estimated correlation between the total test score and a theoretical normally distributed proficiency that was dichotomized to produce the item score. For point-biserial correlations, the item scores are treated as observed binary classifications — the correct or incorrect answer.

Point-biserial correlations are computed as:

$$r_{ptbis} = \frac{(\mu_+ - \mu_-)}{\sigma_{tot}}\sqrt{pq} \qquad (6.1)$$

where μ_+ is the mean criterion score of examinees answering the item correctly; μ_- is the mean criterion score of the examinees answering the item incorrectly; σ_{tot} is the standard deviation of the criterion score of all examinees answering the item; p is the proportion of examinees answering the item correctly; and q equals (1- p).

The relationship of biserial and point-biserial correlations is presented in the following formula, which can also be used to compute the biserial correlation:

$$r_{bis} = \frac{\sqrt{pq}}{Y_{zp}} r_{ptbis} \tag{6.2}$$

where Y_{ZP} is the Y ordinate (height) of the standard normal curve at the z-score associated with the p-value for the item. The r_{bis} values are always greater than the r_{ptbis} values.

The polyserial correlation measures the relationship between a polytomous item and the criterion score. A polytomous item is an item that is scored with more than two ordered categories, such as the ELA essay. Polyserial correlations used for the CAHSEE are based on a polyserial regression model (Drasgow, 1988; Lewis & Thayer, 1996), which assumes that performance on an item is determined by the examinee's position on an underlying latent variable that is normally distributed at a given criterion score level. Based on this model, the polyserial correlation can be estimated as:

$$r_{polyreg} = \frac{\beta\, \sigma_{tot}}{\sqrt{\beta^2 \sigma_{tot}^2 + 1}} \tag{6.3}$$

where β is a series of parameters estimated from the data using maximum likelihood and σ_{tot} is the standard deviation of the criterion score.

Item discrimination indices are bounded by -1 and +1. The higher the value, the better the item distinguishes between higher- and lower-scoring examinees. Positive values indicate that the students who do well on the test have a higher probability of answering the questions correctly, while negative values also indicate that the students who do poorly on the test have a higher probability of answering the questions correctly. Therefore, negative correlations can indicate serious problems with the item content (e.g., multiple correct answers or unusually difficult or complex content), or that students have not been taught the instructional content pertaining to that item.

Percentage of students choosing each response option:

These statistics indicate the percentage of students that selects each of the available answer options.

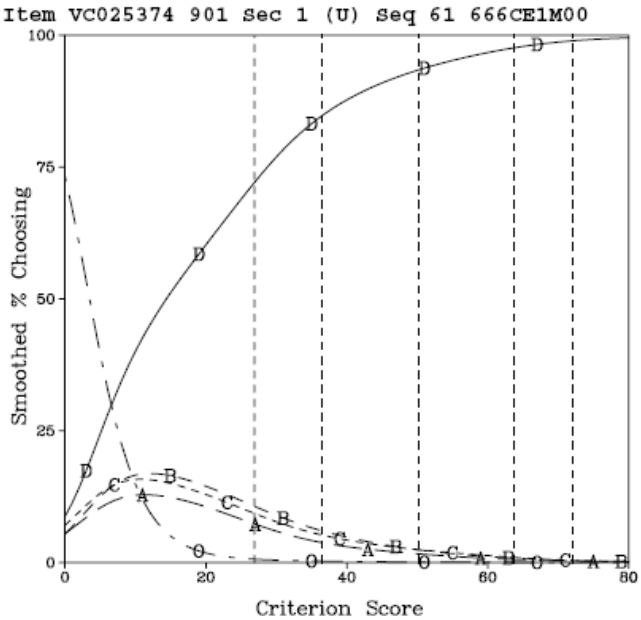
Percentage of students omitting an item:

This statistic is useful for identifying problems with test features such as testing time and item/test layout. Typically, the expectation is that if students have an adequate amount of testing time and are motivated to respond to the test questions, at least 95 percent of students should attempt to answer each question. When omit percentages exceed 5 percent for a series of items at the end of a timed section, this may indicate that there was insufficient time for students to complete all items. Alternatively, if the omit percentage is greater than 5 for a single item, this could be an indication of an item/test page layout problem. For example, students might accidentally skip an item that follows a lengthy stem.

Examples of the item analysis produced for the CAHSEE are provided in Figures 6.1 and 6.2 for two items that have been released. For each item, statistical information is presented on the right and the graphical display is presented on the left. The keyed answer is flagged with an asterisk. Statistical information includes the number and percentage of examinees choosing each option and their mean and standard deviation on the total or criterion score. The proportion of the top 20 percent choosing each option is presented on the far right column. Numbers of students reaching (Rch) or not reaching (NR) the item are also included. In addition, p-value (average item score), r-biserial (correlation with criterion), and percent reached are presented. Point-biserial correlations are not provided in these plots, but they are included in the CAHSEE client item bank.

The graphs represent response curves for the key and the item distractors, with the horizontal axis indicating the criterion score and the vertical axis (Smoothed % Choosing) displaying the examinee's probability of answering the item correctly or the probability of choosing a distractor at each criterion score. The dashed vertical lines indicate 10th, 25th, 50th, 75th, and 90th percentiles of the distribution of the criterion score so that the information represented by the graph can be referenced to student abilities. For ELA, the criterion score is a weighted sum of the MC number right score and the CR score. For mathematics, the criterion is the total number right score. Figure 6.1 provides an example of a mathematics item. It is an easy item, with a p-value of 0.89. The lowest-scoring students have about an 11 percent probability of getting the item correct. Figure 6.2 provides an example of an ELA item with moderate difficulty. The item in Figure 6.1 has an r-biserial of 0.59 and the item in Figure 6.2 has an r-biserial of 0.46. It can be seen that for both items, the probability of answering the item correctly increases as the total score/criterion increases. This indicates that both items discriminate well among students across the range of the total score.

Figure 6.1: CAHSEE Item Analysis — Mathematics Item

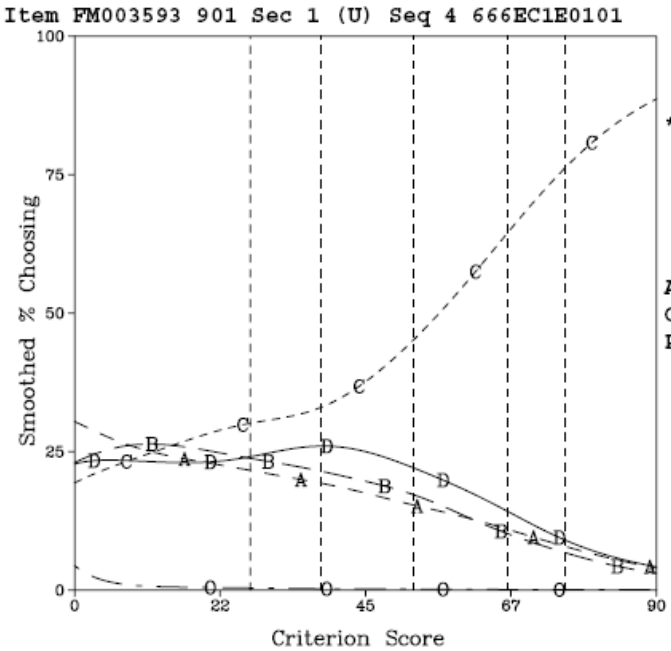


Multiple Choice, 4 choice

	N	%Tot	Criterion Mean	SD	Top 20%
A	5036	2.8	32.9	12.6	0.2
B	7374	4.1	33.3	12.3	0.3
C	6757	3.8	34.1	13.0	0.3
*D	159654	88.9	51.9	15.8	99.1
Omt	802	0.4	20.6	13.5	0.0
NR	0				
Rch	179623	100.0	49.8	16.6	

Average Item Score 0.89
Correlation with Crit. 0.59
Percent Reached 100.00

Figure 6.2: CAHSEE Item Analysis — ELA Item



Multiple Choice, 4 choice

	N	%Tot	Criterion Mean	SD	Top 20%
A	4928	15.0	45.4	16.5	6.7
B	5220	15.9	44.4	15.8	5.6
*C	16269	49.7	58.9	17.2	80.7
D	6272	19.1	46.4	15.4	6.9
Omt	70	0.2	39.9	18.0	0.1
NR	0				
Rch	32759	100.0	52.2	17.9	

Average Item Score 0.50
Correlation with Crit. 0.46
Percent Reached 100.00

Summary of Item Statistics

Classical item analyses are performed on the MC items and on the ratings of the writing prompts. Each statistic is designed to provide key information about the quality of each item from an empirical perspective. Summary statistics for the operational and field-test items are provided for ELA and mathematics overall and by content strand. Tables 1 and 2 of Appendices 6.B through 6.H provide summary statistics for the point-biserial correlations and the IRT b-values for the operational items, and Tables 1

and 2 of Appendices 6.I to 6.J provide the same information for the field-test items. When there is only one item in the strand (e.g., WA in Table 6.B.1), the standard deviation (SD) of the statistic is not available.

Table 8 of the same appendices summarizes the statistics for ELA essays used in July 2011 through May 2012 administrations. These statistics include the polyserial correlation, the DIF results for specified group comparisons, and the IRT b-values and step parameters for each score point.

Procedures for Documenting Items That Fail to Meet the Desired Psychometric Criteria

Classical item analyses are performed using the Generalized Analysis System (GENASYS). As part of the psychometric review process, Statistical Analysis staff review the GENASYS output for each item. Items with psychometric characteristics that fall outside of the expected range of values, using the criteria specified in Table 6.2, are flagged for review by test development staff. Additional items that exhibit questionable performance are flagged as well. Items are flagged for review to verify that each item is correctly keyed, that there is one clear and correct answer, and that each item is printed correctly.

Table 6.2: Flagging Criteria for Classical Item Analyses

Flag Type	Value
Low Average Item Score for dichotomous (MC) items Low Average Item Score for polytomous (CR) items	<0.25 <0.30
Low Correlation with Criterion for dichotomous (MC) items Low Correlation with Criterion for polytomous (CR) items	<0.30 <0.60
High Percent Omits	>5%
High Percent Not Responding	>5%
High Ability Group Defined as Top 20% (to identify distractors chosen by high ability examinees)	20%
High Average Item Score for dichotomous (MC) items High Average Item Score for polytomous (CR) items	>0.95 >0.70

Differential Item Functioning Analyses

One of the goals of test development is to assemble a set of items that provides an estimate of a student's ability that is as fair and accurate as possible for all groups within the population. To this end, DIF studies were conducted following the classical item analyses. DIF statistics are used to identify those items that identifiable groups of students (e.g., females, African Americans, Hispanics) with the same underlying level of ability have different probabilities of answering correctly. If the item is differentially more difficult for an identifiable subgroup, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF–flagged items might be related to actual differences in relevant knowledge or skill (item impact) or a statistical Type I error. As a result, DIF statistics are used to identify potential sources of item bias. Subsequent review by content experts and bias and

sensitivity committees is required to determine the source and meaning of any observed differences.

ETS uses two DIF detection methods: the Mantel-Haenszel approach and the standardization approach. As part of the Mantel-Haenszel procedure, the statistic described by Holland and Thayer (1988), known as MH D-DIF, is used.[5] This statistic is expressed as the differences between the focal and reference groups after conditioning on total test score. This statistic is reported on the delta scale, which is a normalized transformation of item difficulty (proportion correct) with a mean of 13 and a standard deviation of 4. Negative MH D-DIF statistics favor the reference group and positive values favor the focal group. The classification logic used for flagging items is based on a combination of absolute differences and significance testing. Items that are not statistically different based on the MH D-DIF ($p > 0.05$) are considered to have similar performance between the two studied groups; these items are considered to be functioning appropriately. For items where the statistical test indicates significant differences ($p < 0.05$), the effect size is used to determine the direction and severity of the DIF. For the ELA CR item, the Mantel-Haenszel procedure was executed where item categories are treated as integer scores and a chi-square test was carried out with one degree of freedom. The male and white groups are considered as reference groups, and the female and other ethnic groups are categorized as focal groups.

Based on these DIF statistics, items are classified into one of three categories and assigned values of A, B, or C (see Table 6.3). Items classified into category A contain negligible DIF, items in category B exhibit slight or moderate DIF, and items in category C have moderate to large values of DIF. Negative values imply that, conditional on the matching variable, the focal group has a lower mean item score than the reference group. In contrast, a positive value implies (conditional on the matching variable) that the reference group has a lower mean item score than the focal group.

[5] The formula for the estimate of constant odds ratio is:

$$\alpha_{MH} = \frac{\left(\sum_m \frac{R_{rm} W_{fm}}{N_m}\right)}{\left(\sum_m \frac{R_{fm} W_{rm}}{N_m}\right)},$$

where

R_{rm} = number in reference group at ability level m answering the item right,
W_{fm} = number in focal group at ability level m answering the item wrong,
R_{fm} = number in focal group at ability level m answering the item right,
W_{rm} = number in reference group at ability level m answering the item wrong,
N_m = total group at ability level m.

This can then be used in the following formula (Holland & Thayer, 1988):

$$MH\ D\text{-}DIF = -2.35 \ln[\alpha_{MH}].$$

Table 6.3: DIF Categories

DIF Category	Definition
A (negligible)	Absolute value of the MH D-DIF is not significantly different from zero, or is less than one. Positive values are classified as "A+" and negative values as "A-."
B (slight to moderate)	Absolute value of the MH D-DIF is significantly different from zero but not from one, and is at least one; OR Absolute value of the 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 (moderate to large)	Absolute value of the MH D-DIF is significantly different from one, and is at least 1.5. Positive values are classified as "C+" and negative values as "C-."

For CR items, the MH D-DIF statistic is not calculated; instead, the standardization procedure is used in conjunction with the Mantel chi-square statistic. Analogous flagging rules have been developed that are used to classify the CR items into A, B, or C DIF categories. The flagging criteria for CR items are:

A) If the Mantel chi-square p-value > 0.05, or if the Mantel chi-square p-value < 0.05 but the absolute value of the Standardized Mean Difference (SMD)/SD $\leq$ 0.17, the item is classified as A.

B) If the Mantel chi-square p-value < 0.05 and 0.17 < |SMD/SD| $\leq$ 0.25, then the item is classified as B.

C) If the Mantel chi-square p-value < 0.05 and |SMD/SD| > 0.25, then the item is classified as C.

Positive values favor the focal group and negative values favor the reference group.

All DIF analyses were performed according to the procedures specified in the document "CAHSEE DIF Procedures" dated February 22, 2002. DIF comparison groups are based on gender (Male compared to Female), ethnicity (White compared to American Indian, Asian, Pacific Islander, Filipino, Combined Asian,[6] Hispanic, and African American), and English language proficiency (English proficient compared to EL[7]). Operational items flagged for C DIF are reviewed by an expert committee consisting of CAHSEE item development staff, the CDE staff responsible for the CAHSEE, external educators identified by the CDE, and additional CDE content experts in ELA and mathematics, as needed, to ensure that the items are free from any bias before being used to compute final test scores. Some items were flagged for DIF in more than one administration. If an item passed the review process by the CAHSEE DIF Review Panel, it is not reviewed a second time.

Details of the results for each administration can be found in Tables 5, 6, and 7 of Appendixes 6.B to 6.H for the operational items and Appendices 6.I and 6.J for field-testing items in census administrations.

[6] Combined Asian group includes students from Asian, Filipino, and Pacific Islander groups.
[7] According to "CAHSEE DIF Procedures," DIF analyses were performed for EL on mathematics items only. From September 2005, DIF analyses were also performed for EL on ELA items.

Item Response Theory Analyses (Calibration, Scaling and Equating)

The CAHSEE tests are equated to a reference form using a common-item nonequivalent groups design and methods based on IRT. The “base” or “reference” scale for the CAHSEE was established in the February 2004 administration. The 2011 items were placed on the reference scale through a set of linking items that appeared in the 2010 operational forms and were re-administered in 2011. The procedures used for equating the CAHSEE involve three steps: item calibration, item parameter scaling, and true-score equating.

Measurement Model

Items are calibrated using the Rasch model for the MC questions and the Rasch partial-credit model for the ELA CR item. The section *Model for Generating Item Statistics* in Chapter 2 provides further details on both models.

ETS uses GENASYS for the IRT item calibration and equating work. As part of this system, a proprietary version of the PARSCALE computer program (Muraki & Bock, 1995) is used and parameterized to produce one-parameter calibrations. Research conducted at ETS suggests that PARSCALE calibrations produce results that are virtually identical to results based on WINSTEPS (Way, Kubiak, Henderson, & Julian, 2002), which is a program often used to perform Rasch scaling.

Item Calibration and Scaling

The samples used for item calibration, scaling, and equating include scanned and scored student records provided in statistical file extracts. Typically, over 90 percent of the student data are available for equating. Students taking special test forms (large-print, audio CD, and Braille) are excluded from the equating sample. Incomplete data records are also removed. In addition, data records are eliminated based on analyses of performance on different sections of the tests. Specifically, outlier scores are identified for mathematics and ELA by comparing scores on the first and second sections of the test, and for ELA by comparing scores on the MC and CR components of the test. Finally, the equating samples exclude students who did not indicate a valid test form code. For the purposes of item calibration, scaling, and equating and the production of scoring tables for score reporting, only operational items are included. Field-test items are analyzed and calibrated separately.

The PARSCALE program is run in two stages. In the first stage, the estimation imposes normal constraints on the updated subject prior distribution (θ). The estimates resulting from this first stage are used as starting values for a second PARSCALE run, in which the subject prior distribution is updated after each expectation-maximization (EM) cycle. For both stages, the multiplicative metric of the scale is controlled by the use of the fixed discrimination parameter.

The resulting calibrations are then transformed to the existing scale, using the Stocking and Lord (1983) test characteristic equating procedure. Because only a constant is added to the new item parameter estimates, this procedure is essentially equivalent to setting the means of the new estimates equal to their anchor item values. The linking process is iterative and involves an inspection of differences between the

new estimates and the anchor estimates for the linking items. Items with large weighted root-mean-square differences (WRMSD) between item characteristic curves (ICCs) based on the old and new difficulty estimates are eliminated from the Stocking and Lord equating and the linking constants are re-estimated. The differences are calculated as follows:

$$WRMSD = \sqrt{\sum_{j=1}^{N_g} w_j [P_n(\theta_j) - P_r(\theta_j)]^2} \tag{6.4}$$

where abilities are grouped in the intervals of 0.005 between –4.0 and 4.0, θ_j is the mean of the abilities in the interval j, N_g is the number of intervals, w_j is a weight equal to the proportion of estimated abilities from the transformed new form in interval j, $P_n(\theta_j)$ is the probability of correct response for the transformed new form item at ability level θ_j, and $P_r(\theta_j)$ is the probability of correct response at ability level θ_j for the reference form (i.e., the item bank estimates). Any linking items for which the difference WRMSD is greater than 0.125 are eliminated from the anchor set. This criterion was established in early CAHSEE calibrations, has produced reasonable results over time, and has been used satisfactorily for CAHSEE administrations and other testing programs.

For each administration, plots of new transformed difficulty estimates against the reference estimates are then produced and inspected. Results across administrations have consistently indicated high correlations between the new and reference difficulty estimates, and typically no more than two or three items are eliminated from the linking because of large WRMSD differences. In general, the correlation between the new and existing difficulty estimates tends to be slightly higher for mathematics than for ELA, possibly because ELA items are passage-dependent and more susceptible to context and position effects.

True-Score Equating

Once the items for the 2011–12 administration were calibrated and linked to the operational theta scale, IRT true-score equating procedures were utilized to transform the new form to the base form scale established in February 2004. The true-score equating procedure is based on the relationship between raw scores and ability. For mathematics, which consists entirely of MC items, this is the well-known relationship defined in Lord (1980; eq. 4–5):

$$\xi(\theta) = \sum_{i=1}^{n} P_i(\theta) \tag{6.5}$$

where $P_i(\theta)$ is the probability of a correct response to item i at ability level θ (defined by the Rasch model) and $\xi(\theta)$ is the corresponding true score, and the summation is over the n items in the test.

For ELA, $\xi(\theta)$ is based on a weighted sum of MC and CR items, and the relationship can be defined as:

$$\xi(\theta) = w_{mc} \sum_{i=1}^{n_{mc}} P_i(\theta) + w_{cr} \sum_{j=1}^{n_{cr}} \sum_{x=1}^{m} s_x P_{xj}(\theta) \tag{6.6}$$

where w_{mc} = 1.000, w_{cr} = 4.500, s_x is the score value for category x, n_{mc} is the number of MC items in the test, n_{cr} is the number of CR items in the test, m is the number of score categories in each CR item, and $P_{xj}(\theta)$ is the probability of a score in category x at ability θ (defined by the Rasch partial-credit model). For the ELA writing item, there are eight possible scores: 0, 1, 1.5, 2, 2.5, 3, 3.5, and 4.

For each integer score ξ_n on the new form, the true-score equating procedure first solves for the corresponding ability level using equations 6.5 (for mathematics) or 6.6 (for ELA). Next, the procedure uses that ability level (θ) to find the corresponding score, $\xi_{b,}$ on the base form. Finally, each score ξ_b is transformed to the CAHSEE reporting-score scale using the raw-score-to-scale-score conversion table developed for the February 2004 administration and linear interpolation.

For both ELA and mathematics, the 2011–12 forms were linked back to scale established in February 2004.

Complete raw-to-scale score conversion tables for the 2011–12 CAHSEE are presented in Appendix 6.O. The raw scores and corresponding unrounded converted scale scores are listed in those tables. For all the 2011–12 CAHSEE administrations, scale scores were adjusted at both ends of the scale so that the minimum was 275 and the maximum reported scale score was 450. The scale scores defining the cut scores for passing and ESEA proficiency levels are indicated on the conversion tables.

Equating Braille, Large-Print and Audio CD Forms

The large-print and audio CD versions of the test forms were identical to the standard form administered for all administrations. The July 2011 operational form was used as the Braille, Braille large-print, and Braille audio CD forms for the July, October, and November 2011 administrations for both ELA and mathematics. The December 2011 operational form was used for the December 2011 and February 2012 administrations, and the March operational form was used for the March and May 2012 administrations of the Braille, Braille large-print, and Braille audio CD forms for both tests. The Braille versions did not have field-test items. These special versions included operational items from standard administrations; therefore, no special equating analyses were required.

Raw-Score to Scale-Score Conversion Tables and Conditional Standard Errors of Measurement

Following the equating analyses, raw-score to scale-score conversion tables and conditional standard errors of measurement (CSEM) were produced. CSEM for CAHSEE scale scores are based on IRT and are calculated by the IRTEQUATE module in GENASYS. For mathematics, where reported scores are based on number-correct scores, the calculation of the CSEM based on Rasch model difficulty estimates is straightforward. However, for ELA, reported scores are based on a weighted composite of the MC and CR items. Because the raw-to-scale score conversions for the base form are nonlinear, the scale-score CSEM estimated in GENASYS are characterized by minor irregularities that are smoothed in a subsequent step. Operational, large-print, audio books, and audio CD score conversions and the smoothed CSEM at score points for the ELA and mathematics

tests are presented in Tables 1 to 14 of Appendix 6.O; the scoring tables for the Braille forms are presented in Tables 15 to 20.

Appendix 6.P presents the equations for calculating the standard errors of theta (θ) based on weighted raw scores using the Rasch and Rasch partial-credit models.

Scaling Field-Test Items

The MC field-test items are embedded in operational forms and their IRT parameters are placed on the same base scale as the operational items. The field-test item scaling is carried out by fixing the IRT item parameters of the operational items and estimating the item parameters for the field-test items. Consistent with the calibration process for operational items, the PARSCALE program is constrained for the Rasch model by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588) and by fixing the lower asymptote for all MC items to zero. Items with biserial correlations less than or equal to 0.10 are not included in the calibration.

Item Response Theory Model-Data Fit Analyses

Because the Rasch model is used in equating the CAHSEE, an important part of IRT analyses is the assessment of model-data fit. Statistics describing the fit of the Rasch model to the data, reported in letter categories of A, B, C, D, and F (IRT flag), are produced. A description and examples of this model-data fit-rating scheme are provided in Appendix 6.A. In general, items with flagging categories of A, B, or C are all considered acceptable. Ratings of D are considered questionable, and the ratings of F indicate a poor model fit. All items receiving a rating of F are also rated as Do Not Use. The test developers are asked to avoid the items flagged as D if possible and to carefully review them if they must be used. Test developers are instructed to not select items rated F for operational test assembly without a review by a psychometrician.

For the census (February and March) administrations, the evaluation of model fit is performed twice in an administration cycle. This assessment is first performed on operational items before scoring tables are produced and released and is performed again on field-test items as part of the FIA. The flags produced as a result of this assessment are added to the item bank. For the non-census administrations (July, October, November, December, and May), the assessment is performed only on operational items before scoring tables are produced and released.

The distributions of the operational items across the IRT model-data fit classifications are presented in Tables 3 and 4 of Appendices 6.B through 6.H. The numbers and percentages of field-test items that received IRT model-data fit ratings of A, B, C, or D and were recommended for use in future operational tests are summarized in Table 6.4. The distributions for the field-test items are presented in Tables 3 and 4 of Appendices 6.I and 6.J.

Table 6.4: Field-Test Items Recommended for Future Forms

Administration Date	Subject	Number of Items (Percent)
February 2011	ELA	694 (80)
	Mathematics	624 (72)
March 2011	ELA	1,049 (86)
	Mathematics	936 (86)

Summaries of Scaled Item Response Theory B-Values

Once the IRT b-values are placed on the base scale, analyses are performed to assess the overall test difficulty, the difficulty level of strands, and the distribution of items in a particular range of item difficulty. Summary statistics for operational items are provided for ELA and mathematics overall and by content strand in Tables 1 and 2, respectively, of Appendices 6.B through 6.H. In addition, results are provided for mathematical reasoning, a secondary strand based on items that are integrated into other mathematics content strands. Summary statistics for field-test items are provided for ELA and mathematics overall and by content strand in Tables 1 and 2, respectively, of Appendices 6.I to 6.J.

Reliability Analyses

Reliability focuses on the extent to which differences in test scores reflect true differences in the knowledge, ability, or skill being tested rather than fluctuations due to chance or factors other than those of interest. The reliability analyses included in this section are reliability statistics and standard errors of measurement (SEMs) for total test, strands, and subgroups; intercorrelations between the two content areas and between different sub-strands; inter-rater agreement on the essay prompts; and the results from the classification accuracy and consistency analyses.

Test Reliabilities and Standard Errors of Measurement

The variance in the distributions of test scores, essentially the differences among individuals, is partly due to real differences in the knowledge, skill, or ability being tested (true variance) and partly due to random errors in the measurement process (error variance). The number used to describe reliability is an estimate of the proportion of the total variance that is true variance. Several different ways of estimating this proportion exist. The estimates of reliability reported in this report are internal-consistency measures. Therefore, they apply only to the test form being analyzed. They do not take into account form-to-form variation due to equating limitations or lack of parallelism, nor are they responsive to day-to-day variation due, for example, to the examinee's state of health or the testing environment. Reliability coefficients range from 0 to 1. The higher the reliability coefficient for a set of scores, the more likely individuals would be to obtain very similar scores if they took another form of the test. The formula for the internal consistency reliability, as measured by Cronbach's Alpha (Cronbach, 1951), is reported below:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma_i^{\,2}}{\sigma_x^2}\right] \tag{6.7}$$

where n is the number of items, σ_i^2 is the variance of scores on the i-th item, and σ_x^2 is the variance of the total score (sum of scores on the individual items).

When test scores are a composite of the MC and CR items, the reliability estimates can be computed by the following formula (Feldt & Brennan, 1989):

$$\alpha_c = 1 - \frac{\sum_{j=1}^{k} \left(w_j \sigma_{e_j}\right)^2}{\sigma_c^2} \tag{6.8}$$

where k is the number of part scores in the composite, w_j is the weight associated with the j-th part score, $\sigma_{e_j}^2$ is the SEM of the j-th part score, and σ_c^2 is the variance of the composite score.

The reliability of the CR items can be estimated indirectly by examining the correlation between the MC and CR item components in relation to the MC reliability. The lower-bound reliability for a CR item in a test with MC items and only one CR item can be found using the squared correlation between the MC and CR item portions of the test and dividing by the reliability of the MC portion of the test:

$$\alpha_{cr} = \frac{r_{mc.cr}^2}{\alpha_{mc}} \tag{6.9}$$

Reliability estimates for the ELA section across the seven administrations ranged from 0.87 to 0.94. Reliability estimates for the mathematics section across the seven administrations ranged from 0.85 to 0.95. Reliabilities for the MC strands for ELA ranged from 0.37 to 0.84, while the strand reliabilities for mathematics ranged from 0.32 to 0.85. The reliabilities for the ELA CR (unweighted) item score ranged from 0.26 to 0.43.

The SEM provides a measure of score instability in the score metric. The formula for computing the SEM is:

$$\sigma_e = \sigma_x \sqrt{1-\alpha} \tag{6.10}$$

where reliability is the reliability estimated using formulas 6.7, 6.8, or 6.9 above, and σ_x is the standard deviation of the score being examined.

The SEM is particularly useful in determining the confidence interval (CI) that captures an examinee's true score. Assuming that measurement error is normally distributed, it can be said that upon infinite replications of the testing occasion, approximately 95 percent of the CIs of ±1.96 SEM around the observed score would contain an examinee's true score (Crocker & Algina, 1986).

The SEMs for the ELA total raw scores and the mathematics raw scores across the 2011–12 administrations ranged from 3.81 to 4.48 and 3.56 to 4.16, respectively. The reliabilities and SEMs of the test scores may be found in Appendix 6.K, Tables 1 to 7.

Strand Intercorrelations, Reliabilities and Standard Errors of Measurement

Intercorrelations are the correlations between the raw scores obtained on the different tests or strands by individual test takers. Intercorrelations, reliabilities, and SEM estimates for both assessments and for the content strands within each assessment are reported in Appendix 6.K. Note: the scoring weight of 4.5 was not applied to ELA essay scores when computing the reliability and intercorrelation statistics for the essay items.

Across the 2011–12 administrations, the correlations between ELA strands (including essay) ranged from 0.36 to 0.80, while the correlations between mathematics strands ranged from 0.25 to 0.82. The reliabilities of the ELA strands ranged from 0.37 to 0.84, and the reliability of the mathematics strands ranged from 0.32 to 0.85. The SEM for the MC strands for ELA ranged from 0.96 to 2.01, while the SEM for mathematics ranged from 1.35 to 2.08. The SEM for the ELA CR items ranged from 0.38 to 0.53.

Subgroup Reliabilities and Standard Errors of Measurement

Reliabilities (R_{XX}) and SEM estimates are reported for subgroups from the February and March 2012 census administrations, where larger case counts are available. Table 6.K.8 shows R_{XX} and SEM for the MC total and the composite score for gender, ethnicity, English-proficient and EL, and accommodation and non-accommodation groups for the ELA February 2012 administration. Table 6.K.9 provides the same information for the mathematics test of the February 2012 administration. Tables 6.K.10 and 6.K.11 provide R_{XX} and SEM for the subgroups for ELA and mathematics for the March 2012 administration. The sample sizes for these R_{XX} also appear in these tables.

Reliability estimates for the ELA composite (MC + Essay) scores are similar for the gender groups across the two administrations. The SEM is slightly lower for the female group than for the male group. In general, R_{XX} are similar for the ethnic groups; the SEM are similar for most groups with a range of 2.81 (Asian, March) to 3.61 (African American, February) for the MC total and a range of 3.61 (Asian, Filipino, White, March) to 4.24 (African American, February) for the composite. All R_{XX} are greater than or equal to 0.88.

The R_{XX} for the accommodation group are the same for March and slightly lower for the accommodation group for February. The SEM are higher for the accommodation group. R_{XX} for the EL group are lower than that for the English-proficient group. SEM is higher for the EL group than the English-proficient group.

R_{XX} and SEM for the mathematics total score are similar for gender groups. R_{XX} for ethnic groups are similar; the SEM are similar for most groups with a range of 2.77 (Asian, March) to 3.97 (African American, February). The R_{XX} for the accommodation groups are slightly lower than for the nonaccommodation groups, and the SEM are slightly higher. R_{XX} for the English-proficient group is slightly higher than that for the EL group; whereas, the SEM is higher for the EL group than for the English-proficient group. All R_{XX} are greater than or equal to 0.91.

Writing Prompt and Rater Agreement Summary

The CAHSEE ELA CR section consists of a single writing prompt that may be passage-based or stand-alone. All of the following categories of writing are randomly rotated to appear in the test administrations: response to literature or analytic essay (expository writing), biography, persuasion, or business letter. The standards require students to combine the rhetorical strategies of narration, exposition, persuasion, and description to produce texts of at least 1,500 words each.

Every response to the ELA writing prompt is rated by two different readers on a scale between 1 and 4. Non-valid responses are scored as zero. The scores from both readers must be the same or adjacent (within one score point of each other). If the difference between the two scores is more than one score point, it is considered discrepant, and a scoring leader provides a third score, which becomes the score of record. The ratings obtained from the first two readers are used to carry out inter-rater agreement and generalizability analyses to assess the reliability of the writing scores.

Rater Agreement

An important part of the analysis of any multiple-rated CR item is the degree to which the individual rater scores agree. Tables 1, 3, 5, 7, 9, 11, and 13 in Appendix 6.L present the possible score combinations, the distribution of the two ratings, and the distribution of differences between the first and second ratings of each CR item. As the diagonals of the tables show, the majority of raters assigned the same score. When ratings differed, most were by only one score point. Only the February and March 2012 administrations had about 1 percent of CR item scores resulted in discrepant scores.

Tables 2, 4, 6, 8, 10, 12, and 14 in Appendix 6.L present the mean of the first and second ratings for each CR item and the corresponding standard deviation, mean absolute difference, and the correlation between the first and second ratings. The mean absolute difference between the first and second ratings for the seven administrations ranged from 0.17 to 0.30. The correlation between first and second ratings across the seven administrations ranged from 0.64 to 0.75.

The even-numbered tables also summarize the reasons given for CR items that received a score of zero. Zero scores are given if a CR item is left blank, or if the response is illegible, does not address the topic, is a cartoon/drawing, uses inappropriate content, or is not written in English. During the seven administrations, the percentage of zero scores received on the CR item ranged from 1 to 4 percent, with the majority of zero scores due to CR items that were either left blank or written off topic. The lowest percentage of zero scores for the CR item was in the December 2011 and the March 2012 administrations and the largest percentage was in the May 2012 administration.

Generalizability Analyses

Generalizability analyses were performed to estimate the proportion of variance explained by possible sources of variation, including raters and persons (desired variance). A person crossed with rater design, or P x R design, was used for the

generalizability analyses. This design assumes that the examinee could be rated by any rater in the pool of raters. Theoretically, this is correct, but the practicality of producing scores in a timely manner prevents this from happening operationally. Therefore, the generalizability and dependability coefficients produced will be conservative estimates of score generalizability. The analyses were conducted using the GENOVA software group (Brennan, 2001; Crick & Brennan, 1983). A generalizability study (g-study) was performed to estimate variance components for selected sources of variation, also known as “facets.” A decision study (d-study) was performed to estimate the generalizability and dependability coefficients for the operational design.

The results of the generalizability analyses are presented in Appendix 6.M. Generalizability coefficients for the CR item ranged from 0.77 to 0.86 across the seven administrations. The rater facet was consistently estimated to have minimal effect on score variance across the seven administrations.

The fluctuation in generalizability coefficients across administrations can be attributed to the changes in the amount of variation resulting from differences among students. Generally, as facets other than universe and person are found to contribute to the score variance, the generalizability coefficient will decrease, indicating that score differences are a function of more than individual differences in the construct being assessed. Therefore, more of the student’s score is due to error, and score reliability is reduced. However, when the person variance is examined across administrations, the percentage of variance attributable to individual differences is very similar and large. The final interaction term is confounded with undifferentiated error that is not accounted for in the current design and represents the second largest source of score variance.

Decision Classification Analyses

The method used for estimating the reliability of classification decisions is described in Livingston and Lewis (1995) and is implemented using the proprietary computer program RELCLASS-COMP (Version 4.14). The program provides two statistics that describe the reliability of classifications based on test scores from an administration of one form. Decision *consistency* describes the extent to which examinees are classified in the same way as they would be on the basis of a single form of a test other than the one for which data are available. Decision *accuracy* describes the extent to which examinees are classified in the same way as they would be on the basis of the average of all possible forms of a test.

RELCLASS-COMP estimates decision *consistency* using an estimated multivariate distribution of reported classifications on the current form of the examination and classifications on parallel forms. RELCLASS-COMP estimates decision *accuracy* using an estimated multivariate distribution of reported classifications on the current form of the examination and the classifications based on an all-forms average (true score). In each case, the proportion of classifications with exact agreement is the sum of the entries in the diagonal of the contingency table representing the multivariate distribution. Reliability of classification at the cut score is estimated by collapsing the multivariate distribution at the passing score boundary into an *n* by *n* table (where *n* is

the number of proficiency levels) and summing the entries in the diagonal. Note that the proportions may not add up to 1 due to rounding.

The reliability of classification results for the ESEA cut scores at the Advanced and Proficient levels for both accuracy and consistency are reported in Appendix 6.N, Tables 1, 3, 5, 7, 9, 11, and 13. Across the seven administrations, the accuracy of classifying students into the Advanced level versus Proficient level or below for ELA ranged from 0.92 to 0.99, while the accuracy for mathematics ranged from 0.94 to 0.99. The consistency of classifying students into the Advanced level versus Proficient level or below for ELA ranged from 0.89 to 0.98, while the consistency for mathematics ranged from 0.92 to 0.99.

The accuracy of classifying students into the Proficient level or above versus Below Proficient level for ELA ranged from 0.92 to 0.97, while the accuracy for mathematics ranged from 0.89 to 0.98. The consistency of classifying students into the Proficient level or above versus Below Proficient level for ELA ranged from 0.90 to 0.98, while the consistency for mathematics ranged from 0.91 to 0.97.

The reliability of classification results for the cut points at the Pass/Not Pass level for both accuracy and consistency are reported in Appendix 6.N, Tables 2, 4, 6, 8 , 10, 12, and 14. Across the seven administrations, the decision accuracy for ELA at the Pass/Not Pass levels ranged from 0.90 to 0.94, while the decision accuracy for mathematics at the Pass/Not Pass levels ranged from 0.90 to 0.95. The decision consistency for ELA at the Pass/Not Pass levels ranged from 0.86 to 0.92, while the decision consistency for mathematics at the Pass/Not Pass level ranged from 0.86 to 0.93. The magnitude of the estimates (0.86 or above) reflect a high level of accuracy and consistency in the student classifications.

Validity Evidence

Validity refers to the degree to which evidence supports a proposed interpretation or use of a set of scores, and it is one of the most fundamental considerations in developing and evaluating tests (AERA, APA, & NCME, 1999). Validity is not based on a single study or type of study but involves an ongoing process of gathering evidence supporting the interpretation or use of the resulting test scores. The process is iterative in nature and begins with the test design and continues throughout the entire assessment process, including design, content specifications, item development, examination of psychometric quality, and inferences made from the results.

This section presents the evidence gathered to support the intended uses and interpretations of scores for the CAHSEE testing program. The description is organized in the manner prescribed in the *Standards for Educational and Psychological Testing* (1999), published jointly by AERA, APA, and NCME. These standards require a clear definition of the purpose of the test, which includes a description of the qualities, called constructs, which are to be assessed by a test, and the population to be assessed. The standards also require how the scores are to be interpreted and used. In addition, the *Standards* identify five kinds of evidence that can provide support for score interpretations and uses, which are as follows:

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. Additional validity evidence

These kinds of evidence are also defined as important elements of validity information in documents developed by the U.S. Department of Education for the peer review of testing programs administered by states in response to the ESEA of 2001 (USDOE, 2009).

The next section defines the purpose of the CAHSEE, followed by a description and discussion of the kinds of validity evidence that have been gathered.

Test Purpose

As mentioned in Chapter 1, the purpose of the CAHSEE is to assess student achievement in public high schools and to ensure that students who graduate from public high schools can demonstrate minimum competency in reading, writing, and mathematics as defined by the grade ten content standards in ELA and mathematics. Additionally, the CAHSEE is used in determining AYP that applies toward meeting the requirement of the federal ESEA of 2001, which is to have all students score Proficient or above by 2014.

Constructs to Be Measured

The CAHSEE ELA and mathematics examinations are designed to show how well students perform relative to the California content standards. These content standards were approved by the SBE; they describe what students should know and be able to do. The ELA examination measures reading and writing standards through grade ten. The mathematics examination measures grades six and seven mathematics and Algebra I standards.

The CAHSEE test blueprints provide the number of items per standard that will appear on an operational form. The CAHSEE item specifications provide the general characteristics of the items for each content standard, including an operational definition of the construct, appropriate or inappropriate item types or content, administration instructions, and the rules used to score examinee responses. By following the test blueprints and item specifications to construct the CAHSEE tests, as many aspects of the measurement procedure as possible are controlled so that the testing conditions will remain the same over test administrations (Cronbach, 1971; Cronbach, Gleser, Nanda, & Rajaratnam, 1972) to minimize construct irrelevant score variance (Messick, 1989). The blueprints for the CAHSEE ELA and mathematics examinations can be found on the CDE CAHSEE Program Resources Web page at http://www.cde.ca.gov/ta/tg/hs/resources.asp.

Intended Test Population

The target population for the CAHSEE consists of students who are either enrolled in California public high schools in grades ten, eleven, or twelve, or are enrolled in adult schools operated by a school district. Beginning in the 2005–06 school year, with the exception of eligible SWDs, no student will receive a public high school diploma without passing the CAHSEE and meeting all other state and district requirements for graduation.

Scores Generated and the Interpretations and Uses of These Scores

Total scores expressed as scale scores, student performance levels, and strand scores for each strand are generated for each subject area test. On the basis of a student's total score, an inference is drawn about how much knowledge and skill in the subject area the student has achieved. The total score also is used to classify students in terms of their levels of knowledge and skills in a subject area. These three levels are called performance levels and are as follows: Advanced, Proficient, and Pass.

Strand results show an individual student's raw score and percent-correct score. A detailed description of the uses and applications of CAHSEE scores is presented in Chapter 8.

The CAHSEE tests provide results or score summaries that are used for different purposes. The three major purposes are:

- Communicating with parents and guardians about the students' achievements and if the students fulfill one or both part(s) of the CAHSEE for the graduation requirement.

- Informing decisions needed to provide additional assistance for students who did not pass one or both part(s) of the CAHSEE.

- Providing data for state and federal accountability programs for schools.

These are the only uses and interpretations of scores for which validity evidence has been gathered. If the test user wishes to interpret or use the scores in other ways, the user is cautioned that the validity of doing so has not been established (APA, AERA, & NCME, 1999, Standard 1.3). The user is advised to gather evidence to support these additional interpretations or uses (APA, AERA, & NCME, 1999, Standard 1.4).

Evidence Based on Content

According to the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (1999), analyses that demonstrate a strong relationship between a test's content and the construct that the test was designed to measure can provide important evidence of validity. In current K–12 testing, the construct of interest usually is operationally defined by state content standards and the test blueprints that specify the content, format, and scoring of items that are admissible measures of the knowledge and skills described in the content standards. Evidence that the items

meet these specifications and represent the domain of knowledge and skills referenced by the standards supports the inference that students' scores on these items appropriately can be regarded as measures of the intended construct.

As noted in the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (1999), evidence based on test content may involve logical analyses of test content in which experts judge the adequacy with which the test content conforms to the test specifications and represents the intended domain of content. Such reviews can also be used to determine whether the test content contains material that is not relevant to the construct of interest. Analyses of test content may also involve the use of empirical evidence of item quality.

Also to be considered in evaluating test content are the procedures used for test administration and test scoring. As Kane (2006, p. 29) has noted, although evidence that appropriate administration and scoring procedures have been used does not provide compelling evidence to support a particular score interpretation or use, such evidence may prove useful in refuting rival explanations of test results. Evidence based on content includes the following:

Descriptions of the state standards—As was noted in Chapter 1, the SBE adopted rigorous content standards in ELA and mathematics. These standards were designed to guide instruction and learning for all students in the state and to bring California students to world-class levels of achievement.

Specifications and blueprints—ETS maintains item development specifications for both CAHSEE tests. The item specifications describe the characteristics of the items that should be written to measure each content standard. A thorough description of the specifications can be found in Chapter 4. Once the items are developed, ETS selects all CAHSEE test items to conform to the SBE-approved California content standards and test blueprints. Test blueprints for the components of the CAHSEE program were proposed by ETS, reviewed and approved by the CDE, and presented to the SBE for adoption. There have been no recent changes in the blueprints. The content blueprints for the CAHSEE can be found on the CDE CAHSEE Program Resources Web page at http://www.cde.ca.gov/ta/tg/hs/resources.asp.

Item development process—Detailed descriptions of the content and psychometric criteria applicable to the construction of the 2011-12 CAHSEE are presented in Chapter 4.

Item review process—Chapter 3 explains the extensive item review process applied to items written for use in the CAHSEE. In brief, items written for the CAHSEE undergo multiple review cycles and involve multiple groups of reviewers. The content review is conducted by external reviewers. The content review committees are responsible for reviewing all newly developed items for alignment to the California content standards. Additionally, the SPAR is responsible for reviewing and approving test items before they are used operationally or in field tests. The SPAR examines the items for intrusiveness into students' personal lives such as student and family beliefs, morality, religion, or sexuality. More information about the SPAR is given in Chapter 3.

Form construction process—For each test, the content standards, blueprints, and test specifications are used as the basis for choosing items. Additional targets for item difficulty and discrimination that are used for test construction were defined in light of

what are desirable statistical characteristics in test items and statistical evaluations of the CAHSEE items. Guidelines for test construction were established with the goal of maintaining parallel forms to the greatest extent possible from year to year. Details can be found in Chapter 4.

Alignment study—Strong alignment between standards and assessments is fundamental to meaningful measurement of student achievement and instructional effectiveness. An alignment study evaluates the extent to which there is overlap between the test content and the standards. This is used to establish whether the test material reflects the skills and knowledge that students are expected to demonstrate. A universal test design study evaluates a test for appropriate format, scope, and content for various student populations, such as students with limited English proficiency and SWDs. The results of the alignment study and universal test design study provide useful information regarding test validity.

The Human Resources Research Organization (HumRRO) performs yearly alignment studies on the CAHSEE to verify the quality of the tests. The most recent alignment study results are presented in Taylor, Johnstone, and Hardoin (2011). HumRRO and the National Center for Educational Outcomes staff served as the reviewers to establish an independent evaluation of content alignment.

The Webb Alignment method (2005) was used to evaluate the alignment of the 2011 ELA and mathematics tests to the California content standards. The Webb method consists of four dimensions. Each dimension provides information about the extent of alignment between the assessment and content standards.

Categorical concurrence is a general measure of content match between the test and state standards indicating the number of items addressing each content strand.

DOK measures the level of cognitive processing required by items compared to the processing outlined by the content standards.

Range-of-knowledge correspondence focuses on the breadth of content assessed compared to the state standards.

Balance-of-knowledge representation considers the number of items matched to each content standard per strand.

The review was conducted for the ELA and mathematics items based on the March 2011 administration. The results indicated that the mathematics items assessed most standards broadly and at an appropriate level of complexity, except for MR. The standards for MR require a higher level of processing (i.e., DOK level 3 or 4). Not many items assessing higher-order thinking skills were found. The ELA items showed lower alignment to some standards, especially on the DOK assessed. Items in the Reading Comprehension and Writing Applications have narrowly covered standards within these strands.

Data were presented for the results of each alignment study HumRRO has conducted. The mathematics studies were conducted in 2005, 2008, and 2011, and the ELA studies were conducted in 2005, 2008, 2009, and 2011. The results show fairly consistent results across studies for both content areas, indicating test form alignment is relatively stable over time. Most outcomes surpass the minimum criterion specified for the alignment measure.

The same reviewers and test forms used for the alignment study were employed for the accessibility review of universal test design. The panelists were asked to rate the quality of the items in terms of lack of flaws and accessibility to all students. The test forms demonstrated many instances of conformity to universal design practices, including appropriate grade level vocabulary and sentence complexity, inclusion of commonly used words, sensitivity to test-taker characteristics, and identifiable questions. Some concerns were raised about aspects of the visual presentation of the items.

The independent evaluation reports are available on the CDE CASHEE Independent Evaluation Reports Web page at http://www.cde.ca.gov/ta/tg/hs/evaluations.asp.

Evidence Based on Relations to Other Variables

Empirical results concerning the relationships between scores on a test and measures of other variables external to the test can also provide evidence of validity when these relationships are found to be consistent with the definition of the construct that the test is intended to measure. As indicated in the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (1999), the variables investigated can include other tests that measure the same constructs and different constructs, criterion measures that scores on the test are expected to predict, and demographic characteristics of examinees that are expected to be related and unrelated to test performance.

Relationship between CAHSEE and CST Results

Studies have been undertaken to examine the relationship between student scores on the CST and CAHSEE. Student performance on the CAHSEE and CST has been examined using special populations of students. The CDE's independent evaluator, HumRRO, has addressed the consistency of test results for students struggling to pass the CAHSEE and for SWDs. Results suggest the CST results were good indicators of how students performed on the CAHSEE, demonstrating consistency of test results across testing programs.

The most recent Independent Evaluation report examined the relationship between the seventh grade ELA and mathematics scores from the 2003 CST administration and success in passing the 2008 CAHSEE (Wise, 2009). Researchers used the average of students' ELA and mathematics CST scores to show the likelihood of passing the CAHSEE. Students who were near or above the median on the seventh grade tests had a very high likelihood of meeting the CAHSEE requirement; whereas, students who scored well below the median on the CST tests did not have a high likelihood of meeting the CAHSEE requirement. These results demonstrate that it is possible to identify students early on who may need additional help to pass the CAHSEE.

Additional research showing the relationship between the CAHSEE and CST results comes from the Independent Evaluation report of 2007. HumRRO investigated the performance of students who were repeat test-takers by examining the average CST scores for all students and for grade twelve students taking the CAHSEE. The latter group had mean CST scores between 0.75 and 1.00 standard deviations below the

mean scores for all students on the grade eleven ELA, Algebra I, Geometry, and Algebra II tests. The correlations between the CST and CAHSEE scores for CAHSEE repeat test-takers were in the low to moderate range (0.28 to 0.38). Generally, the 2006 CST end-of-course results proved to be a good predictor of 2007 CAHSEE results (Wise & Rui, 2007).

To examine the performance of SWDs, the 2007 CAHSEE results were compared with the 2006 CST results. Results indicate that the majority of students who took the grade nine or ten CST in 2006 scored in the Far Below Basic and Below Basic categories for ELA and had very little success in passing the CAHSEE. Students who scored in the higher CST performance categories (i.e., Basic, Proficient, Advanced) had an increased chance of passing the CAHSEE. Similar results were found for the grade ten SWDs in 2007 who had taken the general mathematics test and the Algebra I end-of-course tests in 2006. Approximately 80 percent of students scored in the Far Below Basic and Below Basic categories on the CST mathematics test in 2006 had low rates of passing the CAHSEE mathematics test in 2007, although they had more success in passing the CAHSEE if they were in the bottom two categories on the Algebra I test than did students in the same categories on the general mathematics test (Wise, 2007).

Differential Item Functioning

DIF statistics are used to identify those items that identifiable groups of students (e.g., females, African Americans, Hispanics) with the same underlying level of ability have different probabilities of answering the items correctly. If the item is differentially more difficult for an identifiable subgroup, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might be related to actual differences in relevant knowledge or skill (item impact) or statistical Type I error. As a result, DIF statistics are used to identify potential sources of item bias. Subsequent review by content experts and bias/sensitivity committees is required to determine the source and meaning of any differences that are seen.

For the CAHSEE, DIF comparison groups are based on gender (male compared to female), ethnicity (white compared to American Indian, Asian, Pacific Islander, Filipino, combined Asian, Hispanic, and African American), and English language proficiency (English proficient compared to EL).

Across all 2011–12 administrations few operational items were flagged for significant DIF per administration. Of the 28 items that were flagged, 20 were ELA items and 8 were mathematics items. All items were submitted for DIF panel review and determined to be valid measures of the intended construct. DIF analysis of the CAHSEE items is described earlier in this chapter (refer to *Differential Item Analysis* section). Details of the results of the DIF analysis for each administration can be found in Tables 5, 6, and 7 of Appendixes 6.B through 6.H.

Intercorrelations between Content Areas

To the degree that students' content area scores correlate as expected, evidence of the validity regarding those scores as measures of the intended constructs is provided.

The correlations between scores on the ELA and mathematics tests are presented in Appendix 6.K. Results appear to be consistent with expectations. In general, students' ELA scores correlate moderately with their mathematics scores (range of 0.61 to 0.73) for the non-census administrations and correlate higher for the census administrations where there is larger variance in the ability of students which leads to higher correlations (i.e., 0.81 and 0.79 for the February and March administrations, respectively).

In addition, intercorrelations between the strands and the total test scores are presented in the same tables. In general, moderate correlations between the test scores and strand scores of the same content area are expected since, by design, the strands measure various aspects of the same construct; whereas, intercorrelations between the test scores and strand scores of different content areas are expected to be less strong because the strands measure aspects of different constructs. The findings reflect these expectations. Using the February administration as an example, the mean intercorrelations between the content area scores and the strand scores were 0.85 and 0.89 for ELA and mathematics, respectively. The mean intercorrelation between the ELA score and mathematics strands was 0.73, and the mean intercorrelation between the mathematics score and the ELA strands was 0.68.

Generalizability Analyses for Writing Prompts

Generalizability analyses were performed on student responses to the ELA CR item to assess the proportion of variance explained by raters and persons. The details on the methodology are described in this chapter under the heading Generalizability Analyses. The results can be found in Appendix 6.M.

A decision study (d-study) was conducted to look at the generalizability-coefficients (g-coefficients) for the writing scores; the g-coefficients ranged from 0.77 to 0.86 across administrations. The largest variance component was attributed to the "person" variation, which is the desired variation to occur among the examinee or "person" scores. Variation attributable to the construct-irrelevant rater variable was negligible.

Evidence Based on Response Processes

As noted in the AERA, APA, and NCME's *Standards for Educational and Psychological Testing* (1999), additional support for a particular score interpretation or use can be provided by theoretical and empirical evidence indicating that examinees are using the intended response processes when responding to the items in a test. This evidence may be gathered from interacting with examinees in order to understand what processes underlie their item responses. Finally, evidence may also be derived from evidence provided by observers or judges involved in the scoring of examinee responses.

Inter-Rater Agreement

Rater consistency for the ELA writing prompt is critical to the CAHSEE writing scores and their interpretations. These findings provide evidence of the degree to which raters agree in their observations about the qualities evident in students' essay responses. As described in this chapter under *Writing Prompt and Rater Agreement*

Summary, two raters scored each examinee response. The raters demonstrated exact agreement for 71 to 83 percent of student papers across the administrations and demonstrated exact or adjacent agreement for 99 to 100 percent of the papers across the administrations. Details of the analyses are provided in Appendix 6.L in this chapter.

Evidence Based on Internal Structure

As suggested by the *Standards for Educational and Psychological Testing* (1999), evidence of validity can also be obtained from studies of the properties of the item scores and the relationship between these scores and scores on components of the test. To the extent that the score properties and relationships found are consistent with the definition of the construct measured by the test, support is gained for interpreting these scores as measures of the construct.

For the CAHSEE, it is assumed that a single construct underlies the total scores obtained on each test. Evidence to support this assumption can be gathered from the results of item analyses, evaluations of internal consistency, intercorrelations of strands, and test dimensionality.

With respect to the strands that are reported, these scores are intended to reflect the examinees' knowledge and/or skill in an area that is part of the construct underlying the total test. Analyses of the intercorrelations among the strands themselves and between the strands and total test score can be used for this purpose. Information about the internal consistency of the items on which each strand is based also is useful to provide. The relevant findings are described in the paragraphs that follow.

Classical and IRT Item Statistics

Point-biserial correlations calculated for the items in a test show the degree to which the items discriminate between students with low and high scores on a test. To the degree that the correlations are high, evidence that the items assess the same construct is provided. The distributions of point-biserial correlations for the items in the CAHSEE are presented in Appendices 6.B through 6.H, Table 1 for ELA and Table 2 for mathematics. This mean correlation ranged from 0.31 to 0.46 for ELA and from 0.27 to 0.46 for mathematics.

Also germane to the validity of a score interpretation are the ranges of difficulties for the items on which a test score will be based. The finding that items have difficulties that span the range of examinee ability provides evidence that examinees at all levels of ability are adequately measured by the items. Information on the distributions of item *b*-values is presented in Appendices 6.B through 6.H, Table 1 for ELA and Table 2 for mathematics. The data indicate that the tests had a range of item *b*-values.

Reliability of Test Scores

Reliability is a prerequisite for *validity*. The finding of reliability in student scores supports the validity of the inference that the scores reflect a stable construct. Findings concerning the reliabilities at the total-test level, as well as reliability results for the strands, were discussed in detail in the Reliability Analyses section in this

chapter. This section will summarize briefly evidence supporting the reliability of the CAHSEE test scores.

Overall reliability—The CAHSEE scores exhibit high reliability as evidenced by the mean internal consistency (coefficient α) values for the population of students. The mean reliability across the February and March census administrations was 0.94 for ELA and 0.95 for mathematics.

Subgroup reliability—High reliability of CAHSEE scores was observed for groups defined by gender, ethnicity, accommodation status, and English Language Proficiency status. Reliabilities and SEM estimates are reported for demographic groups for the February and March 2012 census administrations, where larger samples were available. For both administrations and content areas, results indicate similar reliability estimates within gender and ethnic groups. The reliabilities for the accommodation groups are slightly lower than those for the non-accommodation groups, and the reliabilities for the EL group are lower than those for the English-proficient group. It can be noted that a reduced range (i.e., a reduced standard deviation of scores) is normally associated with a reduction in reliability, and this likely occurred for some of the CAHSEE subgroups.

Strand reliability—The reliabilities of CAHSEE content strands invariably are lower than those for the total tests because they are based on very few items. Consistent with the findings of previous years, the strand reliabilities also are affected by the number of items in each strand with strand scores based on fewer items having somewhat lower reliabilities than strand scores based on more items. Because the reliabilities of scores at the strand level are lower, schools should supplement the score results with other information when interpreting the results.

Reliability of Performance Classification—The methodology used for estimating the reliability of classification decisions is described in the section Decision Classification Analyses in this chapter. Results for the reliability of classification reflect a high level of agreement in the student classifications, consistent with levels seen in previous years.

Intercorrelations of Strands

Evidence that strand scores have the intended meaning is provided by the finding that higher correlations are observed among scores obtained on strands designed to assess similar skills than are seen among scores obtained on strands designated to assess different skills. This is related to the ideas of convergent and divergent validity that Campbell and Fiske (1959) outlined.

Intercorrelations between CAHSEE ELA and mathematics raw scores at the strand level are provided for each administration in Appendix 6.K. For the census administrations, the ELA MC strands exhibit mean correlations of 0.72 (SD = 0.05) and 0.69 (SD = 0.06) for the February and March administrations, respectively. The mathematics strands exhibit mean correlations of 0.74 (SD = 0.06) and 0.74 (SD = 0.04) for the February and March administrations, respectively.

The mean intercorrelations between the ELA and mathematics strand scores are 0.63 (SD = 0.06) and 0.60 (SD = 0.06) for the February and March administrations,

respectively. The finding that the relationship is stronger between strands of the same content area, compared to the strands of different content areas, is consistent with the concept that ELA and mathematics measure different constructs, while the strands within a content area relate to the same construct.

Test Dimensionality

Gaffney and Perryman (2009) analyzed the factor structure for low stakes and high stakes tests to determine whether consequences at the student level affect the factor structure. Specifically, exploratory factor analysis was used to analyze the factor structure of the grade ten CST and the CAHSEE ELA component scores. Four ELA strand scores for each test were included in the analyses. For reading, the scores were for a) word analysis, fluency and systematic vocabulary development; and b) reading comprehension; and for writing, the scores were for a) writing strategies, and b) writing applications. Principal components extraction was followed by Promax rotation. Results indicated that the component scores for the low and high stakes subtests clearly loaded on separate factors in the two factor solution.

Additional Validity Evidence

In addition to the validation documentation gathered and maintained by the CDE, other empirical information in support of the CAHSEE is described below.

- Chapter 4 outlines the statistical criteria used to ensure that the test forms are comparable across administrations, providing additional evidence supporting inferences based on the comparability of the scores on different test forms.

- The section on IRT Analyses in the current chapter describes the calibration, scaling, and equating procedures used to place scores on the base scale developed in February 2004, thereby supporting the validity of inferences based on comparability of test scores.

- The current chapter also provides validity evidence supporting the quality of the test items with summary information for classical item analyses, DIF, and model-data fit for the operational and field-test items.

- Demographic summary information in Chapter 8 provides validity evidence supporting the quality of the test forms. The relative results for demographic subgroups for CAHSEE are, at a general level, consistent with results of other assessment programs, such as the Standardized Testing and Reporting (STAR) program. For example, the finding that socio-economic status is related to achievement is a common finding across testing programs.

Conclusions

Conclusions about the validity of CAHSEE for a particular use depend upon the definition of that use, but a wide variety of evidence is available for examining the validity of the CAHSEE testing program. As summarized in this chapter, this evidence covers the design of the content of the test, the alignment of the items to the state

standards, judgmental and statistical review of item quality, the accuracy of classification decisions based on this assessment, and the credibility of statistical analyses based on CAHSEE results.

Appendix 6.A: CAHSEE Item Review—Description and Examples of Classification Categories

The categories used by ETS statisticians to classify items for the CAHSEE, based on an evaluation of how well each item fits the Rasch model, are described below. The flagging scheme has categories of A, B, C, D, and F. Descriptors for each category are provided below. The IRT ICCs and empirical data (item-ability regressions) for six recently field-tested items are shown on the next page (Figure 6.3), starting from the upper-left corner. These six items illustrate the span of the rating categories. The item number in the calibration and the ETS identification number for each item (accession number) are listed next to one of the descriptions for the five possible categories provided below. This number can be used to identify the corresponding item-ability regression plot in the figure.

Flag A (Item 93, FM002619; Item 95, FM002640)

- Good fit of theoretical curve to empirical data along the entire ability range; may have some small divergence at the extremes.
- Small Chi-square value relative to the other items in the calibration with similar sample sizes.

Flag B (Item 96, FM002641)

- Good fit at the passing score.
- Theoretical curve within error range across most of ability range; may have some small divergence at the extremes.
- Acceptable Chi-square value relative to the other items in the calibration with similar sample sizes.

Flag C (Item 97, FM002767)

- Acceptable fit at the passing score.
- Theoretical curve within error range at some regions and slightly outside of error range at remaining regions of ability range.
- Moderate Chi-square value relative to the other items in the calibration with similar sample sizes.

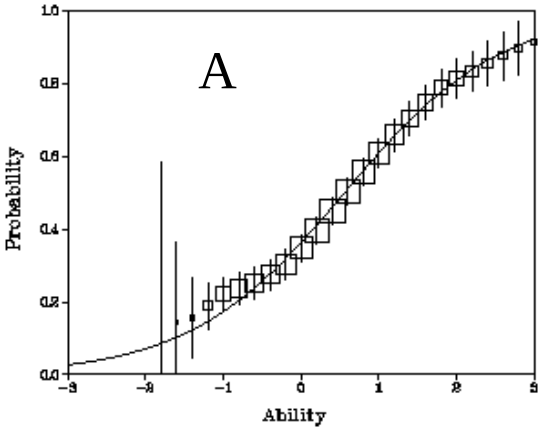
Flag D (Item 94, FM002620)

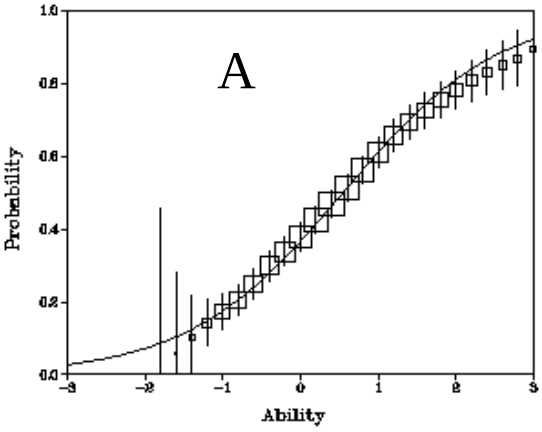
- Fit at the passing score may be slightly out of error range.
- Theoretical curve outside of error range at some regions across ability range.
- Empirical curve may have a zero slope at and around the passing score.
- Large Chi-square value relative to the other items in the calibration with similar sample sizes.
- Items receiving a D fit value also receive a Use Status of R. These items require additional examination and can be used if deemed appropriate.

Flag F (Item 98, FM002768)

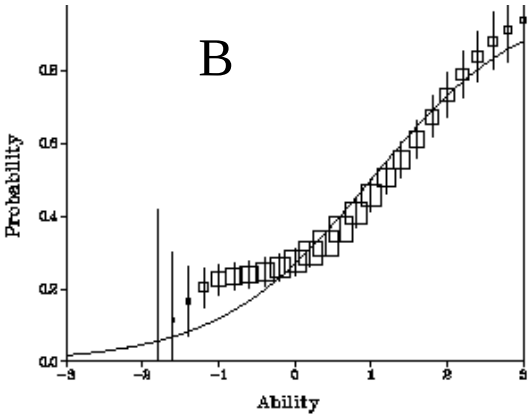
- Fit at the passing score outside error range.
- Theoretical curve outside of error range at most regions across ability range.
- Probability of answering item correctly may be greater at lower ability than higher ability (U shaped empirical curve).
- Very large Chi-square value (sometimes larger than three digits) relative to the other items with similar sample sizes.
- Items receiving an F fit value also receive a Use Status code of X, an indication the item should not be used in its current form and must be revised and field-tested before operational use.

Figure 6.3: Samples of CAHSEE Item-Fit Rating Categories

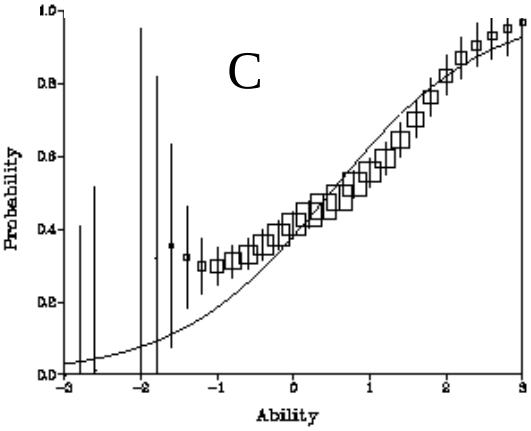




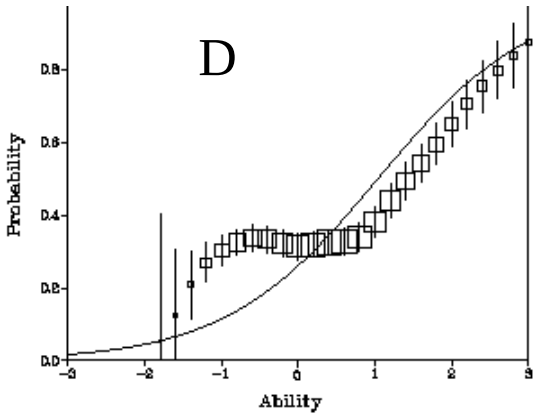
No.93 FM002619901 4 Choice P+=0.494
a=0.588 F, b=0.561, c=0.000 F, CHI=11.85

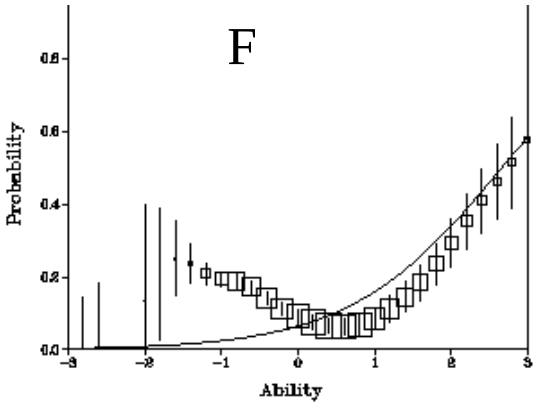


No.96 FM002641901 4 Choice P+=0.402
a=0.588 F, b=1.003, c=0.000 F, CHI=116.63



No.97 FM002767901 4 Choice P+=0.506
a=0.588 F, b=0.486, c=0.000 F, CHI=151.58





Appendix 6.B: Summary Statistics for Operational Items—July 2011

Table 6.B.1: Summary of Operational Item Statistics—ELA

Content area[1]	Number of items	IRT b value				Point-biserial/Pearson Correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	0.00	0.77	-2.16	1.39	0.31	0.07	0.14	0.42
RC	18	-0.19	0.67	-1.43	0.85	0.32	0.07	0.22	0.42
RL	20	-0.30	0.86	-2.16	0.75	0.30	0.06	0.19	0.39
RW	7	0.03	0.86	-0.80	1.20	0.32	0.06	0.21	0.40
WA	1	0.68	-	0.68	0.68	0.62	-	0.62	0.62
WC	15	0.27	0.80	-0.90	1.39	0.28	0.08	0.14	0.37
WS	12	0.38	0.40	-0.58	0.91	0.31	0.07	0.18	0.42

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.B.2: Summary of Operational Item Statistics—Mathematics

Content area[1]	Number of items	IRT b value				Point-biserial Correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.24	0.61	-1.93	1.06	0.27	0.08	0.09	0.47
A1	12	0.19	0.33	-0.29	0.76	0.22	0.08	0.09	0.35
AF	20	-0.31	0.61	-1.93	0.49	0.31	0.07	0.20	0.41
MG	18	-0.08	0.70	-1.48	1.06	0.27	0.10	0.13	0.47
MR	8	0.13	0.41	-0.48	0.60	0.24	0.08	0.11	0.35
NS	17	-0.44	0.46	-1.18	0.33	0.26	0.06	0.16	0.36
PS	13	-0.51	0.62	-1.19	0.72	0.30	0.09	0.11	0.39

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.B.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	36	49
B	Use	25	34
C	Use	12	16
Total		73	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.B.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	39	49
B	Use	21	26
C	Use	20	25
Total		80	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.B.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE045229	L630J007	0	2	A	S	S	S	S	S	A	A	C-
ELA	VE044826	L616H003	0	38	A	S	S	S	S	S	B+	C+	A
ELA	VC392671	L5OSA620	0	72	B-	S	S	S	S	S	B-	C-	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring. S indicates that DIF was not performed due to insufficient sample size.

Table 6.B.6: Distribution of Operational Item DIF Classifications—ELA

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C DIF Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	2	3
B-	3	4	0	0	0	0	0	0	0	0	0	0	1	1	3	4	3	4		
A	69	95	0	0	0	0	0	0	0	0	0	0	70	96	67	92	67	92		
B+	1	1	0	0	0	0	0	0	0	0	0	0	2	3	1	1	2	3		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1
Small N[3]	0	0	73	100	73	100	73	100	73	100	73	100	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	3	4

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.B.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C DIF Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	3	4	0	0	0	0	0	0	0	0	0	0	2	3	2	3	0	0		
A	75	94	0	0	0	0	0	0	0	0	0	0	78	98	77	96	78	98		
B+	2	3	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	3		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	80	100	80	100	80	100	80	100	80	100	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	0	0

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.B.8: Listing of CR Item Statistics—ELA

Accession number	VE025116
CAHSEE ID	L6OSA1410
Polyserial correlation	0.68
IRT b-value	0.68332
Step category 1[1]	1.46710
Step category 2	1.38703
Step category 3	3.53953
Step category 4	-1.01293
Step category 5	-0.79612
Step category 6	-2.48664
Step category 7	-2.09796
DIF category, Male-Female	A
DIF category, White-American Indian	S[3]
DIF category, White-Asian	S
DIF category, White-Pacific Islander	S
DIF category, White-Filipino	S
DIF category, White-Combined Asian	S
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

[3]S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.C: Summary Statistics for Operational Items—October 2011

Table 6.C.1: Summary of Operational Item Statistics—ELA

Content area[1]	Number of items	IRT b value				Point-biserial/Pearson correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	-0.04	0.62	-1.42	1.50	0.36	0.07	0.22	0.48
RC	18	-0.22	0.58	-1.23	1.19	0.37	0.06	0.28	0.48
RL	20	-0.21	0.63	-1.42	1.27	0.38	0.07	0.24	0.47
RW	7	-0.01	0.63	-1.03	0.68	0.34	0.07	0.23	0.43
WA	1	0.69	-	0.69	0.69	0.66	-	0.66	0.66
WC	15	0.00	0.60	-1.11	0.99	0.33	0.07	0.22	0.44
WS	12	0.38	0.50	-0.58	1.50	0.35	0.07	0.23	0.46

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.C.2: Summary of Operational Item Statistics—Mathematics

Content area[1]	Number of items	IRT b value				Point-biserial correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.19	0.61	-1.82	1.21	0.34	0.08	0.13	0.50
A1	12	0.25	0.40	-0.19	1.15	0.29	0.08	0.15	0.41
AF	20	-0.31	0.58	-1.33	0.85	0.37	0.08	0.13	0.50
MG	18	-0.03	0.52	-1.07	0.88	0.34	0.08	0.15	0.45
MR	8	-0.28	0.64	-0.92	1.21	0.36	0.11	0.20	0.45
NS	17	-0.36	0.75	-1.82	1.21	0.31	0.06	0.20	0.44
PS	13	-0.42	0.52	-1.08	0.51	0.35	0.06	0.24	0.46

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.C.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	15	21
B	Use	16	22
C	Use	42	58
Total		73	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.C.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	14	18
B	Use	28	35
C	Use	34	43
D	Review	4	5
Total		80	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.C.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VC392117	L563G005	0	3	A	A	C-	S	S	C-	A	A	B-
ELA	VE044532	L603J001	0	6	B-	A	B-	S	S	C-	C-	C-	A
ELA	VE338869	L794J014	0	33	A	A	C-	S	S	B-	A	A	A
ELA	VE046175	L687H010	0	45	A	A	B+	S	S	C+	A	A	A
ELA	VE046481	L6OSA1188	0	65	A	A	C+	S	S	A	A	A	A
ELA	VE339067	L7OSA1603	0	72	A	A	B-	S	S	B-	A	C-	A
ELA	VC138573	L3OSA425	0	75	A	A	C-	S	S	B-	A	A	A
MATH	VC023358	M14115	0	15	C-	A	A	S	B-	A	A	A	A
MATH	VE339532	M60492	0	24	A	A	C-	S	A	B-	A	A	A
MATH	VC024755	M21686	0	37	A	A	A	S	C+	A	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring. S indicates that DIF was not performed due to insufficient sample size.

Table 6.C.6: Distribution of Operational Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient English Learner		Total C DIF Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	3	4	0	0	0	0	2	3	1	1	2	3	0	0	5	7
B-	1	1	0	0	6	8	0	0	0	0	8	11	1	1	2	3	2	3		
A	72	99	73	100	58	79	0	0	0	0	61	84	71	97	69	95	71	97		
B+	0	0	0	0	5	7	0	0	0	0	1	1	0	0	0	0	0	0		
C+[2]	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	2	3
Small N[3]	0	0	0	0	0	0	73	100	73	100	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	7	10

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentages may differ slightly from the sum of the category percentages due to rounding.

Table 6.C.7: Distribution of Operational Item DIF Classifications—Mathematics

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient English Learner		Total C DIF Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2	3
B-	3	4	0	0	4	5	0	0	9	11	3	4	0	0	0	0	1	1		
A	74	93	80	100	74	93	0	0	63	79	76	95	80	100	80	100	79	99		
B+	2	3	0	0	1	1	0	0	7	9	1	1	0	0	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1
Small N[3]	0	0	0	0	0	0	80	100	0	0	0	0	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	3	4

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.C.8: Listing of CR Item Statistics—ELA

Accession number	VE025112
CAHSEE ID	L6OSA1406
Polyserial correlation	0.72
IRT b-value	0.69495
Step category 1[1]	1.73000
Step category 2	1.44832
Step category 3	3.68268
Step category 4	-0.92904
Step category 5	-0.82689
Step category 6	-2.57011
Step category 7	-2.53497
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	S[3]
DIF category, White-Filipino	S
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group

[3] S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.D: Summary Statistics for Operational Items—November 2011

Table 6.D.1: Summary of Operational Item Statistics—ELA

Content area[1]	Number of items	IRT b value				Point-biserial/Pearson correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	-0.07	0.65	-1.83	1.28	0.37	0.06	0.23	0.49
RC	18	-0.04	0.52	-0.84	0.89	0.38	0.07	0.29	0.48
RL	20	-0.30	0.72	-1.83	0.73	0.35	0.06	0.23	0.45
RW	7	-0.57	0.45	-1.00	0.25	0.40	0.06	0.31	0.47
WA	1	0.92	-	0.92	0.92	0.68	-	0.68	0.68
WC	15	0.03	0.71	-0.83	1.28	0.38	0.05	0.27	0.46
WS	12	0.35	0.34	-0.40	0.84	0.39	0.07	0.30	0.49

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.D.2: Summary of Operational Item Statistics—Mathematics

Content area[1]	Number of items	IRT b value				Point-biserial correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.24	0.61	-1.80	1.61	0.34	0.08	0.12	0.51
A1	12	0.25	0.48	-0.62	1.14	0.30	0.09	0.12	0.41
AF	20	-0.32	0.35	-1.03	0.14	0.38	0.06	0.28	0.46
MG	18	-0.10	0.83	-1.80	1.61	0.32	0.09	0.18	0.51
MR	8	-0.24	0.38	-0.92	0.25	0.36	0.07	0.22	0.45
NS	17	-0.37	0.48	-1.16	0.52	0.32	0.06	0.22	0.38
PS	13	-0.59	0.56	-1.22	0.45	0.37	0.08	0.22	0.47

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.D.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	26	36
B	Use	44	60
C	Use	3	4
Total		73	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.D.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	30	38
B	Use	34	43
C	Use	15	19
D	Review	1	1
Total		80	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.D.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE338647	L773J004	0	1	A	A	C-	B-	B-	B-	A	A	A
ELA	VE338655	L773J014	0	3	A	A	B-	C-	B-	B-	A	A	A
ELA	VC019864	L269E013	0	27	A	A	B-	C-	B-	B-	A	A	A
MATH	FM004402	M10416	0	65	A	A	B-	A	C-	B-	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.

Table 6.D.6: Distribution of Operational Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient English Learner		Total C Items Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	1	1	2	3	0	0	0	0	0	0	0	0	0	0	3	4
B-	1	1	0	0	8	11	2	3	5	7	4	5	0	0	1	1	3	4		
A	72	99	73	100	61	84	68	93	68	93	68	93	73	100	72	99	69	95		
B+	0	0	0	0	3	4	1	1	0	0	1	1	0	0	0	0	1	1		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	3	4

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.D.7: Distribution of Operational Item DIF Classifications—Mathematics

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient English Learner		Total C Items Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1
B-	2	3	0	0	2	3	2	3	2	3	3	4	0	0	0	0	3	4		
A	78	98	80	100	77	96	78	98	74	93	76	95	80	100	80	100	77	96		
B+	0	0	0	0	1	1	0	0	3	4	1	1	0	0	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	1	1

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.D.8: Listing of CR Item Statistics—ELA

Accession number	VE025138
CAHSEE ID	L6OSA1444
Polyserial correlation	0.74
IRT b-value	0.91759
Step category 1[1]	2.28232
Step category 2	1.47211
Step category 3	3.22897
Step category 4	-0.89427
Step category 5	-0.92991
Step category 6	-2.51280
Step category 7	-2.64641
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	A
DIF category, White-Filipino	A
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

Appendix 6.E: Summary Statistics for Operational Items—December 2011

Table 6.E.1: Summary of Operational Item Statistics—ELA

Content area[1]	Number of items	IRT b value				Point-biserial/Pearson correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	-0.03	0.66	-1.78	1.42	0.31	0.08	0.12	0.49
RC	18	-0.12	0.61	-1.49	0.96	0.34	0.08	0.20	0.49
RL	20	-0.24	0.67	-1.78	1.19	0.33	0.08	0.15	0.44
RW	7	-0.04	0.65	-0.85	0.61	0.29	0.06	0.19	0.37
WA	1	0.59	.	0.59	0.59	0.61	.	0.61	0.61
WC	15	-0.05	0.76	-1.06	1.42	0.30	0.09	0.12	0.43
WS	12	0.43	0.41	-0.33	1.12	0.28	0.08	0.15	0.40

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.E.2: Summary of Operational Item Statistics—Mathematics

Content area[1]	Number of items	IRT b value				Point-biserial correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.15	0.67	-2.32	1.32	0.28	0.08	0.08	0.43
A1	12	0.32	0.39	-0.25	0.87	0.19	0.07	0.08	0.30
AF	20	-0.36	0.52	-1.06	0.61	0.32	0.07	0.18	0.41
MG	18	-0.05	0.83	-2.15	1.32	0.27	0.09	0.09	0.43
MR	8	-0.32	0.39	-0.74	0.32	0.29	0.07	0.17	0.38
NS	17	-0.12	0.57	-1.19	0.81	0.29	0.08	0.12	0.41
PS	13	-0.43	0.75	-2.32	0.59	0.28	0.07	0.16	0.36

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.E.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	22	30
B	Use	32	44
C	Use	18	25
D	Review	1	1
Total		73	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.E.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	34	43
B	Use	28	35
C	Use	18	23
Total		80	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.E.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
MATH	VE047456	M50353	0	14	C-	S	S	S	S	S	S	S	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring. S indicates that DIF was not performed due to insufficient sample size.

Table 6.E.6: Distribution of Operational Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	10		
A	67	92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	86		
B+	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	73	100	73	100	73	100	73	100	73	100	73	100	73	100	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	0	0

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.E.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
B-	5	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4		
A	72	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	74	93		
B+	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	4		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	80	100	80	100	80	100	80	100	80	100	80	100	80	100	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	1	1

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.E.8: Listing of CR Item Statistics—ELA

Accession number	VE025129
CAHSEE ID	L6OSA1431
Polyserial correlation	0.68
IRT b-value	0.59014
Step category 1[1]	5.60739
Step category 2	0.97470
Step category 3	3.64746
Step category 4	-1.49131
Step category 5	-1.05040
Step category 6	-3.28758
Step category 7	-4.40026
DIF category, Male-Female	B+
DIF category, White-American Indian	S[3]
DIF category, White-Asian	S
DIF category, White-Pacific Islander	S
DIF category, White-Filipino	S
DIF category, White-Combined Asian	S
DIF category, White-Hispanic	S
DIF category, White-African American	S
Least favorable DIF category among all focal groups[2]	B+

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

[3] S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.F: Summary Statistics for Operational Items—February 2012

Table 6.F.1: Summary of Operational Item Statistics—ELA

		IRT b value				Point-biserial/Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	73	0.00	0.69	-1.35	1.50	0.46	0.07	0.30	0.62
RC	18	-0.03	0.59	-1.07	0.94	0.48	0.07	0.37	0.62
RL	20	-0.21	0.75	-1.35	1.50	0.44	0.07	0.31	0.53
RW	7	-0.41	0.66	-1.12	0.91	0.46	0.07	0.36	0.54
WA	1	0.93	-	0.93	0.93	0.73	-	0.73	0.73
WC	15	0.16	0.57	-0.71	1.07	0.45	0.10	0.30	0.61
WS	12	0.35	0.74	-0.77	1.50	0.46	0.05	0.35	0.53

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.F.2: Summary of Operational Item Statistics—Mathematics

		IRT b value				Point-biserial correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	80	-0.24	0.67	-2.04	1.52	0.46	0.08	0.23	0.61
A1	12	0.19	0.69	-1.40	1.11	0.48	0.06	0.38	0.55
AF	20	-0.35	0.54	-1.09	0.59	0.48	0.06	0.36	0.57
MG	18	-0.09	0.59	-0.80	1.52	0.49	0.07	0.33	0.61
MR	8	-0.21	0.28	-0.68	0.16	0.46	0.07	0.36	0.55
NS	17	-0.29	0.70	-2.04	0.64	0.45	0.08	0.23	0.56
PS	13	-0.61	0.73	-1.91	0.60	0.41	0.09	0.24	0.55

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.F.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	23	32
B	Use	30	41
C	Use	19	26
Total		73	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.F.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	25	31
B	Use	29	36
C	Use	26	33
Total		80	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.F.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	FM004202	L132B001	0	28	C-	A	B-	A	A	B-	A	A	B-
ELA	VE044898	L618H014	0	42	A	A	B-	B-	C-	B-	A	A	A
ELA	VE337919	L720K001	0	56	A	A	B-	A	C-	B-	A	A	A
ELA	VE339072	L7OSA1608	0	67	A	A	C-	B-	C-	C-	A	A	A
MATH	FM007036	M24029	0	31	A	A	B-	B-	C-	B-	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.

Table 6.F.6: Distribution of Operational Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	1	1	0	0	1	1	0	0	3	4	1	1	0	0	0	0	0	0	4	5
B-	2	3	0	0	5	7	3	4	2	3	5	7	0	0	0	0	3	4		
A	69	95	73	100	65	89	68	93	64	88	66	90	73	100	73	100	69	95		
B+	1	1	0	0	2	3	2	3	4	5	1	1	0	0	0	0	1	1		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	4	5

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.F.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	1
B-	1	1	0	0	3	4	2	3	3	4	3	4	0	0	1	1	0	0		
A	79	99	80	100	76	95	76	95	75	94	76	95	79	99	79	99	80	100		
B+	0	0	0	0	1	1	2	3	1	1	1	1	1	1	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	1	1

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.F.8: Listing of CR Item Statistics—ELA

Accession number	VE025126
CAHSEE ID	L6OSA1425
Polyserial correlation	0.81
IRT b-value	0.92875
Step category 1[1]	2.20762
Step category 2	1.47811
Step category 3	3.47785
Step category 4	-0.95795
Step category 5	-0.92785
Step category 6	-2.62766
Step category 7	-2.65012
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	B+
DIF category, White-Filipino	B+
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	B+

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

Appendix 6.G: Summary Statistics for Operational Items—March 2012

Table 6.G.1: Summary of Operational Item Statistics—ELA

		IRT b value				Point-biserial/Pearson Correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	73	-0.15	0.81	-1.95	1.63	0.44	0.07	0.27	0.56
RC	18	-0.07	0.80	-1.54	1.63	0.45	0.08	0.32	0.56
RL	20	-0.55	0.94	-1.95	1.62	0.43	0.05	0.31	0.49
RW	7	-0.35	0.54	-1.06	0.54	0.42	0.07	0.33	0.50
WA	1	0.93	-	0.93	0.93	0.71	-	0.71	0.71
WC	15	-0.07	0.77	-1.92	0.77	0.44	0.06	0.31	0.52
WS	12	0.37	0.42	-0.49	0.97	0.44	0.08	0.27	0.54

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.G.2: Summary of Operational Item Statistics—Mathematics

		IRT b value				Point-biserial Correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	80	-0.20	0.70	-2.07	1.34	0.45	0.08	0.26	0.61
A1	12	0.22	0.42	-0.49	1.02	0.47	0.07	0.33	0.59
AF	20	-0.36	0.52	-1.26	1.02	0.44	0.07	0.28	0.55
MG	18	-0.15	0.89	-1.67	1.34	0.49	0.06	0.34	0.60
MR	8	-0.01	0.88	-1.51	1.02	0.39	0.06	0.28	0.49
NS	17	-0.20	0.80	-2.07	0.65	0.42	0.11	0.27	0.61
PS	13	-0.41	0.67	-1.15	0.88	0.42	0.08	0.26	0.56

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.G.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	33	45
B	Use	22	30
C	Use	18	25
Total		73	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.G.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	37	46
B	Use	16	20
C	Use	27	34
Total		80	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.G.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VC392370	L581G013	0	38	B-	A	C-	C-	C-	C-	C-	B-	B-
ELA	VC392371	L581G014	0	39	A	A	C-	B-	C-	C-	C-	A	A
MATH	VC151297	M30735	0	22	A	A	C-	B-	B-	B-	B-	B-	A
MATH	VC024009	M20742	0	89	C-	A	A	A	A	A	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.

Table 6.G.6: Distribution of Operational Item DIF Classifications—ELA

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	2	3	1	1	2	3	2	3	2	3	0	0	0	0	2	3
B-	1	1	0	0	2	3	1	1	3	4	1	1	0	0	2	3	1	1		
A	72	99	73	100	66	90	71	97	67	92	68	93	70	96	71	97	70	96		
B+	0	0	0	0	3	4	0	0	1	1	2	3	1	1	0	0	2	3		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	2	3

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentages may differ slightly from the sum of the category percentages due to rounding.

Table 6.G.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2	3
B-	5	6	0	0	4	5	2	3	3	4	4	5	1	1	1	1	2	3		
A	73	91	80	100	75	94	76	95	77	96	76	95	77	96	79	99	78	98		
B+	1	1	0	0	0	0	2	3	0	0	0	0	2	3	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	2	3

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentages may differ slightly from the sum of the category percentages due to rounding.

Table 6.G.8: Listing of CR Item Statistics—ELA

Accession number	VE025125
CAHSEE ID	L6OSA1423
Polyserial correlation	0.83
IRT b-value	0.92595
Step category 1[1]	2.35804
Step category 2	1.55477
Step category 3	3.86798
Step category 4	-1.09706
Step category 5	-1.02552
Step category 6	-2.8696
Step category 7	-2.78862
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	A
DIF category, White-Filipino	A
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

Appendix 6.H: Summary Statistics for Operational Items—May 2012

Table 6.H.1: Summary of Operational Item Statistics—ELA

		IRT b value				Point-biserial/Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	73	-0.10	0.55	-1.64	1.02	0.40	0.06	0.25	0.52
RC	18	-0.11	0.49	-1.10	0.88	0.41	0.04	0.32	0.49
RL	20	-0.26	0.65	-1.64	0.84	0.38	0.05	0.29	0.47
RW	7	-0.48	0.43	-1.14	0.21	0.40	0.09	0.25	0.50
WA	1	0.64	.	0.64	0.64	0.69	.	0.69	0.69
WC	15	0.02	0.38	-0.63	0.65	0.39	0.08	0.27	0.52
WS	12	0.18	0.55	-0.66	1.02	0.40	0.06	0.27	0.49

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.H.2: Summary of Operational Item Statistics—Mathematics

		IRT b value				Point-biserial correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	80	-0.22	0.60	-2.14	1.03	0.37	0.08	0.18	0.56
A1	12	0.41	0.41	-0.41	0.97	0.33	0.07	0.18	0.41
AF	20	-0.36	0.47	-1.48	0.29	0.40	0.07	0.22	0.49
MG	18	-0.15	0.49	-1.32	1.03	0.37	0.10	0.19	0.56
MR	8	-0.46	0.78	-2.14	0.26	0.34	0.07	0.25	0.44
NS	17	-0.30	0.74	-2.14	0.58	0.38	0.06	0.25	0.46
PS	13	-0.56	0.46	-1.22	0.43	0.38	0.06	0.30	0.46

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.H.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	15	21
B	Use	23	32
C	Use	35	48
Total		73	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.H.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	15	19
B	Use	21	26
C	Use	42	53
D	Review	2	3
Total		80	100

[1]Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.H.5: Operational Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE339097	L7OSA1646	0	77	A	A	C-	S	S	C-	A	A	A

Note: Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring. S indicates that DIF was not performed due to insufficient sample size.

Table 6.H.6: Distribution of Operational Item DIF Classifications—ELA

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	1	1
B-	0	0	1	1	7	10	0	0	0	0	0	0	0	0	0	0	1	1		
A	73	100	72	99	62	85	0	0	0	0	72	99	73	100	73	100	72	99		
B+	0	0	0	0	3	4	0	0	0	0	0	0	0	0	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	73	100	73	100	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	1	1

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.H.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	2	3	0	0	4	5	0	0	4	5	2	3	1	1	3	4	0	0		
A	78	98	80	100	75	94	0	0	73	91	77	96	79	99	77	96	80	100		
B+	0	0	0	0	1	1	0	0	3	4	1	1	0	0	0	0	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	80	100	0	0	0	0	0	0	0	0	0	0		
Total[4]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	0	0

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.H.8: Listing of CR Item Statistics—ELA

Accession number	VE025117
CAHSEE ID	L6OSA1411
Polyserial correlation	0.76
IRT b-value	0.63632
Step category 1[1]	1.73708
Step category 2	1.36273
Step category 3	3.44115
Step category 4	-0.85235
Step category 5	-0.72539
Step category 6	-2.57006
Step category 7	-2.39316
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	S[3]
DIF category, White-Filipino	S
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.

[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

[3] S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.I: Summary Statistics for Field-Test Items—February 2012

Table 6.I.1: Summary of Field-Test Item Statistics—ELA

		IRT b value				Point-biserial/Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	658	0.73	0.84	-1.73	3.20	0.38	0.12	0.08	0.63
RC	175	0.78	0.76	-0.99	2.78	0.38	0.11	0.08	0.61
RL	228	0.71	0.76	-1.29	3.20	0.39	0.12	0.09	0.59
RW	95	0.72	0.94	-0.96	3.16	0.39	0.12	0.08	0.58
WC	97	0.63	1.11	-1.73	3.20	0.36	0.12	0.11	0.63
WS	63	0.87	0.75	-0.92	2.41	0.37	0.12	0.12	0.55

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WA = Writing Applications, WC = Writing Conventions, WS = Writing Strategies

Table 6.I.2: Summary of Field-Test Item Statistics—Mathematics

		IRT b value				Point-biserial correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	566	0.58	0.87	-1.73	3.66	0.39	0.11	0.07	0.63
A1	102	0.96	0.59	-0.57	2.71	0.37	0.12	0.09	0.62
AF	129	0.38	0.80	-1.61	2.64	0.42	0.10	0.09	0.57
MG	131	0.90	0.99	-1.39	3.66	0.37	0.12	0.07	0.58
MR	56	0.72	0.91	-1.23	2.72	0.33	0.12	0.09	0.57
NS	113	0.36	0.82	-1.73	2.17	0.43	0.10	0.17	0.63
PS	91	0.29	0.84	-1.38	2.72	0.35	0.10	0.09	0.56

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.I.3: IRT Model Data Fit Distribution of Field-Test Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	90	13
B	Use	140	20
C	Use	274	39
D	Review	53	8
F	Do Not Use	101	15
N/A	Dropped from Calibration	36	5
Total		694	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.I.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	50	8
B	Use	101	16
C	Use	233	37
D	Review	63	10
F	Do Not Use	119	19
N/A	Dropped from Calibration	58	9
Total		624	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.I.5: Field-Test Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE579505	L826L001	7	46	A	S	S	S	S	C-	C-	B-	A
ELA	VE579506	L826L002	9	47	A	S	S	S	S	S	A	S	C-
ELA	VE359260	L882K006	11	51	A	S	S	S	S	S	A	S	C-
ELA	VE359196	L873K001	15	46	A	S	S	S	S	S	C-	S	A
ELA	VE359308	L894K012	20	47	A	S	S	S	S	S	C-	S	A
ELA	VE359313	L899K001	21	46	C-	S	S	S	S	S	A	S	B-
ELA	VE359318	L899K006	21	48	C-	S	S	S	S	S	A	S	A
ELA	VE359322	L899K010	22	49	C+	S	S	S	S	S	A	S	A
ELA	VE579682	L842L007	23	47	A	S	S	S	S	S	B-	S	C-
ELA	VE579680	L842L005	25	48	A	S	S	S	S	S	A	S	C-
ELA	VE645206	L890L001	35	46	A	S	S	S	S	S	A	S	C-
ELA	VE359240	L879K001	44	46	A	S	S	S	S	S	A	S	C-
ELA	VE359247	L879K008	44	49	C+	S	S	S	S	S	A	S	A
ELA	VE645129	L885L001	46	46	A	S	S	S	S	S	C-	S	A
ELA	VE580141	L897K002	49	46	A	S	S	S	S	S	A	S	C-
ELA	VE580100	L893K001	55	46	C-	S	S	S	S	S	B-	S	A
ELA	VE359033	L7OSA1811	63	74	A	S	S	S	S	S	A	S	C-
ELA	VE579280	L807L007	68	46	A	S	S	S	S	S	C-	S	A
ELA	VE359026	L7OSA1804	69	74	A	S	S	S	S	S	B-	S	C-
ELA	VE359024	L7OSA1802	71	74	A	S	S	S	S	S	C-	S	A
ELA	VE580056	L890K004	73	47	B-	S	S	S	S	S	C-	S	A
ELA	VE359020	L7OSA1798	73	74	A	S	S	S	S	S	C-	S	B-
ELA	VE579245	L803L002	74	46	B-	S	S	S	S	S	A	S	C-
ELA	VE579257	L803L014	74	49	C+	S	S	S	S	S	A	S	A
ELA	VE580217	L8OSA1947	74	74	C+	S	S	S	S	S	A	S	A
ELA	VE359018	L7OSA1796	75	74	A	S	S	S	S	S	A	S	C+
ELA	VE358957	L7OSA1666	83	74	B-	S	S	S	S	S	C-	S	A
ELA	VE579334	L811L009	85	48	C-	S	S	S	S	S	B-	S	A
ELA	VE046993	L6OSA979	88	74	B+	S	S	S	S	S	A	S	C-
ELA	VE046987	L6OSA973	94	74	B-	S	S	S	S	S	C-	S	A
ELA	VE645391	L8OSA2088	95	74	B+	S	S	S	S	S	A	S	C-
ELA	VE580088	L892K004	96	47	A	S	S	S	S	S	C-	S	B-
ELA	VE046976	L6OSA962	98	74	A	S	S	S	S	S	A	S	C+

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE645390	L8OSA2087	102	74	A	S	S	S	S	S	B-	S	C-
ELA	VE579424	L820L001	103	46	B-	S	S	S	S	S	C-	S	B-
MATH	VE580835	M62027	2	18	C-	S	S	S	S	A	A	B-	A
MATH	VE580422	M61270	2	79	C-	S	S	S	S	B-	A	A	A
MATH	VE580725	M61894	5	21	A	S	A	S	S	A	A	C-	A
MATH	VE580367	M61127	7	35	A	S	S	S	S	C-	A	A	A
MATH	VE645773	M62136	8	18	A	S	S	S	S	A	A	A	C-
MATH	VE580844	M62036	13	20	C-	S	S	S	S	B-	B-	C-	A
MATH	VE580488	M61463	13	60	B-	S	S	S	S	A	A	C-	A
MATH	VE645593	M61439	16	35	C-	S	S	S	S	B+	A	A	A
MATH	VE580830	M62022	17	18	A	S	S	S	S	C+	A	A	A
MATH	VE580836	M62028	19	20	C-	S	S	S	S	A	A	A	A
MATH	VE359336	M40118	21	36	A	S	C-	S	S	B-	A	A	A
MATH	VE645886	M62285	30	59	A	S	B-	S	S	C-	A	A	A
MATH	VE359787	M61169	31	18	C-	S	A	S	S	A	B-	A	A
MATH	VC173026	M30882	32	21	B+	S	A	S	S	A	A	A	C-
MATH	VE359788	M61170	33	19	C-	S	A	S	S	A	A	A	A
MATH	VE580775	M61958	33	20	A	S	B+	S	S	C+	A	A	A
MATH	VE580383	M61171	44	19	C-	S	A	S	S	A	A	B-	A
MATH	VE580467	M61425	49	21	A	S	S	S	S	A	A	A	C-

Table 6.I.6: Distribution of Field-Test Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	4	1	0	0	0	0	0	0	0	0	1	0	12	2	0	0	13	2	29	4
B-	20	3	0	0	2	0	0	0	0	0	2	0	46	7	2	0	52	7		
A	636	92	0	0	13	2	0	0	0	0	30	4	624	90	31	4	624	90		
B+	30	4	0	0	0	0	0	0	0	0	0	0	12	2	0	0	3	0		
C+[2]	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	6	1
Small N[3]	0	0	694	100	679	98	694	100	694	100	661	95	0	0	661	95	0	0		
Total[4]	694	100	694	100	694	100	694	100	694	100	694	100	694	100	694	100	694	100	35	5

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentages may differ slightly from the sum of the category percentages due to rounding.

Table 6.I.7: Distribution of Field-Test Item DIF Classifications—Mathematics

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyse[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	8	1	0	0	1	0	0	0	0	0	2	0	0	0	3	0	3	0	16	3
B-	23	4	0	0	14	2	0	0	0	0	24	4	15	2	29	5	21	3		
A	583	93	0	0	228	37	0	0	0	0	549	88	598	96	574	92	598	96		
B+	10	2	0	0	21	3	0	0	0	0	47	8	11	2	18	3	2	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2	0
Small N[3]	0	0	624	100	360	58	624	100	624	100	0	0	0	0	0	0	0	0		
Total[4]	624	100	624	100	624	100	624	100	624	100	624	100	624	100	624	100	624	100	18	3

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
[3] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[4] Total percentages may differ slightly from the sum of the category percentages due to rounding.

Appendix 6.J: Summary Statistics for Field-Test Items—March 2012

Table 6.J.1: Summary of Field-Test Item Statistics—ELA

		IRT b value				Point-biserial /Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	1002	0.77	0.91	-2.01	4.48	0.36	0.11	0.07	0.59
RC	268	0.80	0.77	-1.29	2.97	0.37	0.11	0.08	0.58
RL	314	0.90	0.85	-1.08	3.56	0.36	0.11	0.09	0.58
RW	112	0.59	0.88	-1.22	2.68	0.37	0.10	0.14	0.59
WC	152	0.35	1.20	-2.01	4.48	0.35	0.11	0.07	0.58
WS	156	0.98	0.81	-1.40	3.05	0.36	0.11	0.11	0.56

[1] RC = Reading Comprehension, RL = Literary Response & Analysis, RW = Word Analysis, WC = Writing Conventions, WS = Writing Strategies

Table 6.J.2: Summary of Field-Test Item Statistics—Mathematics

		IRT b value				Point-biserial/Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	914	0.78	0.94	-1.75	3.65	0.39	0.11	0.07	0.65
A1	150	1.16	0.81	-0.95	3.57	0.37	0.12	0.09	0.61
AF	210	0.67	0.97	-1.75	3.65	0.38	0.12	0.10	0.65
MG	243	0.92	0.94	-1.57	3.59	0.39	0.12	0.08	0.61
MR	74	1.11	1.00	-1.25	3.65	0.32	0.11	0.10	0.53
NS	154	0.70	0.89	-1.42	3.32	0.41	0.11	0.08	0.63
PS	157	0.44	0.92	-1.69	2.96	0.41	0.10	0.07	0.56

[1] A1 = Algebra I, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS = Number Sense, PS = Probability & Statistics

Table 6.J.3: IRT Model Data Fit Distribution of Field-Test Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	146	14
B	Use	265	25
C	Use	397	38
D	Review	89	8
F	Do Not Use	105	10
N/A	Dropped from Calibration	47	4
Total		1049	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.J.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	140	15
B	Use	196	21
C	Use	432	46
D	Review	34	4
F	Do Not Use	112	12
N/A	Dropped from Calibration	22	2
Total		936	100

[1] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.J.5: Field-Test Items Containing Significant DIF

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE579626	L839L002	3	46	A	S	S	S	S	A	C-	S	B-
ELA	VE645384	L8OSA2081	3	74	A	S	S	S	S	B-	C-	S	B-
ELA	VE579625	L839L001	4	46	A	S	S	S	S	A	B-	S	C-
ELA	VE645236	L892L001	11	46	A	S	S	S	S	A	C-	A	B-
ELA	VE579708	L845L003	14	46	A	S	S	S	S	B-	C-	S	B-
ELA	VE579706	L845L001	14	47	A	S	S	S	S	B-	C-	S	B-
ELA	VE645096	L884K005	17	48	A	S	S	S	S	A	B-	S	C-
ELA	VE644707	L850L008	20	48	B-	S	S	S	S	C-	C-	S	C-
ELA	VE644702	L850L003	21	46	C-	S	S	S	S	S	A	S	A
ELA	VE579809	L854L001	24	46	A	S	S	S	S	S	C-	A	A
ELA	VE579816	L854L008	24	48	A	S	S	S	S	S	A	A	C-
ELA	VE644838	L868L015	31	51	C-	S	A	S	S	A	A	A	B-
ELA	VE645177	L888L001	33	46	B-	S	S	S	S	C+	A	S	A
ELA	VE645178	L888L002	34	46	A	S	S	S	S	B-	C-	S	C-
ELA	VE579594	L836L001	35	46	A	S	S	S	S	A	C+	S	B+
ELA	VE579883	L860L002	38	46	A	S	A	S	S	A	B-	B-	C-
ELA	VE579764	L851L001	40	46	C-	S	S	S	S	B-	A	S	A
ELA	VE579772	L851L009	40	49	A	S	S	S	S	A	B-	S	C-
ELA	VE579766	L851L003	41	46	A	S	S	S	S	C-	C-	S	B-
ELA	VE046792	L6OSA778	41	74	A	S	S	S	S	B-	B-	S	C-
ELA	VE579758	L849L009	42	48	C-	S	S	S	S	B-	C-	S	C-
ELA	VE644743	L863L008	44	49	A	S	S	S	S	B-	C-	S	C-
ELA	VE579851	L857L002	46	46	C-	S	S	S	S	S	C-	B-	A
ELA	VE579872	L859L007	50	46	A	S	S	S	S	B-	C-	C-	A
ELA	VE644636	L834L001	52	46	C-	S	S	S	S	A	C-	S	C-
ELA	VE644643	L834L008	52	49	C-	S	S	S	S	A	B-	S	A
ELA	VE644637	L834L002	53	46	C+	S	S	S	S	A	A	A	A
ELA	VE046979	L6OSA965	54	74	A	S	S	S	S	A	A	S	C-

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE644841	L869L002	55	46	C-	S	S	S	S	B-	C-	S	C-
ELA	VE644848	L869L009	57	48	A	S	S	S	S	C-	B-	S	A
ELA	VE644523	L803M008	58	49	A	S	S	S	S	C-	A	S	B-
ELA	VE580292	L8OSA2022	58	74	A	S	S	S	S	B-	C-	S	B-
ELA	VE644517	L803M002	59	47	A	S	S	S	S	A	C-	S	A
ELA	VE358993	L7OSA1770	59	74	A	S	S	S	S	A	A	S	C-
ELA	VE580068	L891K001	60	46	C-	S	S	S	S	A	C-	S	A
ELA	VE644660	L837L006	65	48	A	S	S	S	S	C-	B-	S	A
ELA	VE046977	L6OSA963	65	74	A	S	S	S	S	A	A	S	C+
ELA	VE580040	L889K004	69	47	B-	S	S	S	S	A	C-	S	A
ELA	VE579346	L812L005	74	46	C-	S	S	S	S	A	B-	S	B-
ELA	VE359000	L7OSA1778	74	74	A	S	S	S	S	A	A	S	C-
ELA	VE579239	L801L009	75	50	C-	S	S	S	S	A	A	S	A
ELA	VE579231	L801L001	76	46	A	S	S	S	S	C-	C-	S	C-
ELA	VE645060	L882L001	77	46	B-	S	S	S	S	B-	C-	S	B-
ELA	VE645063	L882L004	77	47	A	S	S	S	S	A	B-	S	C-
ELA	VE580256	L8OSA1986	79	74	A	S	S	S	S	B-	C-	S	B-
ELA	VE644862	L871L005	82	46	C-	S	S	S	S	A	A	S	A
ELA	VE644922	L874L002	84	46	A	S	S	S	S	A	C-	S	C-
ELA	VE580255	L8OSA1985	84	74	A	S	S	S	S	A	A	S	C-
ELA	VE358950	L7OSA1627	85	74	A	S	S	S	S	C-	B-	S	A
ELA	VE645356	L899L009	86	47	A	S	S	S	S	A	C-	S	A
ELA	VE359069	L7OSA1847	88	74	A	S	S	S	S	C-	A	S	A
ELA	VE580280	L8OSA2010	92	74	A	S	S	S	S	C-	B-	S	A
ELA	VE645316	L897L002	96	46	A	S	S	S	S	A	A	S	C-
ELA	VE645319	L897L005	96	47	A	S	S	S	S	C-	A	S	A
ELA	VE644755	L864L006	101	48	A	S	S	S	S	A	C-	S	A
ELA	VE359083	L7OSA1861	103	74	A	S	S	S	S	C-	B-	S	B-

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE644497	L801M012	104	49	A	S	S	S	S	A	C-	S	A
ELA	VE580278	L8OSA2008	104	74	A	S	S	S	S	A	B-	S	C-
ELA	VE359084	L7OSA1862	105	74	A	S	S	S	S	C-	A	S	A
ELA	VE580290	L8OSA2020	106	74	A	S	S	S	S	A	C-	S	A
ELA	VE644764	L865L001	110	46	C-	S	S	S	S	A	C-	S	B-
ELA	VE644767	L865L004	110	47	A	S	S	S	S	A	A	S	C-
ELA	VE580248	L8OSA1978	110	74	A	S	S	S	S	C-	A	S	C-
ELA	VE645365	L8OSA2062	113	74	A	S	C-	S	S	C-	B-	S	A
ELA	VE644791	L866L012	115	46	A	S	S	S	S	A	A	S	C-
ELA	VE046593	L6OSA1300	116	74	A	S	S	S	S	C+	B+	S	A
ELA	VE046595	L6OSA1302	117	74	A	S	S	S	S	C-	C-	S	A
ELA	VE046610	L6OSA1318	124	74	C-	S	S	S	S	A	C-	S	C-
ELA	VE645016	L879L004	126	47	A	S	S	S	S	B-	C-	S	A
ELA	VE580259	L8OSA1989	127	74	A	S	S	S	S	B+	A	S	C+
ELA	VE046698	L6OSA683	128	74	A	S	S	S	S	C-	A	S	A
ELA	VE580249	L8OSA1979	130	74	A	S	S	S	S	C-	A	S	C-
ELA	VE580250	L8OSA1980	133	74	A	S	S	S	S	C-	A	S	A
ELA	VE644891	L872L004	134	47	B-	S	A	S	S	A	C-	S	A
ELA	VE644890	L872L003	135	46	B-	S	S	S	S	C-	C-	S	C-
ELA	VE644559	L805M004	136	47	C-	S	S	S	S	A	B-	S	C-
ELA	VE645059	L881L016	140	51	A	S	S	S	S	C+	A	S	A
ELA	VE046854	L6OSA840	142	74	A	S	S	S	S	A	C-	S	B-
ELA	VE644936	L875L001	144	46	C-	S	A	S	S	A	C-	S	A
ELA	VE580263	L8OSA1993	144	74	A	S	C+	S	S	C+	A	S	A
ELA	VE644939	L875L004	145	46	B-	S	S	S	S	A	C-	S	A
ELA	VE644938	L875L003	146	46	B-	S	S	S	S	B-	C-	S	C-
ELA	VE644937	L875L002	147	46	C-	S	S	S	S	A	A	S	A
ELA	VE644959	L875L024	147	49	A	S	S	S	S	C-	B-	S	A

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	VE046867	L6OSA853	147	74	A	S	S	S	S	C-	A	S	C-
ELA	VE644907	L873L003	148	47	B-	S	S	S	S	B-	C-	S	C-
ELA	VE580252	L8OSA1982	148	74	A	S	S	S	S	C-	A	S	B-
ELA	VE046868	L6OSA854	149	74	A	S	S	S	S	C-	A	S	A
ELA	VE046869	L6OSA855	151	74	B+	S	S	S	S	C-	A	S	A
ELA	VE046870	L6OSA856	152	74	A	S	S	S	S	C-	A	S	A
ELA	VE046959	L6OSA945	154	74	A	S	S	S	S	A	C-	S	C-
ELA	VE645111	L884L003	155	46	A	S	S	S	S	A	A	S	C-
ELA	VE046960	L6OSA946	155	74	A	S	S	S	S	C-	A	S	A
MATH	VE359637	M60829	2	57	C-	S	A	S	S	A	A	C-	A
MATH	VE580840	M62032	3	18	C-	S	A	S	S	A	A	A	A
MATH	VE359713	M60992	5	19	C-	S	A	S	S	A	A	A	A
MATH	VE580760	M61941	7	18	C-	S	A	S	S	A	A	A	A
MATH	VE645794	M62166	8	58	A	S	A	S	S	A	A	C-	A
MATH	VE580536	M61528	10	20	A	S	A	S	S	B-	B-	C-	B-
MATH	VE580935	M62128	10	79	C-	S	A	S	S	A	A	A	A
MATH	VE580551	M61549	13	21	C-	S	B-	S	S	B-	B-	B-	A
MATH	VE580721	M61880	14	36	B-	S	C-	S	S	C-	C-	B-	B-
MATH	VE645529	M61338	16	57	B-	S	C-	S	S	C-	B-	B-	A
MATH	VE580564	M61579	19	18	C-	S	A	S	S	A	A	A	B-
MATH	VE645801	M62173	19	19	C-	S	A	S	S	A	A	A	A
MATH	VE580579	M61600	19	20	A	S	A	S	S	A	B-	A	C-
MATH	VE580854	M62046	21	20	C-	S	A	S	S	A	A	B-	A
MATH	VE645809	M62183	22	21	B-	S	B-	S	S	B-	C-	C-	A
MATH	VE359359	M40429	22	60	B-	S	C-	S	S	C-	C-	A	A
MATH	VE645548	M61357	24	59	A	S	A	S	S	A	B-	C-	A
MATH	VE645855	M62244	25	18	C-	S	A	S	S	A	A	A	A
MATH	VE645777	M62140	26	21	C-	S	A	S	S	A	B-	C-	B-

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
MATH	VE359331	M40023	28	18	C-	S	A	S	S	A	A	A	A
MATH	VC173559	M32607	28	35	A	S	A	S	S	A	A	A	C-
MATH	VE359675	M60878	28	79	A	S	C+	S	S	B+	A	A	A
MATH	VE359553	M60582	36	35	A	S	C+	S	S	B+	A	A	A
MATH	VE580555	M61558	37	35	A	S	C-	S	S	C-	C-	B-	B-
MATH	VE359813	M61229	40	20	A	S	A	S	S	A	A	C-	A
MATH	VE048554	M51453	41	80	A	S	C+	S	S	C+	A	A	A
MATH	VE359355	M40381	43	57	A	S	B-	S	S	B-	C-	A	A
MATH	VE645435	M61033	44	57	A	S	C-	S	S	C-	C-	B-	B-
MATH	VE580539	M61531	47	20	A	S	C-	S	S	B-	B-	C-	B-
MATH	VE580557	M61568	48	21	A	S	B-	S	S	C-	C-	B-	C-
MATH	VE645759	M61914	50	35	C-	S	A	S	S	B-	B-	A	A
MATH	VE359780	M61160	51	18	C-	S	A	S	S	B-	A	B-	C-
MATH	VE359616	M60742	51	59	B-	S	C-	S	S	C-	B-	C-	B-
MATH	VE645897	M62299	52	21	A	S	A	S	S	A	A	A	C-
MATH	VE580816	M62008	53	19	C+	S	C+	S	S	B+	A	A	A
MATH	VE048151	M51051	53	60	A	S	C+	S	S	B+	A	A	B-
MATH	VE645519	M61296	56	80	A	S	C+	S	S	B+	A	A	A
MATH	VE645853	M62242	57	19	B-	S	A	S	S	A	C-	B-	A
MATH	VE580654	M61710	59	21	A	S	A	S	S	A	B-	C-	A
MATH	VE645596	M61442	61	36	C-	S	A	S	S	B-	A	C-	A
MATH	VE048373	M51272	61	59	A	S	A	S	S	A	A	C-	A
MATH	VE048145	M51045	61	60	C-	S	B-	S	S	B-	B-	B-	A
MATH	VE580751	M61932	62	19	A	S	B-	S	S	C-	B-	C-	A
MATH	VE580534	M61526	62	20	C-	S	B-	S	S	B-	A	B-	B-
MATH	VE359885	M61878	63	20	B-	S	A	S	S	A	A	C-	A
MATH	VE645537	M61346	63	59	A	S	A	S	S	A	B-	C-	A
MATH	VE645778	M62141	64	21	A	S	B-	S	S	C-	C-	B-	A

Test	Accession No.	CAHSEE ID	Form	Item No.	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
MATH	VE580814	M62006	65	19	C+	S	A	S	S	A	A	A	A
MATH	VE580533	M61525	65	20	A	S	C-	S	S	C-	A	A	A
MATH	VE359709	M60988	67	35	A	S	C+	S	S	C+	A	A	A
MATH	VE359884	M61877	68	20	B-	S	A	S	S	B-	A	C-	C-
MATH	VE645547	M61356	68	35	C-	S	B-	S	S	B-	B-	B-	A
MATH	VE359783	M61163	70	18	A	S	A	S	S	A	A	A	C-
MATH	VE359891	M61966	70	21	A	S	C-	S	S	C-	A	A	A
MATH	VE580752	M61933	72	19	B-	S	B-	S	S	B-	C-	B-	B-
MATH	VE645622	M61563	73	35	B-	S	B-	S	S	B-	A	C-	A
MATH	VE580543	M61535	74	57	A	S	C+	S	S	B+	A	A	A
MATH	VE359455	M60266	74	80	B+	S	C+	S	S	C+	A	A	A
MATH	VE359527	M60504	78	60	C-	S	C-	S	S	C-	B-	C-	C-

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.
S indicates that DIF was not performed due to insufficient sample size.

Table 6.J.6: Distribution of Field-Test Item DIF Classifications—ELA

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C items Across All Analyse[1]	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-	17	2	0	0	1	0	0	0	0	0	26	2	39	4	1	0	33	3	85	8
B-	42	4	0	0	4	0	0	0	0	0	66	6	84	8	13	1	96	9		
A	978	93	0	0	82	8	0	0	0	0	884	84	920	88	80	8	911	87		
B+	11	1	0	0	2	0	0	0	0	0	36	3	5	0	1	0	7	1		
C+	1	0	0	0	1	0	0	0	0	0	4	0	1	0	0	0	2	0	8	1
Small N[2]	0	0	1049	100	959	91	1049	100	1049	100	33	3	0	0	954	91	0	0		
Total[3]	1049	100	1049	100	1049	100	1049	100	1049	100	1049	100	1049	100	1049	100	1049	100	93	9

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[3] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.J.7: Distribution of Field-Test Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C items Across All Analyse[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-	19	2	0	0	10	1	0	0	0	0	12	1	10	1	18	2	8	1	49	5
B-	51	5	0	0	41	4	0	0	0	0	42	4	52	6	88	9	57	6		
A	851	91	0	0	793	85	0	0	0	0	822	88	872	93	822	88	869	93		
B+	13	1	0	0	83	9	0	0	0	0	57	6	2	0	8	1	2	0		
C+	2	0	0	0	9	1	0	0	0	0	3	0	0	0	0	0	0	0	10	1
Small N[2]	0	0	936	100	0	0	936	100	936	100	0	0	0	0	0	0	0	0		
Total[3]	936	100	936	100	936	100	936	100	936	100	936	100	936	100	936	100	936	100	59	6

[1] Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
[2] Small N indicates that DIF analysis was not performed due to insufficient sample size.
[3] Total percentage may differ slightly from the sum of the category percentages due to rounding.

Appendix 6.K: Intercorrelations, Reliability Estimates and Standard Errors of Measurement

Table 6.K.1: Intercorrelations and Reliability Estimates by Section—July 2011

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.63	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.67	0.34	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.82	0.40	0.51	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.83	0.41	0.51	0.63	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.73	0.33	0.40	0.52	0.52	1.00	-	-	-	-	-	-	-
Writing Conventions	0.71	0.35	0.40	0.45	0.47	0.47	1.00	-	-	-	-	-	-
Mathematics	0.62	0.36	0.43	0.55	0.52	0.54	0.52	1.00	-	-	-	-	-
Probability and Statistics	0.58	0.34	0.41	0.51	0.49	0.49	0.45	0.74	1.00	-	-	-	-
Number Sense	0.47	0.26	0.35	0.41	0.39	0.39	0.41	0.74	0.45	1.00	-	-	-
Algebra & Functions	0.55	0.33	0.37	0.49	0.45	0.48	0.45	0.84	0.55	0.50	1.00	-	-
Measurement & Geometry	0.51	0.29	0.35	0.46	0.42	0.46	0.43	0.77	0.48	0.43	0.55	1.00	-
Algebra 1	0.33	0.19	0.21	0.29	0.27	0.31	0.28	0.60	0.29	0.34	0.41	0.36	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	47.07	2.05	3.65	10.19	11.60	5.33	7.08	37.08	6.82	8.62	9.53	7.72	4.40
SD	11.98	0.55	1.59	3.38	3.40	2.46	2.75	10.57	2.56	2.93	3.50	2.99	2.09
Reliability	0.87	0.26	0.45	0.68	0.66	0.57	0.60	0.85	0.59	0.56	0.65	0.58	0.41
SEM	4.40	0.47	1.17	1.91	2.00	1.61	1.74	4.16	1.65	1.95	2.06	1.93	1.61

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 7,162 examinees completing the ELA section.
Correlations for the mathematics section are reported for 6,908 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 2,979 examinees taking both sections.

Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2011

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.65	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.67	0.35	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.86	0.45	0.52	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.87	0.45	0.56	0.71	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.78	0.41	0.47	0.60	0.58	1.00	-	-	-	-	-	-	-
Writing Conventions	0.77	0.44	0.44	0.57	0.55	0.56	1.00	-	-	-	-	-	-
Mathematics	0.68	0.45	0.49	0.58	0.57	0.59	0.57	1.00	-	-	-	-	-
Probability and Statistics	0.63	0.41	0.46	0.54	0.55	0.54	0.52	0.78	1.00	-	-	-	-
Number Sense	0.57	0.38	0.44	0.48	0.50	0.48	0.48	0.79	0.55	1.00	-	-	-
Algebra & Functions	0.61	0.41	0.44	0.53	0.52	0.53	0.52	0.88	0.64	0.60	1.00	-	-
Measurement & Geometry	0.57	0.38	0.42	0.49	0.47	0.51	0.49	0.85	0.58	0.58	0.68	1.00	-
Algebra1	0.44	0.32	0.30	0.37	0.34	0.40	0.39	0.69	0.40	0.43	0.54	0.53	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	49.83	2.11	3.85	10.72	11.84	5.67	8.24	38.81	6.95	8.82	10.24	8.12	4.69
SD	13.90	0.56	1.68	3.67	4.13	2.62	3.08	12.98	2.82	3.24	4.00	3.53	2.42
Reliability	0.90	0.31	0.49	0.74	0.77	0.63	0.68	0.90	0.66	0.67	0.74	0.70	0.57
SEM	4.37	0.47	1.20	1.88	1.97	1.58	1.75	4.10	1.64	1.86	2.03	1.94	1.58

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 36,320 examinees completing the ELA section.
Correlations for the mathematics section are reported for 35,663 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 20,028 examinees taking both sections.

Table 6.K.3: Intercorrelations and Reliability Estimates by Section—November 2011

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.68	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.74	0.42	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.87	0.49	0.63	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.87	0.49	0.62	0.72	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.80	0.44	0.51	0.63	0.62	1.00	-	-	-	-	-	-	-
Writing Conventions	0.81	0.49	0.52	0.60	0.59	0.62	1.00	-	-	-	-	-	-
Mathematics	0.70	0.48	0.49	0.62	0.61	0.60	0.62	1.00	-	-	-	-	-
Probability and Statistics	0.65	0.43	0.48	0.59	0.57	0.53	0.55	0.79	1.00	-	-	-	-
Number Sense	0.58	0.41	0.40	0.52	0.50	0.47	0.51	0.81	0.57	1.00	-	-	-
Algebra & Functions	0.64	0.43	0.44	0.57	0.55	0.55	0.57	0.89	0.65	0.63	1.00	-	-
Measurement & Geometry	0.59	0.41	0.40	0.51	0.51	0.52	0.53	0.82	0.56	0.54	0.65	1.00	-
Algebra1	0.46	0.33	0.28	0.39	0.39	0.42	0.43	0.70	0.40	0.48	0.56	0.51	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	51.08	2.04	4.76	10.37	12.42	5.93	8.40	40.67	7.58	9.11	10.51	8.63	4.85
SD	14.68	0.60	1.73	3.77	3.86	2.85	3.31	13.07	2.93	3.25	4.16	3.33	2.41
Reliability	0.91	0.35	0.59	0.75	0.74	0.70	0.74	0.90	0.71	0.65	0.76	0.67	0.57
SEM	4.36	0.48	1.11	1.90	1.96	1.57	1.69	4.10	1.59	1.91	2.05	1.90	1.58

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 93,546 examinees completing the ELA section.
Correlations for the mathematics section are reported for 93,317 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 54,617 examinees taking both sections.

Table 6.K.4: Intercorrelations and Reliability Estimates by Section—December 2011

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.61	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.63	0.30	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.85	0.39	0.49	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.85	0.43	0.48	0.66	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.70	0.32	0.36	0.52	0.48	1.00	-	-	-	-	-	-	-
Writing Conventions	0.76	0.40	0.40	0.54	0.55	0.43	1.00	-	-	-	-	-	-
Mathematics	0.61	0.39	0.47	0.53	0.52	0.53	0.47	1.00	-	-	-	-	-
Probability and Statistics	0.54	0.37	0.41	0.45	0.46	0.46	0.43	0.72	1.00	-	-	-	-
Number Sense	0.50	0.31	0.38	0.43	0.43	0.46	0.37	0.78	0.47	1.00	-	-	-
Algebra & Functions	0.58	0.39	0.41	0.50	0.49	0.48	0.46	0.85	0.53	0.57	1.00	-	-
Measurement & Geometry	0.48	0.29	0.41	0.43	0.40	0.42	0.35	0.79	0.47	0.49	0.58	1.00	-
Algebra1	0.24	0.15	0.20	0.21	0.20	0.22	0.16	0.54	0.25	0.27	0.36	0.35	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	48.41	2.04	3.82	10.10	11.72	5.35	8.21	37.21	6.77	7.84	10.24	7.98	4.38
SD	11.97	0.44	1.54	3.45	3.66	2.34	2.80	10.60	2.38	3.12	3.62	2.92	1.96
Reliability	0.87	0.27	0.37	0.69	0.70	0.52	0.61	0.85	0.52	0.63	0.67	0.57	0.32
SEM	4.25	0.38	1.22	1.92	2.01	1.63	1.74	4.14	1.66	1.90	2.08	1.92	1.61

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 2,317 examinees completing the ELA section.
Correlations for the mathematics section are reported for 2,086 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 822 examinees taking both sections.

Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 2012

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.73	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.79	0.50	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.92	0.58	0.71	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.90	0.57	0.71	0.80	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.87	0.56	0.64	0.77	0.72	1.00	-	-	-	-	-	-	-
Writing Conventions	0.87	0.57	0.65	0.75	0.71	0.74	1.00	-	-	-	-	-	-
Mathematics	0.81	0.58	0.61	0.75	0.71	0.72	0.73	1.00	-	-	-	-	-
Probability and Statistics	0.73	0.51	0.57	0.68	0.65	0.64	0.65	0.83	1.00	-	-	-	-
Number Sense	0.72	0.52	0.54	0.66	0.63	0.64	0.65	0.90	0.70	1.00	-	-	-
Algebra & Functions	0.77	0.54	0.58	0.71	0.67	0.68	0.69	0.94	0.75	0.80	1.00	-	-
Measurement & Geometry	0.75	0.53	0.56	0.69	0.65	0.67	0.67	0.92	0.72	0.78	0.82	1.00	-
Algebra 1	0.66	0.50	0.48	0.60	0.56	0.60	0.59	0.86	0.62	0.73	0.77	0.75	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	60.87	2.33	5.33	12.68	14.56	7.67	10.12	52.86	9.34	11.35	13.64	11.51	7.03
SD	16.61	0.65	1.66	4.19	4.14	2.90	3.50	17.04	2.77	3.89	4.71	4.45	3.15
Reliability	0.94	0.43	0.65	0.84	0.83	0.76	0.79	0.95	0.73	0.81	0.85	0.85	0.79
SEM	4.05	0.49	0.97	1.67	1.73	1.44	1.58	3.67	1.43	1.69	1.80	1.74	1.45

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 157,199 examinees completing the ELA section.
Correlations for the mathematics section are reported for 154,107 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 132,889 examinees taking both sections.

Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 2012

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.71	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.76	0.45	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.90	0.54	0.68	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.89	0.54	0.68	0.78	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.85	0.52	0.59	0.72	0.69	1.00	-	-	-	-	-	-	-
Writing Conventions	0.86	0.56	0.60	0.71	0.70	0.71	1.00	-	-	-	-	-	-
Mathematics	0.79	0.55	0.57	0.72	0.68	0.68	0.71	1.00	-	-	-	-	-
Probability and Statistics	0.72	0.48	0.53	0.65	0.63	0.61	0.63	0.85	1.00	-	-	-	-
Number Sense	0.69	0.49	0.50	0.63	0.59	0.59	0.62	0.89	0.71	1.00	-	-	-
Algebra & Functions	0.74	0.50	0.54	0.67	0.63	0.63	0.66	0.92	0.74	0.76	1.00	-	-
Measurement & Geometry	0.73	0.51	0.52	0.66	0.61	0.63	0.65	0.92	0.73	0.76	0.80	1.00	-
Algebra 1	0.65	0.47	0.44	0.58	0.53	0.57	0.59	0.86	0.65	0.71	0.74	0.76	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	65.23	2.39	5.55	13.42	16.07	8.26	11.20	56.46	9.62	11.93	14.73	12.48	7.70
SD	14.63	0.59	1.51	3.67	3.52	2.81	3.18	15.78	2.73	3.56	4.14	4.17	3.07
Reliability	0.93	0.39	0.59	0.81	0.81	0.74	0.78	0.95	0.74	0.78	0.82	0.85	0.78
SEM	3.81	0.46	0.96	1.60	1.53	1.43	1.48	3.56	1.39	1.67	1.75	1.63	1.45

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 387,959 examinees completing the ELA section.
Correlations for the mathematics section are reported for 383,250 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 357,930 examinees taking both sections.

Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 2012

	ELA	Essay	Word Anal.	Read. Comp.	Lit. Resp. & Anal.	Writing Strat.	Writing Conv.	Math	Prob. & Stat.	Number Sense	Alg. & Func.	Meas. & Geom.	A1
ELA	1.00	-	-	-	-	-	-	-	-	-	-	-	-
Essay	0.69	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.75	0.44	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.88	0.50	0.64	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.87	0.51	0.63	0.73	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.81	0.46	0.56	0.68	0.62	1.00	-	-	-	-	-	-	-
Writing Conventions	0.83	0.51	0.57	0.66	0.62	0.65	1.00	-	-	-	-	-	-
Mathematics	0.73	0.50	0.54	0.65	0.63	0.63	0.65	1.00	-	-	-	-	-
Probability and Statistics	0.65	0.44	0.49	0.58	0.58	0.55	0.58	0.81	1.00	-	-	-	-
Number Sense	0.61	0.43	0.46	0.54	0.54	0.52	0.55	0.85	0.62	1.00	-	-	-
Algebra & Functions	0.68	0.46	0.51	0.61	0.59	0.58	0.61	0.90	0.69	0.69	1.00	-	-
Measurement & Geometry	0.63	0.42	0.47	0.57	0.54	0.56	0.57	0.86	0.62	0.63	0.71	1.00	-
Algebra 1	0.50	0.37	0.35	0.45	0.42	0.44	0.46	0.73	0.47	0.56	0.59	0.56	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	48.58	2.08	4.34	9.81	11.47	5.83	7.76	38.67	7.22	8.47	10.26	8.46	4.26
SD	15.87	0.66	1.80	4.07	4.21	2.86	3.49	14.57	2.98	3.64	4.35	3.83	2.53
Reliability	0.92	0.36	0.59	0.78	0.78	0.71	0.75	0.92	0.71	0.75	0.78	0.75	0.64
SEM	4.48	0.53	1.16	1.89	1.97	1.54	1.75	4.05	1.62	1.82	2.02	1.93	1.53

Note: All correlations are for raw scores. Students who took Braille, Large Print and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 39,144 examinees completing the ELA section.
Correlations for the mathematics section are reported for 36,342 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 19,376 examinees taking both sections.

Table 6.K.8: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—February 2012

	Total MC N=154,910		Composite (MC + Essay) N=154,910	
	R_{XX}	SEM	R_{XX}	SEM
Gender				
Male	0.95	3.45	0.94	4.14
Female	0.94	3.31	0.94	3.92
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.41	0.93	4.12
Asian	0.95	3.18	0.95	3.95
Pacific Islander	0.93	3.51	0.93	4.12
Filipino	0.93	3.15	0.92	3.89
Hispanic or Latino	0.94	3.60	0.93	4.15
African American	0.94	3.61	0.93	4.24
White (not of Hispanic origin)	0.94	3.03	0.93	3.82
Two or More Races	0.95	3.18	0.94	3.96
Language Fluency				
English Proficient Students	0.94	3.21	0.93	3.93
English Learner Students	0.88	3.91	0.88	4.41
Special Education Program Participation				
Students Receiving Services	0.92	3.83	0.91	4.51
Students Not Receiving Services	0.94	3.31	0.93	3.97

Table 6.K.9: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—February 2012

	Total N=155,492	
	R_{XX}	SEM
Gender		
Male	0.96	3.66
Female	0.95	3.68
Race/Ethnicity		
American Indian or Alaska Native	0.95	3.78
Asian	0.95	3.02
Pacific Islander	0.94	3.77
Filipino	0.94	3.37
Hispanic or Latino	0.94	3.90
African American	0.94	3.97
White (not of Hispanic origin)	0.95	3.41
Two or More Races	0.96	3.53
Language Fluency		
English Proficient Students	0.95	3.56
English Learner Students	0.91	4.09
Special Education Program Participation		
Students Receiving Services	0.92	4.07
Students Not Receiving Services	0.95	3.62

Table 6.K.10: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—March 2012

	Total MC N=402,061		Composite (MC + Essay) N=402,061	
	R_{XX}	SEM	R_{XX}	SEM
Gender				
Male	0.94	3.28	0.94	3.90
Female	0.93	3.10	0.92	3.70
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.23	0.93	3.88
Asian	0.94	2.81	0.93	3.61
Pacific Islander	0.93	3.25	0.92	3.83
Filipino	0.92	2.92	0.91	3.61
Hispanic or Latino	0.93	3.36	0.93	3.87
African American	0.94	3.40	0.94	3.97
White (not of Hispanic origin)	0.93	2.86	0.91	3.61
Two or More Races	0.95	3.04	0.94	3.77
Language Fluency				
English Proficient Students	0.92	3.03	0.91	3.69
English Learner Students	0.90	3.80	0.90	4.18
Special Education Program Participation				
Students Receiving Services	0.93	3.76	0.92	4.30
Students Not Receiving Services	0.93	3.12	0.92	3.74

Table 6.K.11: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—March 2012

	Total N=401,838	
	R_{XX}	SEM
Gender		
Male	0.95	3.54
Female	0.94	3.58
Race/Ethnicity		
American Indian or Alaska Native	0.94	3.69
Asian	0.94	2.77
Pacific Islander	0.94	3.62
Filipino	0.94	3.20
Hispanic or Latino	0.94	3.74
African American	0.94	3.86
White (not of Hispanic origin)	0.94	3.29
Two or More Races	0.95	3.47
Language Fluency		
English Proficient Students	0.94	3.44
English Learner Students	0.92	4.05
Special Education Program Participation		
Students Receiving Services	0.93	4.06
Students Not Receiving Services	0.94	3.50

Appendix 6.L: Rater Agreement Analyses

Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	217	0	0	0	0	217	3
1	0	224	163	1	0	388	5
2	0	135	4,749	556	10	5,450	76
3	0	3	536	466	35	1,040	15
4	0	0	8	32	27	67	1
Total	217	362	5,456	1,055	72	7,162	100
Percent	3	5	76	15	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	5,683	79	79
1	1,457	20	100
2	22	0	100
3	0	0	100
4	0	0	100

Table 6.L.2: Summary Statistics for the ELA Essay Item—July 2010

	First Rating	Second Rating
Mean	2.05	2.06
Standard Deviation	0.60	0.60
Mean Absolute Difference Between First and Second Ratings: 0.21		
Correlation of First and Second Ratings: 0.70		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	116
Illegible	0
Off Topic	99
Cartoon / Inappropriate	1
Not in English	1
Total	**217**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.3: Agreement of First and Second Ratings on the ELA Essay Item—October 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	902	0	0	0	0	902	2
1	0	983	647	2	1	1,633	4
2	0	623	22,914	3,145	64	26,746	74
3	0	2	3,102	3,101	311	6,516	18
4	0	0	61	251	211	523	1
Total	902	1,608	26,724	6,499	587	36,320	100
Percent	2	4	74	18	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	28,111	77	77
1	8,079	22	100
2	129	0	100
3	1	0	100
4	0	0	100

Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2010

	First Rating	Second Rating
Mean	2.11	2.12
Standard Deviation	0.61	0.61
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.68		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	631
Illegible	0
Off Topic	236
Cartoon / Inappropriate	23
Not in English	12
Total	**902**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.5: Agreement of First and Second Ratings on the ELA Essay Item—November 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	2,910	0	0	0	0	2,910	3
1	0	4,846	2,746	16	1	7,609	8
2	0	2,765	56,276	7,443	147	66,631	71
3	0	29	7,545	6,853	705	15,132	16
4	0	1	152	674	437	1,264	1
Total	2,910	7,641	66,719	14,986	1,290	93,546	100
Percent	3	8	71	16	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	71,322	76	76
1	21,878	23	100
2	344	0	100
3	2	0	100
4	0	0	100

Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2010

	First Rating	Second Rating
Mean	2.05	2.04
Standard Deviation	0.65	0.65
Mean Absolute Difference Between First and Second Ratings: 0.24		
Correlation of First and Second Ratings: 0.70		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	1,761
Illegible	0
Off Topic	1,025
Cartoon / Inappropriate	102
Not in English	22
Total	**2,910**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.7: Agreement of First and Second Ratings on the ELA Essay Item—December 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	31	0	0	0	0	31	1
1	0	82	39	0	0	121	5
2	0	46	1,696	150	2	1,894	82
3	0	0	136	120	7	263	11
4	0	0	0	5	3	8	0
Total	31	128	1,871	275	12	2,317	100
Percent	1	6	81	12	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	1,932	83	83
1	383	17	100
2	2	0	100
3	0	0	100
4	0	0	100

Table 6.L.8: Summary Statistics for the ELA Essay Item—December 2010

	First Rating	Second Rating
Mean	2.04	2.05
Standard Deviation	0.48	0.50
Mean Absolute Difference Between First and Second Ratings: 0.17		
Correlation of First and Second Ratings: 0.65		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	13
Illegible	0
Off Topic	15
Cartoon / Inappropriate	1
Not in English	2
Total	**31**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.9: Agreement of First and Second Ratings on the ELA Essay Item—February 2012

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	2,609	0	0	0	0	2,609	2
1	0	3,875	2,287	21	2	6,185	4
2	0	2,252	73,583	16,016	522	92,373	59
3	0	22	15,635	28,453	4,060	48,170	31
4	0	1	472	4,153	3,236	7,862	5
Total	2,609	6,150	91,977	48,643	7,820	157,199	100
Percent	2	4	59	31	5	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	111,756	71	71
1	44,403	28	99
2	1,037	1	100
3	3	0	100
4	0	0	100

Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2011

	First Rating	Second Rating
Mean	2.33	2.34
Standard Deviation	0.71	0.71
Mean Absolute Difference Between First and Second Ratings: 0.30		
Correlation of First and Second Ratings: 0.69		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	1,698
Illegible	0
Off Topic	813
Cartoon / Inappropriate	76
Not in English	22
Total	**2,609**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.11: Agreement of First and Second Ratings on the ELA Essay Item—March 2012

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	3,467	0	0	0	0	3,467	1
1	0	5,147	3,165	29	2	8,343	2
2	0	3,240	181,896	42,479	1,139	228,754	59
3	0	31	42,096	78,375	9,459	129,961	33
4	0	1	1,093	9,518	6,822	17,434	4
Total	3,467	8,419	228,250	130,401	17,422	387,959	100
Percent	1	2	59	34	4	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	275,707	71	71
1	109,957	28	99
2	2,292	1	100
3	3	0	100
4	0	0	100

Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2011

	First Rating	Second Rating
Mean	2.39	2.39
Standard Deviation	0.65	0.65
Mean Absolute Difference Between First and Second Ratings: 0.30		
Correlation of First and Second Ratings: 0.64		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	2,325
Illegible	0
Off Topic	1,018
Cartoon / Inappropriate	86
Not in English	39
Total	**3,468**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Table 6.L.13: Agreement of First and Second Ratings on the ELA Essay Item—May 2012

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	1,572	0	0	0	0	1,572	4
1	0	1,657	928	6	0	2,591	7
2	0	870	22,687	3,096	53	26,706	68
3	0	9	3,252	3,812	401	7,474	19
4	0	0	58	392	351	801	2
Total	1,572	2,536	26,925	7,306	805	39,144	100
Percent	4	6	69	19	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	30,079	77	77
1	8,939	23	100
2	126	0	100
3	0	0	100
4	0	0	100

Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2011

	First Rating	Second Rating
Mean	2.09	2.08
Standard Deviation	0.70	0.70
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.75		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	1,016
Illegible	0
Off Topic	500
Cartoon / Inappropriate	40
Not in English	16
Total	**1,572**

Note: A final score of 0 is assigned when the first or second rater assigns a 0 score and the adjudicator assigns a 0 score.

Appendix 6.M: Generalizability Analyses

Table 6.M.1: Generalizability Results—July 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	7,057	0.59345	0.24250	69.10
Rater (R)	1	0.21430	0.00001	0.00
PR,*e*	7,057	0.10844	0.10844	30.90
Generalizability Coefficient				0.82
Dependability Coefficient				0.82

Table 6.M.2: Generalizability Results—October 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	35,002	0.61744	0.24907	67.61
Rater (R)	1	0.27998	0.00000	0.00
PR,*e*	35,002	0.11930	0.11930	32.39
Generalizability Coefficient				0.81
Dependability Coefficient				0.81

Table 6.M.3: Generalizability Results—November 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	89,448	0.70528	0.28989	69.79
Rater (R)	1	0.07392	0.00000	0.00
PR,*e*	89,448	0.12551	0.12551	30.21
Generalizability Coefficient				0.82
Dependability Coefficient				0.82

Table 6.M.4: Generalizability Results—December 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	2,278	0.39626	0.15556	64.63
Rater (R)	1	0.03159	0.00000	0.00
PR,e	2,278	0.08515	0.08515	35.38
Generalizability Coefficient				0.79
Dependability Coefficient				0.79

Table 6.M.5: Generalizability Results—February 2012

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	153,818	0.83797	0.34106	68.63
Rater (R)	1	0.56253	0.00000	0.00
PR,*e*	153,818	0.15586	0.15586	31.37
Generalizability Coefficient				0.81
Dependability Coefficient				0.81

Table 6.M.6: Generalizability Results—March 2012

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	381,427	0.68607	0.26567	63.19
Rater (R)	1	0.15154	0.00000	0.00
PR,*e*	381,427	0.15474	0.15474	36.81
Generalizability Coefficient				0.77
Dependability Coefficient				0.77

Table 6.M.7: Generalizability Results—May 2012

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	37,750	0.84825	0.36351	74.99
Rater (R)	1	0.14602	0.00000	0.00
PR,*e*	37,750	0.12123	0.12123	25.01
Generalizability Coefficient				0.86
Dependability Coefficient				0.86

Appendix 6.N: Decision Classification Reliability Analyses

Table 6.N.1: ESEA Reliability Classifications—July 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.00	0.01	0.01	0.02
Proficient (68-75)	0.00	0.02	0.01	0.03
Below Proficient (0-67)	0.00	0.01	0.94	0.95
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.97

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.01	0.01	0.01	0.02
Proficient (68-75)	0.00	0.02	0.01	0.03
Below Proficient (0-67)	0.00	0.02	0.93	0.95
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.96

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.00	0.00	0.00	0.01
Proficient (59-72)	0.00	0.02	0.01	0.03
Below Proficient (0-58)	0.00	0.01	0.96	0.97
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.98

[1]True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.00	0.00	0.00	0.01
Proficient (59-72)	0.00	0.02	0.01	0.03
Below Proficient (0-58)	0.00	0.01	0.95	0.97
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.97

Table 6.N.2: Pass/Not Pass Classifications—July 2011

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.23	0.04	0.27
0-53	0.06	0.67	0.73
Estimated Proportion Correctly Classified			0.90

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.22	0.05	0.27
0-53	0.08	0.64	0.73
Estimated Proportion Consistently Classified			0.86

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.23	0.05	0.27
0-42	0.04	0.68	0.73
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Score	Pass
43-80	0.21	0.06	0.27
0-42	0.07	0.66	0.73
Estimated Proportion Consistently Classified			0.87

Table 6.N.3: ESEA Reliability Classifications—October 2011

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.02	0.01	0.01	0.04
Proficient (69-76)	0.00	0.04	0.02	0.05
Below Proficient (0-68)	0.00	0.02	0.89	0.91
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.96

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.02	0.01	0.01	0.04
Proficient (69-76)	0.00	0.03	0.02	0.05
Below Proficient (0-68)	0.00	0.03	0.87	0.91
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.01	0.01	0.01	0.02
Proficient (58-71)	0.00	0.06	0.01	0.07
Below Proficient (0-57)	0.00	0.02	0.89	0.91
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.97

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.01	0.01	0.00	0.02
Proficient (58-71)	0.00	0.06	0.01	0.07
Below Proficient (0-57)	0.00	0.03	0.88	0.91
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

Table 6.N.4: Pass/Not Pass Classifications—October 2011

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.33	0.04	0.37
0-54	0.05	0.58	0.63
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.31	0.06	0.37
0-54	0.07	0.56	0.63
Estimated Proportion Consistently Classified			0.87

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.34	0.05	0.39
0-41	0.03	0.58	0.61
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.33	0.06	0.39
0-41	0.06	0.56	0.61
Estimated Proportion Consistently Classified			0.88

Table 6.N.5: ESEA Reliability Classifications—November 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.03	0.02	0.01	0.06
Proficient (69-75)	0.00	0.04	0.02	0.06
Below Proficient (0-68)	0.00	0.02	0.85	0.88
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.95

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.03	0.02	0.01	0.06
Proficient (69-75)	0.01	0.03	0.02	0.06
Below Proficient (0-68)	0.01	0.04	0.84	0.88
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.96
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.93

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.01	0.01	0.01	0.02
Proficient (59-72)	0.00	0.07	0.01	0.07
Below Proficient (0-58)	0.00	0.02	0.88	0.90
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.96

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.01	0.01	0.01	0.02
Proficient (59-72)	0.00	0.06	0.01	0.07
Below Proficient (0-58)	0.00	0.03	0.87	0.90
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

Table 6.N.6: Pass/Not Pass Classifications—November 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.36	0.04	0.40
0-55	0.05	0.55	0.60
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.34	0.06	0.40
0-55	0.07	0.53	0.60
Estimated Proportion Consistently Classified			0.87

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.37	0.05	0.43
0-42	0.04	0.54	0.57
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.36	0.07	0.43
0-42	0.06	0.52	0.57
Estimated Proportion Consistently Classified			0.88

Table 6.N.7: ESEA Reliability Classifications—December 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.01	0.01	0.01	0.02
Proficient (68-75)	0.00	0.02	0.01	0.03
Below Proficient (0-67)	0.00	0.01	0.93	0.94
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.97

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.01	0.01	0.01	0.02
Proficient (68-75)	0.00	0.02	0.01	0.03
Below Proficient (0-67)	0.00	0.02	0.92	0.94
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.00	0.00	0.00	0.01
Proficient (57-71)	0.00	0.03	0.01	0.04
Below Proficient (0-56)	0.00	0.01	0.94	0.95
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.98

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.00	0.00	0.00	0.01
Proficient (57-71)	0.00	0.03	0.01	0.04
Below Proficient (0-56)	0.00	0.02	0.93	0.95
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.96

Table 6.N.8: Pass/Not Pass Classifications—December 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.25	0.04	0.29
0-54	0.05	0.66	0.71
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.23	0.05	0.29
0-54	0.08	0.63	0.71
Estimated Proportion Consistently Classified			0.87

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
41-80	0.28	0.06	0.34
0-40	0.04	0.61	0.66
Estimated Proportion Correctly Classified			0.90

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
41-80	0.27	0.07	0.34
0-40	0.07	0.58	0.66
Estimated Proportion Consistently Classified			0.86

Table 6.N.9: ESEA Reliability Classifications—February 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (75-90)	0.22	0.04	0.00	0.26
Proficient (68-74)	0.03	0.10	0.04	0.18
Below Proficient (0-67)	0.00	0.03	0.53	0.57
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.93
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.93

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (75-90)	0.21	0.04	0.01	0.26
Proficient (68-74)	0.05	0.08	0.05	0.18
Below Proficient (0-67)	0.00	0.04	0.52	0.57
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.90
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.90

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.13	0.03	0.00	0.15
Proficient (59-72)	0.02	0.23	0.03	0.28
Below Proficient (0-58)	0.00	0.03	0.54	0.56
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.96
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.94

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.12	0.03	0.00	0.15
Proficient (59-72)	0.03	0.21	0.04	0.28
Below Proficient (0-58)	0.00	0.04	0.53	0.56
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.94
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.92

Table 6.N.10: Pass/Not Pass Classifications—February 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.67	0.03	0.69
0-53	0.03	0.27	0.31
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.65	0.04	0.69
0-53	0.05	0.26	0.31
Estimated Proportion Consistently Classified			0.92

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.68	0.02	0.70
0-42	0.03	0.26	0.30
Estimated Proportion Correctly Classified			0.95

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.67	0.04	0.70
0-42	0.04	0.25	0.30
Estimated Proportion Consistently Classified			0.92

Table 6.N.11: ESEA Reliability Classifications—March 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.24	0.04	0.00	0.28
Proficient (69-75)	0.04	0.13	0.06	0.23
Below Proficient (0-68)	0.00	0.02	0.47	0.49
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.92
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.92

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.23	0.05	0.00	0.28
Proficient (69-75)	0.06	0.11	0.07	0.23
Below Proficient (0-68)	0.00	0.04	0.45	0.49
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.89
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.89

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.17	0.03	0.00	0.20
Proficient (58-71)	0.03	0.28	0.03	0.34
Below Proficient (0-57)	0.00	0.03	0.43	0.46
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.94
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.94

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.16	0.04	0.00	0.20
Proficient (58-71)	0.04	0.25	0.05	0.34
Below Proficient (0-57)	0.00	0.04	0.41	0.46
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.92
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.91

Table 6.N.12: Pass/Not Pass Classifications—March 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.75	0.04	0.78
0-55	0.02	0.19	0.22
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.73	0.05	0.78
0-55	0.03	0.18	0.22
Estimated Proportion Consistently Classified			0.92

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.77	0.02	0.80
0-42	0.03	0.18	0.20
Estimated Proportion Correctly Classified			0.95

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.76	0.03	0.80
0-42	0.03	0.17	0.20
Estimated Proportion Consistently Classified			0.93

Table 6.N.13: ESEA Reliability Classifications—May 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (78-90)	0.03	0.02	0.01	0.05
Proficient (70-77)	0.00	0.05	0.02	0.07
Below Proficient (0-69)	0.00	0.01	0.87	0.88
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.96

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (78-90)	0.03	0.01	0.01	0.05
Proficient (70-77)	0.01	0.04	0.02	0.07
Below Proficient (0-69)	0.00	0.03	0.85	0.88
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.00	0.01	0.01	0.03
Proficient (59-72)	0.00	0.07	0.02	0.09
Below Proficient (0-58)	0.00	0.01	0.88	0.88
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.96

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.00	0.01	0.01	0.03
Proficient (59-72)	0.00	0.06	0.03	0.09
Below Proficient (0-58)	0.00	0.02	0.86	0.88
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Table 6.N.14: Pass/Not Pass Classifications—May 2012

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.27	0.02	0.30
0-56	0.05	0.65	0.70
Estimated Proportion Correctly Classified			0.93

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.26	0.04	0.30
0-56	0.07	0.64	0.70
Estimated Proportion Consistently Classified			0.90

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.34	0.01	0.35
0-42	0.07	0.58	0.65
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.32	0.03	0.35
0-42	0.09	0.56	0.65
Estimated Proportion Consistently Classified			0.88

Appendix 6.O: Scoring Tables for Operational and Special Test Versions

Table 6.O.1: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, July 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	498.8256	450	23
88	475.6605	450	23
87	462.1805	450	23
86	452.5208	450	23
85	444.8414	445	21
84	438.3731	438	17
83	432.7140	433	17
82	427.6419	428	17
81	423.0084	423	15
80	418.7227	419	14
79	414.7220	415	14
78	410.9779	411	13
77	407.4266	407	13
76	404.0518	404(Advanced)	12
75	400.8262	401	12
74	397.7318	398	12
73	394.7539	395	11
72	391.8804	392	11
71	389.1015	389	11
70	386.4064	386	11
69	383.7861	384	10
68	381.2321	381(Proficient)	10
67	378.7391	379	10
66	376.3052	376	10
65	373.9195	374	10
64	371.5787	372	10
63	369.2817	369	9
62	367.0219	367	9
61	364.7967	365	9
60	362.6019	363	9
59	360.4369	360	9
58	358.2973	358	9
57	356.1817	356	9
56	354.0876	354	9
55	352.0123	352	9
54	349.9528	350(Pass)	9
53	347.9124	348	9
52	345.8838	346	9
51	343.8688	344	9
50	341.8624	342	9
49	339.8647	340	9
48	337.8739	338	9
47	335.8887	336	9
46	333.9081	334	9
45	331.9330	332	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	329.9597	330	9
43	327.9881	328	9
42	326.0192	326	9
41	324.0499	324	9
40	322.0809	322	9
39	320.1131	320	9
38	318.1451	318	9
37	316.1787	316	9
36	314.2121	314	9
35	312.2483	312	9
34	310.2846	310	9
33	308.3235	308	9
32	306.3635	306	9
31	304.4072	304	9
30	302.4536	302	9
29	300.5033	301	10
28	298.5523	299	10
27	296.6008	297	10
26	294.6469	295	10
25	292.6859	293	10
24	290.7165	291	10
23	288.7347	289	10
22	286.7246	287	10
21	284.6824	285	11
20	282.6009	283	11
19	280.4739	280	11
18	278.2899	278	11
17	276.0351	276	11
16	273.6952	275	12
15	271.2244	275	12
14	268.6210	275	12
13	265.8596	275	12
12	262.9059	275	12
11	259.7195	275	12
10	256.2520	275	12
9	252.4258	275	12
8	248.1228	275	12
7	243.2505	275	12
6	237.6358	275	12
5	231.0131	275	12
4	222.9459	275	12
3	212.6088	275	12
2	198.1637	275	12
1	173.7346	275	12
0	147.2238	275	12

Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	491.6385	450	19
78	468.3069	450	19
77	454.5011	450	19
76	444.5600	445	17
75	436.7462	437	16
74	430.2582	430	14
73	424.6848	425(Advanced)	13
72	419.7787	420	12
71	415.3758	415	12
70	411.3751	411	11
69	407.6931	408	11
68	404.2742	404	10
67	401.0758	401	10
66	398.0619	398	10
65	395.2042	395	9
64	392.4841	392	9
63	389.8839	390	9
62	387.3873	387	9
61	384.9824	385	9
60	382.6577	383	9
59	380.4058	380(Proficient)	8
58	378.2171	378	8
57	376.0858	376	8
56	374.0061	374	8
55	371.9716	372	8
54	369.9789	370	8
53	368.0215	368	8
52	366.0950	366	8
51	364.2012	364	8
50	362.3332	362	8
49	360.4863	360	8
48	358.6603	359	8
47	356.8508	357	8
46	355.0571	355	8
45	353.2758	353	8
44	351.5033	352	8
43	349.7394	350(Pass)	8
42	347.9833	348	8
41	346.2280	346	8
40	344.4758	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.7237	343	8
38	340.9684	341	8
37	339.2098	339	8
36	337.4421	337	8
35	335.6674	336	8
34	333.8807	334	8
33	332.0833	332	8
32	330.2684	330	8
31	328.4364	328	8
30	326.5831	327	8
29	324.7046	325	8
28	322.8014	323	8
27	320.8687	321	8
26	318.9009	319	8
25	316.8963	317	8
24	314.8495	315	8
23	312.7558	313	8
22	310.6107	311	8
21	308.4084	308	8
20	306.1418	306	9
19	303.8009	304	9
18	301.3778	301	9
17	298.8632	299	9
16	296.2420	296	9
15	293.5024	294	10
14	290.6250	291	10
13	287.5875	288	10
12	284.3639	284	10
11	280.9194	281	11
10	277.2100	277	11
9	273.1798	275	12
8	268.7520	275	12
7	263.8139	275	12
6	258.2085	275	12
5	251.6899	275	12
4	243.8409	275	12
3	233.8807	275	12
2	220.0521	275	12
1	196.7488	275	12
0	171.7079	275	12

Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	500.9245	450	23
88	477.9052	450	23
87	464.1691	450	23
86	454.0914	450	23
85	445.9298	446	21
84	438.9649	439	17
83	432.8239	433	17
82	427.3032	427	17
81	422.2651	422	15
80	417.6233	418	15
79	413.3140	413	14
78	409.2856	409	13
77	405.4995	405(Advanced)	13
76	401.9299	402	12
75	398.5433	399	12
74	395.3173	395	12
73	392.2328	392	11
72	389.2748	389	11
71	386.4285	386	11
70	383.6804	384	11
69	381.0190	381(Proficient)	10
68	378.4391	378	10
67	375.9297	376	10
66	373.4853	373	10
65	371.0983	371	10
64	368.7676	369	9
63	366.4824	366	9
62	364.2431	364	9
61	362.0410	362	9
60	359.8786	360	9
59	357.7500	358	9
58	355.6500	356	9
57	353.5770	354	9
56	351.5291	352	9
55	349.5025	350(Pass)	9
54	347.4975	347	9
53	345.5095	346	8
52	343.5380	344	8
51	341.5793	342	8
50	339.6315	340	8
49	337.6943	338	8
48	335.7642	336	8
47	333.8411	334	8
46	331.9248	332	8
45	330.0112	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.0996	328	8
43	326.1903	326	8
42	324.2802	324	8
41	322.3692	322	8
40	320.4562	320	8
39	318.5419	319	9
38	316.6237	317	9
37	314.7026	315	9
36	312.7777	313	9
35	310.8492	311	9
34	308.9177	309	9
33	306.9815	307	9
32	305.0430	305	9
31	303.1006	303	9
30	301.1580	301	9
29	299.2095	299	9
28	297.2569	297	9
27	295.3010	295	10
26	293.3368	293	10
25	291.3654	291	10
24	289.3851	289	10
23	287.3892	287	10
22	285.3746	285	10
21	283.3365	283	10
20	281.2688	281	11
19	279.1683	279	11
18	276.9968	277	11
17	274.7648	275	11
16	272.4601	275	11
15	270.0670	275	11
14	267.5639	275	11
13	264.8845	275	11
12	262.0321	275	11
11	258.9710	275	11
10	255.6542	275	11
9	252.0181	275	11
8	247.9336	275	11
7	243.2761	275	11
6	237.8968	275	11
5	231.5246	275	11
4	223.7133	275	11
3	213.6229	275	11
2	199.3811	275	11
1	175.0060	275	11
0	147.2238	275	11

Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	493.0270	450	19
78	469.9059	450	19
77	456.1800	450	19
76	446.2788	446	17
75	438.4729	438	16
74	431.9894	432	14
73	426.4150	426	13
72	421.5039	422(Advanced)	12
71	417.0990	417	12
70	413.0873	413	11
69	409.3990	409	11
68	405.9714	406	10
67	402.7651	403	10
66	399.7449	400	10
65	396.8819	397	9
64	394.1537	394	9
63	391.5464	392	9
62	389.0491	389	9
61	386.6404	387	9
60	384.3142	384	9
59	382.0585	382	8
58	379.8690	380(Proficient)	8
57	377.7346	378	8
56	375.6532	376	8
55	373.6173	374	8
54	371.6225	372	8
53	369.6656	370	8
52	367.7386	368	8
51	365.8423	366	8
50	363.9744	364	8
49	362.1288	362	8
48	360.3024	360	8
47	358.4942	358	8
46	356.7005	357	8
45	354.9203	355	8
44	353.1504	353	8
43	351.3872	351	8
42	349.6312	350(Pass)	8
41	347.8805	348	8
40	346.1291	346	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	344.3788	344	8
38	342.6265	343	8
37	340.8695	341	8
36	339.1072	339	8
35	337.3333	337	8
34	335.5508	336	8
33	333.7538	334	8
32	331.9441	332	8
31	330.1140	330	8
30	328.2646	328	8
29	326.3908	326	8
28	324.4880	324	8
27	322.5586	323	8
26	320.5952	321	8
25	318.5927	319	8
24	316.5495	317	8
23	314.4584	314	8
22	312.3153	312	8
21	310.1152	310	8
20	307.8506	308	9
19	305.5138	306	9
18	303.0922	303	9
17	300.5804	301	9
16	297.9619	298	9
15	295.2243	295	10
14	292.3506	292	10
13	289.3169	289	10
12	286.0993	286	10
11	282.6626	283	11
10	278.9656	279	11
9	274.9320	275	12
8	270.5005	275	12
7	265.5568	275	12
6	259.9394	275	12
5	253.4030	275	12
4	245.5259	275	12
3	235.5174	275	12
2	221.5975	275	12
1	198.0754	275	12
0	171.7079	275	12

Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	504.1168	450	23
88	482.5959	450	23
87	468.3464	450	23
86	457.8910	450	23
85	449.3813	449	21
84	442.0909	442	17
83	435.6489	436	17
82	429.8579	430	17
81	424.5779	425	16
80	419.7191	420	15
79	415.2153	415	14
78	411.0111	411	14
77	407.0598	407	13
76	403.3386	403(Advanced)	13
75	399.8126	400	12
74	396.4586	396	12
73	393.2520	393	11
72	390.1772	390	11
71	387.2208	387	11
70	384.3679	384	11
69	381.6082	382(Proficient)	10
68	378.9339	379	10
67	376.3373	376	10
66	373.8116	374	10
65	371.3490	371	10
64	368.9480	369	10
63	366.5989	367	9
62	364.3007	364	9
61	362.0453	362	9
60	359.8353	360	9
59	357.6640	358	9
58	355.5261	356	9
57	353.4206	353	9
56	351.3432	351(Pass)	9
55	349.2914	349	9
54	347.2660	347	9
53	345.2625	345	9
52	343.2788	343	8
51	341.3129	341	8
50	339.3613	339	8
49	337.4258	337	8
48	335.5005	336	8
47	333.5880	334	8
46	331.6863	332	8
45	329.7926	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	327.9061	328	8
43	326.0266	326	8
42	324.1515	324	8
41	322.2808	322	8
40	320.4139	320	9
39	318.5507	319	9
38	316.6887	317	9
37	314.8289	315	9
36	312.9693	313	9
35	311.1100	311	9
34	309.2501	309	9
33	307.3875	307	9
32	305.5222	306	9
31	303.6522	304	9
30	301.7779	302	9
29	299.8949	300	9
28	298.0001	298	9
27	296.0925	296	10
26	294.1684	294	10
25	292.2236	292	10
24	290.2559	290	10
23	288.2594	288	10
22	286.2257	286	10
21	284.1512	284	10
20	282.0140	282	11
19	279.8166	280	11
18	277.5466	278	11
17	275.1930	275	11
16	272.7428	275	11
15	270.1792	275	11
14	267.4815	275	11
13	264.5949	275	11
12	261.5084	275	11
11	258.1901	275	11
10	254.5933	275	11
9	250.6539	275	11
8	246.2876	275	11
7	241.3807	275	11
6	235.7697	275	11
5	229.2013	275	11
4	221.2028	275	11
3	210.9564	275	11
2	196.6700	275	11
1	172.5312	275	11
0	147.2238	275	11

Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.1148	450	19
78	468.8055	450	19
77	454.9757	450	19
76	445.0044	445	17
75	437.1450	437	16
74	430.6180	431	14
73	425.0073	425(Advanced)	13
72	420.0660	420	12
71	415.6306	416	12
70	411.5996	412	11
69	407.8898	408	11
68	404.4450	404	10
67	401.2229	401	10
66	398.1875	398	10
65	395.3101	395	9
64	392.5717	393	9
63	389.9551	390	9
62	387.4435	387	9
61	385.0249	385	9
60	382.6880	383	9
59	380.4250	380(Proficient)	8
58	378.2264	378	8
57	376.0862	376	8
56	373.9986	374	8
55	371.9573	372	8
54	369.9586	370	8
53	367.9959	368	8
52	366.0650	366	8
51	364.1677	364	8
50	362.2968	362	8
49	360.4478	360	8
48	358.6202	359	8
47	356.8099	357	8
46	355.0158	355	8
45	353.2348	353	8
44	351.4637	351	8
43	349.7006	350(Pass)	8
42	347.9463	348	8
41	346.1934	346	8
40	344.4441	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.6952	343	8
38	340.9438	341	8
37	339.1895	339	8
36	337.4263	337	8
35	335.6568	336	8
34	333.8756	334	8
33	332.0842	332	8
32	330.2756	330	8
31	328.4503	328	8
30	326.6041	327	8
29	324.7331	325	8
28	322.8382	323	8
27	320.9136	321	8
26	318.9546	319	8
25	316.9591	317	8
24	314.9220	315	8
23	312.8382	313	8
22	310.7035	311	8
21	308.5122	309	8
20	306.2569	306	9
19	303.9281	304	9
18	301.5170	302	9
17	299.0155	299	9
16	296.4075	296	9
15	293.6817	294	9
14	290.8189	291	10
13	287.7963	288	10
12	284.5883	285	10
11	281.1599	281	11
10	277.4674	277	11
9	273.4541	275	12
8	269.0440	275	12
7	264.1240	275	12
6	258.5364	275	12
5	252.0351	275	12
4	244.2019	275	12
3	234.2533	275	12
2	220.4258	275	12
1	197.0894	275	12
0	171.7079	275	12

Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	17
89	516.1457	450	17
88	503.8915	450	17
87	488.0997	450	17
86	473.0338	450	17
85	461.1005	450	17
84	451.1655	450	17
83	442.7209	443	17
82	435.4468	435	17
81	429.0961	429	16
80	423.4993	423	16
79	418.4746	418	15
78	413.8978	414	14
77	409.6821	410	13
76	405.7618	406(Advanced)	13
75	402.0950	402	12
74	398.6343	399	12
73	395.3486	395	12
72	392.2134	392	11
71	389.2101	389	11
70	386.3214	386	11
69	383.5326	384	10
68	380.8312	381(Proficient)	10
67	378.2125	378	10
66	375.6638	376	10
65	373.1804	373	10
64	370.7546	371	10
63	368.3862	368	9
62	366.0644	366	9
61	363.7896	364	9
60	361.5520	362	9
59	359.3533	359	9
58	357.1867	357	9
57	355.0482	355	9
56	352.9385	353	9
55	350.8516	351(Pass)	9
54	348.7872	349	9
53	346.7419	347	9
52	344.7147	345	8
51	342.7004	343	8
50	340.6998	341	8
49	338.7065	339	8
48	336.7253	337	8
47	334.7461	335	8
46	332.7735	333	8
45	330.8027	331	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.8307	329	8
43	326.8581	327	8
42	324.8807	325	8
41	322.8980	323	8
40	320.9066	321	8
39	318.9087	319	8
38	316.8967	317	9
37	314.8731	315	9
36	312.8338	313	9
35	310.7777	311	9
34	308.7037	309	9
33	306.6048	307	9
32	304.4828	304	9
31	302.3351	302	9
30	300.1605	300	9
29	297.9534	298	9
28	295.7139	296	9
27	293.4363	293	9
26	291.1190	291	10
25	288.7572	289	10
24	286.3380	286	10
23	283.8629	284	10
22	281.3223	281	10
21	278.7118	279	10
20	276.0163	276	11
19	273.2259	275	11
18	270.3242	275	11
17	267.2899	275	11
16	264.1011	275	11
15	260.7079	275	11
14	257.0821	275	11
13	253.1745	275	11
12	248.9127	275	11
11	244.2075	275	11
10	238.9374	275	11
9	232.9361	275	11
8	225.9746	275	11
7	217.7421	275	11
6	207.8858	275	11
5	196.3085	275	11
4	180.1904	275	11
3	167.4042	275	11
2	158.1776	275	11
1	151.7387	275	11
0	147.2238	275	11

Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	494.4096	450	19
78	471.8821	450	19
77	458.1559	450	19
76	448.2663	448	17
75	440.4690	440	16
74	433.9932	434	14
73	428.4233	428	13
72	423.5147	424(Advanced)	12
71	419.1109	419	12
70	415.0994	415	11
69	411.4098	411	11
68	407.9811	408	10
67	404.7709	405	10
66	401.7469	402	10
65	398.8809	399	9
64	396.1496	396	9
63	393.5362	394	9
62	391.0281	391	9
61	388.6130	389	9
60	386.2768	386	9
59	384.0138	384	8
58	381.8145	382	8
57	379.6763	380(Proficient)	8
56	377.5857	378	8
55	375.5423	376	8
54	373.5386	374	8
53	371.5709	372	8
52	369.6365	370	8
51	367.7280	368	8
50	365.8458	366	8
49	363.9882	364	8
48	362.1496	362	8
47	360.3269	360	8
46	358.5191	359	8
45	356.7229	357	8
44	354.9371	355	8
43	353.1582	353	8
42	351.3847	351	8
41	349.6142	350(Pass)	8
40	347.8466	348	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	346.0754	346	8
38	344.3025	344	8
37	342.5246	343	8
36	340.7391	341	8
35	338.9450	339	8
34	337.1367	337	8
33	335.3165	335	8
32	333.4783	333	8
31	331.6245	332	8
30	329.7456	330	8
29	327.8434	328	8
28	325.9121	326	8
27	323.9478	324	8
26	321.9527	322	8
25	319.9164	320	8
24	317.8352	318	8
23	315.7073	316	8
22	313.5230	314	8
21	311.2795	311	9
20	308.9697	309	9
19	306.5848	307	9
18	304.1144	304	9
17	301.5470	302	9
16	298.8729	299	9
15	296.0739	296	10
14	293.1352	293	10
13	290.0328	290	10
12	286.7411	287	11
11	283.2257	283	11
10	279.4445	279	11
9	275.3404	275	12
8	270.8212	275	12
7	265.7688	275	12
6	260.0338	275	12
5	253.3713	275	12
4	245.3601	275	12
3	235.2133	275	12
2	221.1620	275	12
1	197.5615	275	12
0	171.7079	275	12

Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	505.7249	450	23
88	486.0011	450	23
87	471.4952	450	23
86	460.9790	450	23
85	452.4751	450	23
84	445.1962	445	21
83	438.7620	439	19
82	432.9718	433	18
81	427.6884	428	16
80	422.8215	423	15
79	418.3064	418	14
78	414.0919	414	14
77	410.1334	410	13
76	406.3944	406	13
75	402.8546	403(Advanced)	12
74	399.4827	399	12
73	396.2581	396	12
72	393.1637	393	11
71	390.1859	390	11
70	387.3132	387	11
69	384.5321	385	10
68	381.8339	382(Proficient)	10
67	379.2106	379	10
66	376.6577	377	10
65	374.1668	374	10
64	371.7322	372	10
63	369.3516	369	9
62	367.0184	367	9
61	364.7292	365	9
60	362.4788	362	9
59	360.2670	360	9
58	358.0892	358	9
57	355.9408	356	9
56	353.8221	354	9
55	351.7285	352	9
54	349.6580	350(Pass)	9
53	347.6105	348	9
52	345.5811	346	9
51	343.5695	344	9
50	341.5722	342	9
49	339.5874	340	8
48	337.6152	338	8
47	335.6516	336	8
46	333.6977	334	8
45	331.7528	332	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	329.8133	330	8
43	327.8793	328	8
42	325.9502	326	9
41	324.0237	324	9
40	322.0997	322	9
39	320.1784	320	9
38	318.2586	318	9
37	316.3402	316	9
36	314.4220	314	9
35	312.5048	313	9
34	310.5870	311	9
33	308.6697	309	9
32	306.7492	307	9
31	304.8274	305	9
30	302.9017	303	9
29	300.9745	301	9
28	299.0392	299	9
27	297.0954	297	10
26	295.1427	295	10
25	293.1737	293	10
24	291.1850	291	10
23	289.1744	289	10
22	287.1325	287	10
21	285.0540	285	10
20	282.9313	283	11
19	280.7571	281	11
18	278.5221	279	11
17	276.2123	276	11
16	273.8087	275	12
15	271.2823	275	12
14	268.6243	275	12
13	265.8112	275	12
12	262.8109	275	12
11	259.5863	275	12
10	256.0921	275	12
9	252.2667	275	12
8	247.9767	275	12
7	243.1353	275	12
6	237.5757	275	12
5	231.0340	275	12
4	223.0729	275	12
3	212.8633	275	12
2	198.5531	275	12
1	174.2182	275	12
0	147.2238	275	12

Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.7287	450	19
78	469.5120	450	19
77	455.7146	450	19
76	445.7533	446	18
75	437.8935	438	17
74	431.3601	431	15
73	425.7393	426(Advanced)	14
72	420.7850	421	13
71	416.3354	416	12
70	412.2880	412	11
69	408.5617	409	11
68	405.0988	405	10
67	401.8583	402	10
66	398.8043	399	10
65	395.9084	396	10
64	393.1500	393	9
63	390.5134	391	9
62	387.9827	388	9
61	385.5437	386	9
60	383.1871	383	9
59	380.9033	381(Proficient)	9
58	378.6850	379	8
57	376.5239	377	8
56	374.4161	374	8
55	372.3540	372	8
54	370.3343	370	8
53	368.3516	368	8
52	366.3994	366	8
51	364.4805	364	8
50	362.5890	363	8
49	360.7183	361	8
48	358.8694	359	8
47	357.0372	357	8
46	355.2211	355	8
45	353.4180	353	8
44	351.6241	352	8
43	349.8384	350(Pass)	8
42	348.0614	348	8
41	346.2850	346	8
40	344.5119	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.7389	343	8
38	340.9629	341	8
37	339.1836	339	8
36	337.3949	337	8
35	335.5995	336	8
34	333.7918	334	8
33	331.9735	332	8
32	330.1371	330	8
31	328.2836	328	8
30	326.4083	326	8
29	324.5074	325	8
28	322.5820	323	8
27	320.6261	321	8
26	318.6345	319	8
25	316.6060	317	8
24	314.5340	315	8
23	312.4151	312	8
22	310.2439	310	8
21	308.0148	308	9
20	305.7204	306	9
19	303.3500	303	9
18	300.8981	301	9
17	298.3519	298	9
16	295.6992	296	9
15	292.9267	293	10
14	290.0150	290	10
13	286.9422	287	10
12	283.6817	284	10
11	280.1993	280	11
10	276.4504	276	11
9	272.3807	275	12
8	267.9119	275	12
7	262.9319	275	12
6	257.2863	275	12
5	250.7292	275	12
4	242.8466	275	12
3	232.8645	275	12
2	219.0428	275	12
1	195.8376	275	12
0	171.7079	275	12

Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	506.8227	450	23
88	488.4221	450	23
87	473.5652	450	23
86	462.7395	450	23
85	453.8958	450	23
84	446.2486	446	22
83	439.4410	439	20
82	433.2903	433	18
81	427.6741	428	17
80	422.5059	423	16
79	417.7199	418	15
78	413.2604	413	14
77	409.0750	409	13
76	405.1341	405(Advanced)	13
75	401.4016	401	12
74	397.8482	398	12
73	394.4515	394	12
72	391.1940	391	11
71	388.0622	388	11
70	385.0434	385	11
69	382.1302	382(Proficient)	11
68	379.3040	379	10
67	376.5603	377	10
66	373.8897	374	10
65	371.2850	371	10
64	368.7448	369	10
63	366.2582	366	10
62	363.8295	364	9
61	361.4460	361	9
60	359.1077	359	9
59	356.8074	357	9
58	354.5424	355	9
57	352.3100	352	9
56	350.1046	350(Pass)	9
55	347.9315	348	9
54	345.7818	346	9
53	343.6549	344	9
52	341.5460	342	9
51	339.4528	339	9
50	337.3760	337	9
49	335.3093	335	9
48	333.2566	333	9
47	331.2139	331	9
46	329.1793	329	9
45	327.1506	327	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	325.1270	325	9
43	323.1075	323	9
42	321.0901	321	9
41	319.0785	319	9
40	317.0651	317	9
39	315.0554	315	9
38	313.0452	313	9
37	311.0364	311	9
36	309.0286	309	9
35	307.0205	307	9
34	305.0141	305	9
33	303.0069	303	9
32	301.0019	301	9
31	298.9948	299	9
30	296.9874	297	9
29	294.9800	295	10
28	292.9690	293	10
27	290.9556	291	10
26	288.9351	289	10
25	286.8990	287	10
24	284.8481	285	10
23	282.7769	283	10
22	280.6828	281	11
21	278.5602	279	11
20	276.4018	276	11
19	274.1750	275	11
18	271.8849	275	11
17	269.5249	275	11
16	267.0837	275	11
15	264.5223	275	11
14	261.8040	275	11
13	258.9375	275	11
12	255.8987	275	11
11	252.6267	275	11
10	249.0173	275	11
9	245.0706	275	11
8	240.7109	275	11
7	235.8431	275	11
6	230.1796	275	11
5	223.4448	275	11
4	215.2460	275	11
3	204.7121	275	11
2	189.7565	275	11
1	168.6183	275	11
0	147.2238	275	11

Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	18
79	494.0966	450	18
78	471.1887	450	18
77	457.4763	450	18
76	447.5580	448	18
75	439.7210	440	17
74	433.1996	433	15
73	427.5828	428	14
72	422.6264	423(Advanced)	13
71	418.1751	418	12
70	414.1155	414	11
69	410.3790	410	11
68	406.9026	407	10
67	403.6469	404	10
66	400.5776	401	10
65	397.6660	398	10
64	394.8893	395	9
63	392.2322	392	9
62	389.6834	390	9
61	387.2246	387	9
60	384.8478	385	9
59	382.5428	383	9
58	380.3035	380(Proficient)	8
57	378.1209	378	8
56	375.9907	376	8
55	373.9071	374	8
54	371.8646	372	8
53	369.8602	370	8
52	367.8866	368	8
51	365.9433	366	8
50	364.0289	364	8
49	362.1371	362	8
48	360.2647	360	8
47	358.4106	358	8
46	356.5713	357	8
45	354.7457	355	8
44	352.9299	353	8
43	351.1226	351(Pass)	8
42	349.3214	349	8
41	347.5252	348	8
40	345.7292	346	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.9343	344	8
38	342.1372	342	8
37	340.3359	340	8
36	338.5283	339	8
35	336.7108	337	8
34	334.8839	335	8
33	333.0436	333	8
32	331.1896	331	8
31	329.3158	329	8
30	327.4226	327	8
29	325.5043	326	8
28	323.5586	324	8
27	321.5861	322	8
26	319.5781	320	8
25	317.5320	318	8
24	315.4450	315	8
23	313.3096	313	8
22	311.1230	311	8
21	308.8791	309	9
20	306.5703	307	9
19	304.1878	304	9
18	301.7212	302	9
17	299.1639	299	9
16	296.4987	296	9
15	293.7147	294	10
14	290.7924	291	10
13	287.7092	288	10
12	284.4393	284	10
11	280.9477	281	11
10	277.1904	277	11
9	273.1116	275	12
8	268.6344	275	12
7	263.6463	275	12
6	257.9909	275	12
5	251.4226	275	12
4	243.5255	275	12
3	233.5220	275	12
2	219.6624	275	12
1	196.3690	275	12
0	171.7079	275	12

Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	24
89	497.2633	450	24
88	473.8990	450	24
87	460.1413	450	24
86	450.1143	450	24
85	442.0285	442	22
84	435.1318	435	20
83	429.0624	429	18
82	423.6111	424	17
81	418.6410	419	15
80	414.0567	414	14
79	409.7956	410	14
78	405.8103	406(Advanced)	13
77	402.0737	402	13
76	398.5462	399	12
75	395.2019	395	12
74	392.0192	392	11
73	388.9806	389	11
72	386.0682	386	11
71	383.2675	383	11
70	380.5641	381(Proficient)	10
69	377.9537	378	10
68	375.4202	375	10
67	372.9601	373	10
66	370.5635	371	10
65	368.2320	368	9
64	365.9489	366	9
63	363.7160	364	9
62	361.5241	362	9
61	359.3743	359	9
60	357.2595	357	9
59	355.1767	355	9
58	353.1241	353	9
57	351.0964	351(Pass)	9
56	349.0948	349	9
55	347.1157	347	9
54	345.1567	345	8
53	343.2150	343	8
52	341.2890	341	8
51	339.3751	339	8
50	337.4749	337	8
49	335.5828	336	8
48	333.7006	334	8
47	331.8269	332	8
46	329.9582	330	8
45	328.0937	328	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	326.2338	326	8
43	324.3752	324	8
42	322.5184	323	8
41	320.6615	321	8
40	318.8065	319	8
39	316.9486	317	8
38	315.0913	315	9
37	313.2312	313	9
36	311.3699	311	9
35	309.5065	310	9
34	307.6405	308	9
33	305.7720	306	9
32	303.9009	304	9
31	302.0280	302	9
30	300.1518	300	9
29	298.2699	298	9
28	296.3831	296	9
27	294.4897	294	10
26	292.5869	293	10
25	290.6746	291	10
24	288.7504	289	10
23	286.7979	287	10
22	284.8185	285	10
21	282.8069	283	10
20	280.7599	281	11
19	278.6702	279	11
18	276.5256	277	11
17	274.2892	275	11
16	271.9708	275	11
15	269.5539	275	11
14	267.0200	275	11
13	264.3343	275	11
12	261.4335	275	11
11	258.3160	275	11
10	254.9335	275	11
9	251.2218	275	11
8	247.0943	275	11
7	242.4222	275	11
6	236.9785	275	11
5	230.5484	275	11
4	222.6926	275	11
3	212.5789	275	11
2	198.3551	275	11
1	174.0991	275	11
0	147.2238	275	11

Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.1099	450	19
78	468.8545	450	19
77	455.0811	450	19
76	445.1633	445	18
75	437.3538	437	17
74	430.8732	431	15
73	425.3055	425(Advanced)	14
72	420.4037	420	13
71	416.0081	416	12
70	412.0082	412	11
69	408.3305	408	11
68	404.9151	405	10
67	401.7205	402	10
66	398.7112	399	10
65	395.8586	396	9
64	393.1425	393	9
63	390.5468	391	9
62	388.0562	388	9
61	385.6563	386	9
60	383.3380	383	9
59	381.0913	381(Proficient)	8
58	378.9098	379	8
57	376.7841	377	8
56	374.7116	375	8
55	372.6840	373	8
54	370.6980	371	8
53	368.7495	369	8
52	366.8296	367	8
51	364.9426	365	8
50	363.0844	363	8
49	361.2458	361	8
48	359.4290	359	8
47	357.6283	358	8
46	355.8432	356	8
45	354.0708	354	8
44	352.3081	352	8
43	350.5521	351(Pass)	8
42	348.8040	349	8
41	347.0584	347	8
40	345.3140	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.5701	344	8
38	341.8235	342	8
37	340.0727	340	8
36	338.3147	338	8
35	336.5472	337	8
34	334.7696	335	8
33	332.9787	333	8
32	331.1735	331	8
31	329.3484	329	8
30	327.5037	328	8
29	325.6338	326	8
28	323.7358	324	8
27	321.8113	322	8
26	319.8511	320	8
25	317.8520	318	8
24	315.8127	316	8
23	313.7239	314	8
22	311.5839	312	8
21	309.3863	309	8
20	307.1236	307	9
19	304.7881	305	9
18	302.3658	302	9
17	299.8548	300	9
16	297.2339	297	9
15	294.4941	294	9
14	291.6159	292	10
13	288.5755	289	10
12	285.3482	285	10
11	281.8977	282	11
10	278.1809	278	11
9	274.1394	275	12
8	269.6984	275	12
7	264.7439	275	12
6	259.1167	275	12
5	252.5708	275	12
4	244.6867	275	12
3	234.6789	275	12
2	220.7802	275	12
1	197.3483	275	12
0	171.7079	275	12

Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA— July, October and November 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	498.8256	450	23
88	475.6605	450	23
87	462.1805	450	23
86	452.5208	450	23
85	444.8414	445	21
84	438.3731	438	17
83	432.7140	433	17
82	427.6419	428	17
81	423.0084	423	15
80	418.7227	419	14
79	414.7220	415	14
78	410.9779	411	13
77	407.4266	407	13
76	404.0518	404(Advanced)	12
75	400.8262	401	12
74	397.7318	398	12
73	394.7539	395	11
72	391.8804	392	11
71	389.1015	389	11
70	386.4064	386	11
69	383.7861	384	10
68	381.2321	381(Proficient)	10
67	378.7391	379	10
66	376.3052	376	10
65	373.9195	374	10
64	371.5787	372	10
63	369.2817	369	9
62	367.0219	367	9
61	364.7967	365	9
60	362.6019	363	9
59	360.4369	360	9
58	358.2973	358	9
57	356.1817	356	9
56	354.0876	354	9
55	352.0123	352	9
54	349.9528	350(Pass)	9
53	347.9124	348	9
52	345.8838	346	9
51	343.8688	344	9
50	341.8624	342	9
49	339.8647	340	9
48	337.8739	338	9
47	335.8887	336	9
46	333.9081	334	9
45	331.9330	332	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	329.9597	330	9
43	327.9881	328	9
42	326.0192	326	9
41	324.0499	324	9
40	322.0809	322	9
39	320.1131	320	9
38	318.1451	318	9
37	316.1787	316	9
36	314.2121	314	9
35	312.2483	312	9
34	310.2846	310	9
33	308.3235	308	9
32	306.3635	306	9
31	304.4072	304	9
30	302.4536	302	9
29	300.5033	301	10
28	298.5523	299	10
27	296.6008	297	10
26	294.6469	295	10
25	292.6859	293	10
24	290.7165	291	10
23	288.7347	289	10
22	286.7246	287	10
21	284.6824	285	11
20	282.6009	283	11
19	280.4739	280	11
18	278.2899	278	11
17	276.0351	276	11
16	273.6952	275	12
15	271.2244	275	12
14	268.6210	275	12
13	265.8596	275	12
12	262.9059	275	12
11	259.7195	275	12
10	256.2520	275	12
9	252.4258	275	12
8	248.1228	275	12
7	243.2505	275	12
6	237.6358	275	12
5	231.0131	275	12
4	222.9459	275	12
3	212.6088	275	12
2	198.1637	275	12
1	173.7346	275	12
0	147.2238	275	12

Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July, October and November 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	491.6385	450	19
78	468.3069	450	19
77	454.5011	450	19
76	444.5600	445	17
75	436.7462	437	16
74	430.2582	430	14
73	424.6848	425(Advanced)	13
72	419.7787	420	12
71	415.3758	415	12
70	411.3751	411	11
69	407.6931	408	11
68	404.2742	404	10
67	401.0758	401	10
66	398.0619	398	10
65	395.2042	395	9
64	392.4841	392	9
63	389.8839	390	9
62	387.3873	387	9
61	384.9824	385	9
60	382.6577	383	9
59	380.4058	380(Proficient)	8
58	378.2171	378	8
57	376.0858	376	8
56	374.0061	374	8
55	371.9716	372	8
54	369.9789	370	8
53	368.0215	368	8
52	366.0950	366	8
51	364.2012	364	8
50	362.3332	362	8
49	360.4863	360	8
48	358.6603	359	8
47	356.8508	357	8
46	355.0571	355	8
45	353.2758	353	8
44	351.5033	352	8
43	349.7394	350(Pass)	8
42	347.9833	348	8
41	346.2280	346	8
40	344.4758	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.7237	343	8
38	340.9684	341	8
37	339.2098	339	8
36	337.4421	337	8
35	335.6674	336	8
34	333.8807	334	8
33	332.0833	332	8
32	330.2684	330	8
31	328.4364	328	8
30	326.5831	327	8
29	324.7046	325	8
28	322.8014	323	8
27	320.8687	321	8
26	318.9009	319	8
25	316.8963	317	8
24	314.8495	315	8
23	312.7558	313	8
22	310.6107	311	8
21	308.4084	308	8
20	306.1418	306	9
19	303.8009	304	9
18	301.3778	301	9
17	298.8632	299	9
16	296.2420	296	9
15	293.5024	294	10
14	290.6250	291	10
13	287.5875	288	10
12	284.3639	284	10
11	280.9194	281	11
10	277.2100	277	11
9	273.1798	275	12
8	268.7520	275	12
7	263.8139	275	12
6	258.2085	275	12
5	251.6899	275	12
4	243.8409	275	12
3	233.8807	275	12
2	220.0521	275	12
1	196.7488	275	12
0	171.7079	275	12

Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—December 2011 and February 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	17
89	516.1457	450	17
88	503.8915	450	17
87	488.0997	450	17
86	473.0338	450	17
85	461.1005	450	17
84	451.1655	450	17
83	442.7209	443	17
82	435.4468	435	17
81	429.0961	429	16
80	423.4993	423	16
79	418.4746	418	15
78	413.8978	414	14
77	409.6821	410	13
76	405.7618	406(Advanced)	13
75	402.0950	402	12
74	398.6343	399	12
73	395.3486	395	12
72	392.2134	392	11
71	389.2101	389	11
70	386.3214	386	11
69	383.5326	384	10
68	380.8312	381(Proficient)	10
67	378.2125	378	10
66	375.6638	376	10
65	373.1804	373	10
64	370.7546	371	10
63	368.3862	368	9
62	366.0644	366	9
61	363.7896	364	9
60	361.5520	362	9
59	359.3533	359	9
58	357.1867	357	9
57	355.0482	355	9
56	352.9385	353	9
55	350.8516	351(Pass)	9
54	348.7872	349	9
53	346.7419	347	9
52	344.7147	345	8
51	342.7004	343	8
50	340.6998	341	8
49	338.7065	339	8
48	336.7253	337	8
47	334.7461	335	8
46	332.7735	333	8
45	330.8027	331	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.8307	329	8
43	326.8581	327	8
42	324.8807	325	8
41	322.8980	323	8
40	320.9066	321	8
39	318.9087	319	8
38	316.8967	317	9
37	314.8731	315	9
36	312.8338	313	9
35	310.7777	311	9
34	308.7037	309	9
33	306.6048	307	9
32	304.4828	304	9
31	302.3351	302	9
30	300.1605	300	9
29	297.9534	298	9
28	295.7139	296	9
27	293.4363	293	9
26	291.1190	291	10
25	288.7572	289	10
24	286.3380	286	10
23	283.8629	284	10
22	281.3223	281	10
21	278.7118	279	10
20	276.0163	276	11
19	273.2259	275	11
18	270.3242	275	11
17	267.2899	275	11
16	264.1011	275	11
15	260.7079	275	11
14	257.0821	275	11
13	253.1745	275	11
12	248.9127	275	11
11	244.2075	275	11
10	238.9374	275	11
9	232.9361	275	11
8	225.9746	275	11
7	217.7421	275	11
6	207.8858	275	11
5	196.3085	275	11
4	180.1904	275	11
3	167.4042	275	11
2	158.1776	275	11
1	151.7387	275	11
0	147.2238	275	11

Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— December 2011 and February 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	494.4096	450	19
78	471.8821	450	19
77	458.1559	450	19
76	448.2663	448	17
75	440.4690	440	16
74	433.9932	434	14
73	428.4233	428	13
72	423.5147	424(Advanced)	12
71	419.1109	419	12
70	415.0994	415	11
69	411.4098	411	11
68	407.9811	408	10
67	404.7709	405	10
66	401.7469	402	10
65	398.8809	399	9
64	396.1496	396	9
63	393.5362	394	9
62	391.0281	391	9
61	388.6130	389	9
60	386.2768	386	9
59	384.0138	384	8
58	381.8145	382	8
57	379.6763	380(Proficient)	8
56	377.5857	378	8
55	375.5423	376	8
54	373.5386	374	8
53	371.5709	372	8
52	369.6365	370	8
51	367.7280	368	8
50	365.8458	366	8
49	363.9882	364	8
48	362.1496	362	8
47	360.3269	360	8
46	358.5191	359	8
45	356.7229	357	8
44	354.9371	355	8
43	353.1582	353	8
42	351.3847	351	8
41	349.6142	350(Pass)	8
40	347.8466	348	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	346.0754	346	8
38	344.3025	344	8
37	342.5246	343	8
36	340.7391	341	8
35	338.9450	339	8
34	337.1367	337	8
33	335.3165	335	8
32	333.4783	333	8
31	331.6245	332	8
30	329.7456	330	8
29	327.8434	328	8
28	325.9121	326	8
27	323.9478	324	8
26	321.9527	322	8
25	319.9164	320	8
24	317.8352	318	8
23	315.7073	316	8
22	313.5230	314	8
21	311.2795	311	9
20	308.9697	309	9
19	306.5848	307	9
18	304.1144	304	9
17	301.5470	302	9
16	298.8729	299	9
15	296.0739	296	10
14	293.1352	293	10
13	290.0328	290	10
12	286.7411	287	11
11	283.2257	283	11
10	279.4445	279	11
9	275.3404	275	12
8	270.8212	275	12
7	265.7688	275	12
6	260.0338	275	12
5	253.3713	275	12
4	245.3601	275	12
3	235.2133	275	12
2	221.1620	275	12
1	197.5615	275	12
0	171.7079	275	12

Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—March and May 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	506.8227	450	23
88	488.4221	450	23
87	473.5652	450	23
86	462.7395	450	23
85	453.8958	450	23
84	446.2486	446	22
83	439.4410	439	20
82	433.2903	433	18
81	427.6741	428	17
80	422.5059	423	16
79	417.7199	418	15
78	413.2604	413	14
77	409.0750	409	13
76	405.1341	405(Advanced)	13
75	401.4016	401	12
74	397.8482	398	12
73	394.4515	394	12
72	391.1940	391	11
71	388.0622	388	11
70	385.0434	385	11
69	382.1302	382(Proficient)	11
68	379.3040	379	10
67	376.5603	377	10
66	373.8897	374	10
65	371.2850	371	10
64	368.7448	369	10
63	366.2582	366	10
62	363.8295	364	9
61	361.4460	361	9
60	359.1077	359	9
59	356.8074	357	9
58	354.5424	355	9
57	352.3100	352	9
56	350.1046	350(Pass)	9
55	347.9315	348	9
54	345.7818	346	9
53	343.6549	344	9
52	341.5460	342	9
51	339.4528	339	9
50	337.3760	337	9
49	335.3093	335	9
48	333.2566	333	9
47	331.2139	331	9
46	329.1793	329	9
45	327.1506	327	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	325.1270	325	9
43	323.1075	323	9
42	321.0901	321	9
41	319.0785	319	9
40	317.0651	317	9
39	315.0554	315	9
38	313.0452	313	9
37	311.0364	311	9
36	309.0286	309	9
35	307.0205	307	9
34	305.0141	305	9
33	303.0069	303	9
32	301.0019	301	9
31	298.9948	299	9
30	296.9874	297	9
29	294.9800	295	10
28	292.9690	293	10
27	290.9556	291	10
26	288.9351	289	10
25	286.8990	287	10
24	284.8481	285	10
23	282.7769	283	10
22	280.6828	281	11
21	278.5602	279	11
20	276.4018	276	11
19	274.1750	275	11
18	271.8849	275	11
17	269.5249	275	11
16	267.0837	275	11
15	264.5223	275	11
14	261.8040	275	11
13	258.9375	275	11
12	255.8987	275	11
11	252.6267	275	11
10	249.0173	275	11
9	245.0706	275	11
8	240.7109	275	11
7	235.8431	275	11
6	230.1796	275	11
5	223.4448	275	11
4	215.2460	275	11
3	204.7121	275	11
2	189.7565	275	11
1	168.6183	275	11
0	147.2238	275	11

Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—March and May 2012

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	18
79	494.0966	450	18
78	471.1887	450	18
77	457.4763	450	18
76	447.5580	448	18
75	439.7210	440	17
74	433.1996	433	15
73	427.5828	428	14
72	422.6264	423(Advanced)	13
71	418.1751	418	12
70	414.1155	414	11
69	410.3790	410	11
68	406.9026	407	10
67	403.6469	404	10
66	400.5776	401	10
65	397.6660	398	10
64	394.8893	395	9
63	392.2322	392	9
62	389.6834	390	9
61	387.2246	387	9
60	384.8478	385	9
59	382.5428	383	9
58	380.3035	380(Proficient)	8
57	378.1209	378	8
56	375.9907	376	8
55	373.9071	374	8
54	371.8646	372	8
53	369.8602	370	8
52	367.8866	368	8
51	365.9433	366	8
50	364.0289	364	8
49	362.1371	362	8
48	360.2647	360	8
47	358.4106	358	8
46	356.5713	357	8
45	354.7457	355	8
44	352.9299	353	8
43	351.1226	351(Pass)	8
42	349.3214	349	8
41	347.5252	348	8
40	345.7292	346	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.9343	344	8
38	342.1372	342	8
37	340.3359	340	8
36	338.5283	339	8
35	336.7108	337	8
34	334.8839	335	8
33	333.0436	333	8
32	331.1896	331	8
31	329.3158	329	8
30	327.4226	327	8
29	325.5043	326	8
28	323.5586	324	8
27	321.5861	322	8
26	319.5781	320	8
25	317.5320	318	8
24	315.4450	315	8
23	313.3096	313	8
22	311.1230	311	8
21	308.8791	309	9
20	306.5703	307	9
19	304.1878	304	9
18	301.7212	302	9
17	299.1639	299	9
16	296.4987	296	9
15	293.7147	294	10
14	290.7924	291	10
13	287.7092	288	10
12	284.4393	284	10
11	280.9477	281	11
10	277.1904	277	11
9	273.1116	275	12
8	268.6344	275	12
7	263.6463	275	12
6	257.9909	275	12
5	251.4226	275	12
4	243.5255	275	12
3	233.5220	275	12
2	219.6624	275	12
1	196.3690	275	12
0	171.7079	275	12

Appendix 6.P: Standard Errors of Theta Based on Weighted Raw Scores

Let:

i = 1 represent dichotomous items (with scores U_j) scaled with Rasch model, with ICC $P_{j1}(\theta)$

i = 2 represent polytomous items (with scores Y_j) scaled with Rasch partial-credit model;

item j has m_j levels; score for k-th level is k-1, with ICC $P_{j2k}(\theta)$

w_{ji}= weight for the j-th item of type i

$$S = \text{rawscore} = \sum_{j=1}^{n_1} w_{j1} U_j + \sum_{j=1}^{n_2} w_{j2} Y_j$$

$$\hat{\theta} = \text{MLE of } \theta$$

$$\text{SEM}(\hat{\theta}) \approx [I(\theta, S)]^{-1/2}$$

$$I(\theta,S) = \frac{\left[\dfrac{d\,E(S\,|\,\theta)}{d\theta}\right]^2}{\sigma^2(S\,|\,\theta)}$$ see Lord (1980), pp. 67 and 73.

$$E(S\,|\,\theta) = \sum_{j=1}^{n_1} w_{j1} P_{j1}(\theta) + \sum_{j=1}^{n_2} w_{j2} \sum_{k=1}^{m_j} (k-1) P_{j2k}(\theta)$$

$$\frac{dE(S\,|\theta)}{d\theta} = \sum_{j=1}^{n_1} w_{j1} P'_{j1}(\theta) + \sum_{j=1}^{n_2} w_{j2} \sum_{k=1}^{m_j} (k-1) P'_{j2k}(\theta)$$

where

$$P'_{j1}(\theta) = P_{j1}(\theta)\left(1 - P_{j1}(\theta)\right)$$

$$P'_{j2k}(\theta) = P_{j2k}(\theta)\left[k - \sum_{r=1}^{m_j} r P_{j2r}(\theta)\right]$$

$$\sigma^2(S\,|\theta) = \sum_{j=1}^{n_1} \left(w_{j1}\right)^2 P_{j1}(\theta)\left(1 - P_{j1}(\theta)\right) + \sum_{j=1}^{n_2} \left(w_{j2}\right)^2 \sigma^2\left(Y_j\,|\theta\right)$$

$$\sigma^2\left(Y_j|\theta\right) = \left[\sum_{k=1}^{m_j} (k-1)^2 P_{j2k}(\theta)\right] - \left[\sum_{k=1}^{m_j} (k-1) P_{j2k}(\theta)\right]^2$$

Note also Lord (1980), Eq. (5-23) and Eq. (6-6) used for transforming the standard errors to other metrics.

Chapter 7: Performance Standards

Background and Procedures

The CAHSEE was offered for the first time in spring 2001 (March and May) to volunteer ninth-graders (class of 2004). At that time, the SBE set the passing score at 350 on a scale of 250 to 450 based on recommendations produced in a standard-setting workshop conducted by the American Institutes for Research (AIR) with test results from March 2001. In order to conduct a census testing of all tenth-graders, Assembly Bill (AB) 1609 was passed in October 2001 to prohibit ninth-graders from taking the CAHSEE. The SBE directed the CDE to conduct a standard-setting on the first census administration of the CAHSEE, which was in spring 2003.

Subsequently, in July 2003, the SBE made the passing of both the CAHSEE ELA and mathematics examinations a diploma requirement for the Class of 2006 and adopted revised test blueprints for the CAHSEE. In September 2003, a standard-setting workshop was performed using CAHSEE test items that were constructed to meet the new content blueprints and difficulty specifications. The Bookmark Method for setting passing scores was implemented. The Bookmark Method has widespread support in the measurement profession and has withstood legal challenges (see, for example, Lewis, et al., 1998 & Mitzel, et al., 2001).

The Bookmark Method typically uses three rounds of standard-setting in which panelists are instructed to set one passing score at the Just Sufficient level on the items reviewed (ELA or mathematics). During the standard-setting workshop, the panelists were asked to review the item booklet, in which items were ordered from easiest to hardest based on item difficulty, and then to find the point that defined the knowledge and skills needed to just pass the CAHSEE.

Two standard-setting panels for each subject were convened for the workshop. One panel in each subject was primarily composed of California educators (e.g., English or mathematics teachers), while the other panel was primarily composed of community members (e.g., business representatives, district or school administrators, and college professors). Panelists were selected based on their knowledge of the subject matter assessed, familiarity with students in the respective grade levels, an understanding of large-scale assessments, and an appreciation of the consequences of setting these passing scores. Panelists represented diverse geographic regions, major racial/ethnic subgroups, and both genders.

The standard-setting panelists first were given rigorous training, which included an overview of the CAHSEE and the Bookmark Method. Once panelists were comfortable with the procedure, they were asked to place a bookmark at the point in the ordered test book at which they felt students had demonstrated sufficient knowledge and skills in a certain subject area. Panelists typically placed the first bookmark independently and then received information on how their bookmark placement compared with those of their peers. There was then a small-group discussion followed by a second bookmark placement. Finally, there was a large-group discussion followed by a presentation of consequence or impact data—for

example, what percentage of students would pass the test given the current median passing score—followed by the third (last) bookmark placement.

CAHSEE Passing Scores

In November 2003, after reviewing the results of the September standard-setting study, the SBE set new passing scores on the CAHSEE. These scores corresponded to specific score levels on the test forms used in the standard settings: 55 percent correct on the mathematics portion (44 raw score points out of 80) and 60 percent of the points on the ELA portion (54 raw score points out of 90), corresponding to theta values of -0.0701 for mathematics and 0.5356 for ELA. Following the February 2004 administration, the cut-score theta levels were applied to determine the raw score equivalent of the CAHSEE passing scores and to establish the new CAHSEE reporting scale.

For the February 2004 administration, items were calibrated and linked to the CAHSEE item bank scale. IRT-based equating procedures were used to determine the theta levels corresponding to each raw score. The passing scores on the February test forms were set at the raw scores corresponding to the theta levels that were closest to the cut-score theta levels (0.5356 for ELA and -0.0701 for mathematics). Table 7.1 shows that the raw scores associated with passing on the February 2004 form were 54 for ELA and 43 for mathematics.[8]

Table 7.1: Passing Scores on the February 2004 CAHSEE

English-Language Arts			Mathematics	
Raw Score	**Theta**		**Raw Score**	**Theta**
53	0.4724		42	-0.1249
54	0.5290	⇐ **350** ⇒	43	-0.0685
55	0.5863		44	-0.0118

The new reporting scales for the ELA and mathematics assessments were established. The reporting scale was obtained by establishing linear scaling parameters to transform the theta values corresponding to each raw score to the reporting scale. The transformation constants associated with these scoring tables are as follows:

ELA: Scale Score = Theta * 33.7230 + 332.1605

Mathematics: Scale Score = Theta * 32.2900 + 352.2119

The resulting scale has several notable characteristics:

1. The Passing raw scores (54 and 43, respectively) are set to a scale score of 350.
2. The Proficient cut score to be used for ESEA accountability purposes has been set at 380 for both ELA and mathematics. On the theta scale, the Proficient cut-scores were 1.4152 and 0.8762, respectively. These translate to scale scores of 379.88 for ELA and 380.50 for mathematics on the new scale. To simplify the

[8] Since raw scores are integers and the theta metric is continuous, the cut-score theta levels will usually fall between the thetas at adjacent raw scores. In this situation, the raw scores corresponding to the theta level closest to the values of 0.5356 for ELA and -0.0701 for mathematics were chosen as the raw-score cut points.

communication of these important ESEA cut-scores to a wide audience, 380 is used for both ELA and mathematics.

3. The Advanced cut scores to be used for ESEA accountability purposes were thetas of 2.1056 for ELA and 2.1456 for mathematics, which translate to 403 and 422, respectively.
4. The minimum and maximum scale scores were set at 275 and 450 for both ELA and mathematics.

Since the February 2004 administration, the raw-score to scale-score conversions have been maintained through the use of IRT-based equating procedures. The following rule is followed in determining the raw cut score a student must achieve to be classified as Passing, Proficient, or Advanced. For each administration, scale scores and raw scores that correspond to these cuts are identified. First, the scale score that is equal to or greater than the cut point (e.g., 350) is located. If 350 is not found in the conversion, the next higher score (e.g., 351) is used as the cutoff point. The corresponding raw score associated with that rounded scale score is defined to be the raw score equivalent of that cut.

Results

Table 7.2 is a summary of the raw score cuts and overall passing rates based on equating samples for each administration during the 2011–12 school year. For each administration, the passing rate for each proficiency level is the percentage of examinees earning a score at or above the raw score cut for that level. For example, for the July 2011 administration, 27 percent of examinees who took the ELA examination received a raw score of 54 or higher and thereby passed the ELA. Of those who passed, 5 percent of examinees received a score of 68 or higher and passed the Proficient cut point, while 2 percent scored 76 or above and passed the Advanced cut point.

Among examinees who took the mathematics exam, 27 percent passed by attaining a raw score of 43 or higher. Of those who passed, 3 percent were at or above the Proficient cut point with a raw score of 59 or higher, while 1 percent were at or above the Advanced cut point with a raw score of 73 or higher.

Table 7.2: Summary of Cut Scores and Passing Rates Based on Equating Samples

Administration	Levels	ELA		Mathematics	
		Raw Score Cut	Percent Passing[1]	Raw Score Cut	Percent Passing[1]
July 2011	Pass	54	27	43	27
	Proficient And Above	68	5	59	3
	Advanced And Above	76	2	73	1
October 2011	Pass	55	36	42	38
	Proficient And Above	69	9	58	9
	Advanced And Above	77	4	72	2
November 2011	Pass	56	39	43	42
	Proficient And Above	69	11	59	9
	Advanced And Above	76	5	73	2
December 2011	Pass	55	29	41	35
	Proficient And Above	68	6	57	4
	Advanced And Above	76	2	72	1
February 2012	Pass	54	68	43	69
	Proficient And Above	68	43	59	43
	Advanced And Above	75	25	73	15
March 2012	Pass	56	78	43	79
	Proficient And Above	69	50	58	54
	Advanced And Above	76	28	72	20
May 2012	Pass	57	29	43	34
	Proficient And Above	70	12	59	11
	Advanced And Above	78	5	73	3

[1] Percentage of examinees at or above the cut score.

Chapter 8: Scoring and Reporting

ETS conforms to high standards of quality and fairness (ETS, 2002) when scoring tests and reporting scores. Such standards dictate that ETS provide accurate and understandable assessment results to the intended recipients. It is also the mission of ETS to provide appropriate guidelines for score interpretation and to provide cautions about the limitations in the meaning and use of the test scores. Finally, attempts are made to ensure sufficient data are collected for the major subgroups of students. These data are necessary for conducting analyses that ensure equitable results for various groups of test takers.

Procedures for Maintaining and Retrieving Individual Scores

Items for the CAHSEE mathematics and ELA examinations, except for the writing prompts in ELA, are MC. Students are presented with an item and asked to select the correct option from four possible choices. Students mark their choices on an answer document. All MC items are machine scored. Responses to the writing task are scored by trained readers.

In order to score and report the CAHSEE results, ETS follows an established set of written procedures. These specifications are presented in the next sections.

Scoring and Reporting Specifications

ETS develops standardized scoring procedures and specifications so that test materials are processed and scored accurately. These documents include the following:

- General Reporting Specifications—Provides the calculation rules for the information presented on CAHSEE summary reports and defines the appropriate codes to use when a student does not take or complete a test or when a score will not be reported.

- Score Key and Score Conversion—Defines file formats and information that is provided for scoring; defines the process of converting raw scores to scale scores.

- Form Planner Specifications—Describes the contents of files that contain keys required for scoring. Specifically, the form planner specification document defines each column in the form planner, the values contained in the columns, and the naming conventions for the form planners. The form planners contain the information about an assembled test form, including the test name and administration month and year. The form planners also contain information about each item, including item identification, sequence number, item type (i.e., operational, linking, field test), scoring key, strand-score identification, standards, classical statistics (i.e., p-value, biserial and point biserial correlations, percentage of students choosing each option), and IRT statistics (i.e., b-value, IRT fit rating).

- Strand Names and Item Numbers—Identifies the reporting strands for each test and the number of items in each strand.

- Matching Criteria for MC and Writing Answer Documents—Describes the method used to match students' writing and MC results.

The scoring specifications are reviewed and revised by the CDE and ETS. After a version that both parties agree to is finalized, the CDE issues a formal approval of the scoring and reporting specifications.

Scanning and Scoring

Answer documents are scanned by Pearson and scored by ETS in accordance with the scoring specifications that have been approved by the CDE. Each school district must return scorable materials within five working days after the administration date for each test and nonscorable materials within ten working days.

Pearson scans the mathematics and ELA answer document sheets and transmits electronic files of all information captured to ETS for scoring and reporting. This includes demographic information and MC item data. Pearson also sends the data file images to ETS for use in the OSN system. These files consist of the images of the students' constructed responses as well as the unique PAS identification numbers. The identification numbers allow for the absolute matching of CR scores to other student data including demographic information and responses from MC items.

Scoring Multiple-Choice Items

ETS maintains all scoring keys with its SKM system. Prior to scoring, information on all test items and the test keys are loaded into the SKM system from test form planners created from the item bank. The keys are "locked" to assure that they cannot be used in scoring until the appropriate quality control checks have been completed. When the final quality control check is completed and no errors have been identified, the test keys are "unlocked" for use in the scoring process.

Scoring Writing Tasks

All student responses to the ELA writing tasks are scored in the OSN system, a distributed, Web-based scoring system that enables a large number of raters to view and score assigned responses from remote locations. All identifying information from the responses sent to raters is removed so that neither the identity of the student nor the student's school is revealed to the rater; the rater sees only the student response.

Types of Scores and Strand Scores

Raw Score

For mathematics, the total raw score equals the sum of the examinee's scores on the MC test items. In ELA, the total raw score equals the weighted sum of the examinee's scores on both the MC items and the writing task. The weighting scheme for raw scores is described in detail in Chapter 2.

Strand Score

The items on the mathematics and ELA examinations are aggregated into groups, referred to as Reporting Strands. A strand score is obtained by summing an examinee's scores on the items in each reporting strand. A description of the CAHSEE reporting strands is provided in Table 2.2 of Chapter 2.

Scale Score

Raw scores on each CAHSEE examination are transformed to three-digit scale scores using the equating process described in Chapter 6. Scale scores range from 275 to 450 for both mathematics and ELA. The scale scores of students who have been tested in different administrations in a given content area can be compared. However, the raw scores cannot be meaningfully compared, because these scores are affected by the difficulty of the test taken as well as the ability of the student.

Passing Scores

A passing score is 350 for both ELA and mathematics. A student will pass either of the tests if the total score is 350 or higher.

As part of the reporting requirements for the ESEA, cut scores defining Proficient and Advanced performance on the CAHSEE were set for both ELA and mathematics. The ESEA Proficient cut score is 380 for both ELA and mathematics. The Advanced cut score is 403 for ELA and 422 for mathematics. These values are used to classify tenth-grade students taking the CAHSEE into the Proficient and Above category as part of California's assessment of AYP.

Score for Writing Applications

The score for Writing Applications shows the total number of points that an examinee received on the essay. Each essay is rated by two readers on a 1–4 scale. Students can also receive a "non-scorable" (NS) score of 0 if they do not write enough to score, write off the topic, write illegibly, or write in a language other than English. The scores from both readers must be the same or within one score point of each other. If the difference between the two scores is more than one score point,[9] the score for that essay is considered discrepant, and a scoring leader provides a third score, which

[9] If an essay received an NS from a reader and a score of 1, 2, 3, or 4 from another reader, scores are considered discrepant and the essay will receive a score from the third reader.

becomes the score of record. Otherwise, the writing score is obtained by averaging two scores.

Score Verification Procedures

ETS takes various measures to ascertain that the scoring keys are applied to the student responses as expected, and the student scores are computed accurately.

Scoring Key Verification Process

Scoring keys, provided in the form planners, are produced and verified thoroughly by performing various quality control checks before and after they are loaded into the SKM system. The form planners contain the information about an assembled test form including scoring keys, test name, administration year, strand-score identification, and the standards and statistics associated with each item. The various checks that are performed before the keys are finalized are listed below:

1. The form planners are checked for accuracy against the Form Planner Specifications document and the Score Key and Score Conversion document before the keys are loaded into the SKM system.

2. The printed lists of the scoring keys are checked again once the keys have been loaded into the SKM system.

3. The sequence of linking items[10] in the form planners is matched with their sequence in the actual test booklets.

4. The entire scoring system is tested using early returned answer documents.

5. Throughout the answer document scanning/scoring process, a number of records are randomly pulled from each scanned batch and are hand-scored by the resolutions team using a template provided by SKM staff. This QC step verifies that the scanned file matches the hard copy document and that electronic-scoring results are consistent with the hand-scoring results.

6. Classical item analyses are run on an early sample of data to provide an additional check of the keys. Although rare, if an item is found to be problematic (e.g., very difficult, low correlation with criterion), a follow-up process is carried out to determine whether it should be excluded from further analyses.

[10] Linking items are used to link the scores on the current administration's test forms to scores obtained on the base forms to adjust for the difficulty level of the forms across administrations. This is accomplished during the equating process, as discussed in Chapter 2.

Monitoring and Quality Control of Writing Scoring

Students' responses to the ELA writing task are read by two readers, and their writing scores are based on the average of the two ratings. The next sections provide details of the process employed by ETS to score the writing tasks.

The Online Scoring Network™

All student responses to the ELA writing tasks are scored in the OSN system, a distributed, Web-based scoring system that enables a large number of scorers to view and score assigned responses from remote locations. The distributed OSN is as reliable as a center-based solution and is more flexible:

- All scorers are trained with a consistent set of materials and must pass a certification test before they are admitted to the CAHSEE reader pool.
- Scorers must successfully score a calibration set of papers before each scoring session. Failure to do so locks them out of scoring for that session.
- Trained scoring leaders remotely monitor the scoring progress of each scorer using virtual monitoring tools and real-time score-performance data.
- The system is password protected and scorers can access the system only during their scheduled reading blocks.
- The OSN improves the efficiency of the scoring process. Specifically, essays are sorted and distributed by topic to promote efficiency. In addition, each response is systematically routed for first and second reads and discrepancy resolutions, thereby assuring that it is read by different scorers.
- The OSN allows authorized personnel to actively monitor the scoring process dynamically.

Training Scoring Leaders and Readers

Individuals who are selected to serve as scoring leaders or readers must be college graduates who possess at least a bachelor's degree. Each prospective scorer is required to participate in a systematic and multi-tiered virtual training program that assures the application of uniform scoring standards. Scoring leaders are experienced scorers who have had additional training. During operational scoring, scoring leaders monitor and assist readers throughout each scoring session over the Internet.

Scoring Leader Training

Scoring leaders are trained using the Live Meeting approach, which is a dynamic and collaborative Web-based training model that is effective, efficient, and secure. ETS trains scoring leader personnel in advance of each administration.

During the training session, scoring leaders review a comprehensive set of CAHSEE training materials assembled by the CAHSEE Scoring Manager. The Scoring Manager presents scoring guides, anchor papers representative of each score point, and decision set papers that cover ambiguities and problematic approaches to each CR topic. Throughout the training session, scoring leaders can engage the Scoring Manager by way of a conference call or live meeting.

Reader Training

Once recruited, CAHSEE readers are trained by using extensive materials provided over the CAHSEE tutorial Web site. Pre-operational training focuses on general instruction on how to apply the program's scoring guide criteria and on practice in certification tests. Readers log on to a tutorial site for Web-based training, where they are provided information about the writing task, scoring rubrics, scoring rationales, anchor papers, and the standards to be maintained. Anchor papers are selected to demonstrate clear examples of each score point.

At the completion of the training process, readers take a certification test that consists of a set of pre-scored responses. If readers achieve the required exact agreement rate, they are admitted to the reader pool. Readers are permitted to score only those topics on which they have been trained and certified.

System Training

Scorers have different levels of experience working on computers. OSN has established simple step-by-step procedures that guide readers through every aspect of the scoring process. In addition, an ETS Web site currently includes one section dedicated to OSN system functionality and provides detailed guides for readers.

All CAHSEE readers log on, certify, and perform a system test well in advance of operational scoring. This assures that the equipment is functioning as it should, and that the readers are familiar with the OSN system and the scoring process well before they view live items.

Scorer Qualification

ETS maintains a pool of more than 8,000 readers, more than 1,000 of whom are currently trained and experienced in the scoring of CAHSEE writing tasks. Most are educators who hold full-time teaching positions. Preference is given to those with a bachelor's degree in English or a related field. The recruiting and training plan includes a commitment to the CDE to maintain a 20 percent participation rate by California English teachers in the CAHSEE reader pool.

Scorers continue to qualify throughout the operational scoring process. The structure of a typical CAHSEE scoring session, whether a half day or full day, is standardized. Prior to each scoring session, certified readers are required to demonstrate ongoing proficiency by scoring a set of calibration papers, which consists of pre-scored responses to one topic arranged in an electronic folder. The calibration results are automatically calculated. Failure to satisfy this requirement prohibits a reader from scoring that day. Readers must calibrate for each day of operational scoring. During

the scoring session, scoring leaders will monitor scorer performance dynamically and target sub-standard scorers for additional training and calibration.

Accuracy Monitoring

The monitoring functions of the OSN provide a useful method for overseeing the accuracy of scoring and the performance of individual topics. The OSN produces a variety of reports with extensive data on both readers and topics, as well as an overview of the progress and accuracy of the overall scoring process. Most reader performance data are available immediately. A content specialist or a scoring leader is able to view statistical tabulations of reader performance within any given time period. Scoring leaders even have the capability of monitoring readers while they are actively scoring a group of essays.

The OSN can produce reports to show the degree to which readers are consistent in scores that they assign. The consistency is measured in terms of the percentage of instances in which the first and second readers' scores are exact, adjacent, and discrepant; this is a commonly used measure of inter-rater reliability. In addition, the overall mean and the percentage of scores awarded at each score point reveal whether the reader is fulfilling the performance standard of using the full range, or whether the reader is scoring too low, too high, or too exclusively in the middle. If a reader's rate of agreement begins to decline, the reader is retrained by a scoring leader and closely monitored thereafter. If the reader's performance does not improve, the reader is released.

In addition to a statistical depiction of reader performance, the OSN monitoring function also provides a statistical portrait of topic performance. Test development staff is able to see over time whether a given topic is performing well by considering:

- The number/percentage of exact, adjacent, and discrepant scores awarded for papers on that topic
- The average rate at which papers for this topic are read
- The mean score overall
- The percentage of scores awarded at each point
- The number and percentage of scoring iterations (indicating how many third and fourth readings were required)

Data are also available on a particular type of paper (monitor, production, etc.) within a specified time period. As is the case with reader performance, the goal is to assure a higher rate of exact scores than adjacent scores, a low percentage of discrepancies, and a distribution of scores over the whole range. An additional aim is to have comparable mean scores across topics.

Quality Control in Raw-to-Scale Score Conversions

ETS psychometricians employ special procedures that adjust for item difficulty differences across test forms. As a result of this process, scoring tables are produced. These tables map the current administration's raw score to an appropriate scale score.

The Information Technology (IT) Division utilizes these tables to generate scale scores for each student.

After score conversion is completed, both the IT Division and the Statistical Analysis Division independently generate a raw-score-to-scale-score mapping report and verify the accuracy of the score conversion against the original scoring tables.

Overview of Score Aggregation Procedures

In order to provide meaningful results to the stakeholders, the CAHSEE scores for a given content area are aggregated at the school, independently testing charter school, district, and state levels. The aggregated scores are generated for individual and group scores. The following section presents the types of aggregation performed on CAHSEE scores.

Individual Scores

The tables referenced in this section provide state-level summary statistics describing student performance in each CAHSEE administration.

Score Distributions and Summary Statistics

Summary statistics for individual scores are presented in Tables E.3.1 to E.4.3, in the Executive Summary chapter. Included in the tables are the number of items in each test, the number of students taking each test, and the means and standard deviations of student scores expressed in terms of both raw scores and scale scores. The percentage of students passing each CAHSEE content area is presented in Table E.2.

Frequency distributions of scale scores for ELA and mathematics are presented in Tables 1 and 2 in Appendices 8.A to 8.G. The results are reported in terms of score intervals. The passing line indicates the Pass/Not Pass cuts. Similar distributions of scale scores for ELA and mathematics, with the ESEA cuts indicated, are presented in Tables 3 and 4 in Appendices 8.A to 8.G. The first line indicates the Advanced level cut, while the second line indicates the Proficient level cut.

The numbers in the summary tables may not match exactly the results reported on the CDE's Web site, as there may be small differences in the samples used to compute the statistics. The statistics in these tables may differ slightly from the statewide statistics reported on the CDE Web site, because school districts may conduct data correction after the data file is generated for the analyses in this chapter.

Group Scores

Statistics summarizing student performance by content area and test administration for selected groups of students are provided in Tables 5 (ELA) and 6 (mathematics) in Appendices 8.A to 8.G. In the tables, students are grouped by demographic characteristics, including grade, gender, ethnicity, English proficiency, need for special education services, and economic status.[11] The tables show the numbers of valid cases in each group as well as scale score means and standard deviations for

[11] Students' economic status was determined by considering the education level of their parents and whether or not they are eligible to participate in the NSLP.

each demographic group. Table 2.3 in Chapter 2 defines the demographic groups included in the tables.

Similar summary statistics for demographic groups are presented in Tables 7 and 8 in Appendices 8.A to 8.G. These summary statistics display the percentages of examinees classified as Below Proficient, Proficient, and Advanced according to the ESEA classifications. Selected percentiles, scale score means, and standard deviations for the subgroups are presented for all students in Tables 9 and 10 for ELA and mathematics, respectively.

Table 8.1 provides a listing of the frequency distribution and demographic summary tables found in the Appendices. To simplify the presentation of these data, all tables for this section are located from Appendix 8.A to Appendix 8.G.

Table 8.1: Listing of Frequency Distribution and Demographic Summary Tables

Table[1]	Content	Label
8.x.1	Frequency Distributions Highlighted at Pass line – ELA	Frequency Distributions – ELA
8.x.2	Frequency Distributions Highlighted at Pass line – Mathematics	Frequency Distributions – Mathematics
8.x.3	Frequency Distributions Highlighted at ESEA cuts – ELA	Frequency Distributions – ELA
8.x.4	Frequency Distributions Highlighted at ESEA cuts – Mathematics	Frequency Distributions – Mathematics
8.x.5	Scale Score Summary Statistics and Passing Rates for All Examinees - ELA	Demographic Summary for All Examinees – ELA
8.x.6	Scale Score Summary Statistics and Passing Rates for All Examinees – Mathematics	Demographic Summary for All Examinees – Mathematics
8.x.7	ESEA Summary for all Examinees – ELA	ESEA Demographic Summary for All Examinees – ELA
8.x.8	ESEA Summary for all Examinees – Mathematics	ESEA Demographic Summary for All Examinees – Mathematics
8.x.9	Scale Score Percentiles for All Examinees – ELA	Examinee Demographics Showing Mean Scale Score at Each Percentile – ELA
8.x.10	Scale Score Percentiles for All Examinees – Mathematics	Examinee Demographics Showing Mean Scale Score at Each Percentile – Mathematics

[1] x = Administration, where tables A = July, B = October, C = November, D = December, E = February, F = March, G = May.

Reports to be Produced and Scores for Each Report

The scores of the CAHSEE ELA and mathematics tests provide results or score summaries that are reported for different purposes. The three major purposes include:

1. Communicating with parents and guardians about students' achievements and whether the students fulfill one or both part(s) of the CAHSEE graduation requirement.

2. Informing decisions needed to provide additional assistance for students who did not pass one or both part(s) of the CAHSEE.

3. Providing data for state and federal accountability programs for schools.

A detailed description of the uses and applications of the CAHSEE score reports is presented in the next section.

Types of Score Reports

There are three categories of CAHSEE score reports. These categories and the specific reports in each category are given in the table below.

Table 8.2: Types of CAHSEE Reports

1. Individual Reports	▪ CAHSEE Student and Parent Report
2. Aggregate Reports	▪ CAHSEE Subgroup Summary (including the Ethnicity for Economic Status) ◦ School-Level Reports ◦ District-Level Reports ◦ State-Level Reports
3. Quarterly Reports	▪ DataQuest Reporting ▪ Detail File for Academic Performance Index (API) and Adequate Yearly Progress (AYP)

These reports are sent to the independently testing charter schools, county offices of education, or school districts, who in turn forward them to the appropriate schools. In the case of the CAHSEE Student and Parent Report, the reports are mailed to the child's parents or guardians and a copy is retained in the student's file. Internet reports are described on the CDE Web site and are accessible to the public online at http://cahsee.cde.ca.gov/.

Score Report Contents

The CAHSEE Student and Parent Report provides the student's scale score and the score in relation to the passing score for each test taken by the student. Scale scores are reported on a scale ranging from 275 to 450.

The score report also provides strand scores that indicate how the student performed on each of the content strands assessed. The number of questions in the strand and the number answered correctly are provided on the score report. ELA strand score results are grouped under the broader categories of Reading and Writing. Reports for SWDs and EL who use modifications include a notation that indicates that the student was tested with modifications. Modifications change what is being tested and therefore, change the meaning of these scores. If students use modifications, their scores are counted differently from non-modified test scores on summary reports. Students who use a modification and earn the equivalent of a passing score on one or both parts of the CAHSEE do not pass but may be eligible for a local waiver of the CAHSEE requirement. Tests for students who use accommodations that are specified in their IEPs or Section 504 plans are reported in the same way as non-accommodated tests.

Aggregate reports comprise a series of student, demographic, and geographic summaries that inform LEAs and state officials and their constituencies of student performance across defined subgroups. Aggregated reports following each administration are prepared in PDF format and are printed, packed, and shipped by

expedited delivery to LEAs. Reports are also posted in each LEA's secure folder on the CAHSEE Web site. The aggregated results following each administration are not publicly distributed. At the CDE's direction, these reports include complete performance distributions for each of the designated subgroups, regardless of the size of the subgroup. The aggregate results available to the public via DataQuest do not include distributions for any subgroup consisting of 10 or fewer students.

ETS also provides the CDE with data files that meet the quarterly CAHSEE results reporting requirements as well as the state (API) and federal (AYP) accountability reporting timeline requirements. CD-ROMs or DVDs containing these data files with encryption software are delivered to the CDE so that the CDE can generate reports as well as conduct additional analyses to inform decision makers about student achievement and programs delivered to students.

Score Report Applications

The results for the CAHSEE are used primarily to identify students who are not developing high school graduate-level competencies in reading, writing, and mathematics that are essential after high school. Beginning in the 2005–06 school year, with the exception of eligible SWDs, no student received a public high school diploma without passing the CAHSEE and meeting all other state and district requirements for graduation.

Counties, school districts, and schools are encouraged to use the summary results and other standards-based evidence of student achievement to develop and implement an ongoing process for refining classroom instruction and school programs. The goal is to work with school staff to identify patterns of student performance and identify program areas needing improvement.

In addition, the state and federal governments use the CAHSEE results for grade ten as a measure of school and school district accountability. The state accountability program is the Public Schools Accountability Act; the federal accountability program is the ESEA. The use of CAHSEE results for these accountability programs is intended to be independent of how the CAHSEE is used for individual student accountability.

Criteria for Interpreting Test Scores

A school district may use CAHSEE results to help make decisions about student graduation. However, it is important to remember that a single test can provide only limited information. Other relevant graduation requirements should be considered as well. It is also important to note that a student's score in a content area contains measurement error and could vary if the student were retested.

Criteria for Interpreting Score Reports

The information presented on various reports must be interpreted with caution when making performance comparisons. When comparing scale score and performance level results for the CAHSEE, the user is limited to comparisons within the same content area. This is because the underlying scales are different for each content area. Comparing scores obtained in different content areas should be avoided because the

results are not on the same scale. Comparisons between raw scores and cluster scores should be limited to comparisons within not only content area but also test administration. The user may compare scores for the same content area within a school, between schools, or between a school and its district, its county, or the state. The user can also make comparisons within the same content area across administrations.

Appendix 8.A: Frequency Distributions and Demographic Summaries—July 2011

Table 8.A.1: Frequency Distributions, ELA—July 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	7	0	7	99
440-449	8	0	15	99
430-439	13	0	28	99
420-429	26	0	54	99
410-419	39	1	93	99
400-409	59	1	152	98
390-399	63	1	215	97
380-389	142	2	357	95
370-379	211	3	568	92
360-369	495	7	1,063	85
350-359[1]	881	12	1,944	73
340-349	1,277	18	3,221	55
330-339	1,334	19	4,555	36
320-329	988	14	5,543	23
310-319	744	10	6,287	12
300-309	449	6	6,736	6
290-299	255	4	6,991	2
280-289	109	2	7,100	1
270-279	66	1	7,166	0

[1]Passing Score = 350

Table 8.A.2: Frequency Distributions, Mathematics—July 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	13	0	13	99
440-449	9	0	22	99
430-439	15	0	37	99
420-429	17	0	54	99
410-419	11	0	65	99
400-409	46	1	111	98
390-399	61	1	172	98
380-389	65	1	237	97
370-379	171	2	408	94
360-369	393	6	801	88
350-359[1]	1,066	15	1,867	73
340-349	1,375	20	3,242	53
330-339	1,625	24	4,867	30
320-329	983	14	5,850	15
310-319	667	10	6,517	6
300-309	275	4	6,792	2
290-299	82	1	6,874	1
280-289	13	0	6,887	0
270-279	23	0	6,910	0

[1]Passing Score = 350

Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	7	0	7	99
440-449	8	0	15	99
430-439	13	0	28	99
420-429	26	0	54	99
410-419	39	1	93	99
403-409[1]	41	1	134	98
390-402	81	1	215	97
380-389[2]	142	2	357	95
370-379	211	3	568	92
360-369	495	7	1,063	85
350-359	881	12	1,944	73
340-349	1,277	18	3,221	55
330-339	1,334	19	4,555	36
320-329	988	14	5,543	23
310-319	744	10	6,287	12
300-309	449	6	6,736	6
290-299	255	4	6,991	2
280-289	109	2	7,100	1
270-279	66	1	7,166	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	13	0	13	99
440-449	9	0	22	99
430-439	15	0	37	99
422-429[1]	10	0	47	99
410-421	18	0	65	99
400-409	46	1	111	98
390-399	61	1	172	98
380-389[2]	65	1	237	97
370-379	171	2	408	94
360-369	393	6	801	88
350-359	1,066	15	1,867	73
340-349	1,375	20	3,242	53
330-339	1,625	24	4,867	30
320-329	983	14	5,850	15
310-319	667	10	6,517	6
300-309	275	4	6,792	2
290-299	82	1	6,874	1
280-289	13	0	6,887	0
270-279	23	0	6,910	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.A.5: Demographic Summary for All Examinees ELA—July 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	7,166	1,944	27	5,222	73	337	52	57	58	44	47	2.1
Grade[3]												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-	-
Twelfth	6,195	1,580	26	4,615	74	335	51	56	57	43	46	2.0
Adult Education	971	364	37	607	63	347	57	63	63	52	52	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	3,790	1,042	27	2,748	73	336	53	56	57	43	47	2.0
Female	3,369	900	27	2,469	73	338	51	58	59	46	47	2.1
Unknown	7	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	21	6	29	15	71	346	61	60	65	46	47	2.3
Asian	708	174	25	534	75	334	44	56	56	46	50	1.9
Pacific Islander	57	17	30	40	70	335	52	54	55	41	48	2.1
Filipino	126	29	23	97	77	337	50	57	58	46	45	2.1
Hispanic or Latino	4,859	1,195	25	3,664	75	335	52	56	57	43	46	2.0
African American	687	226	33	461	67	340	57	59	60	47	48	2.1
White (not of Hispanic origin)	536	235	44	301	56	349	60	61	63	51	55	2.2
Two or More Races	172	62	36	110	64	342	55	59	60	47	50	2.1
Language Fluency												
English-Only Students	2,087	737	35	1,350	65	342	58	59	60	46	49	2.1
Initially Fluent English Proficient (IFEP)	211	89	42	122	58	344	59	61	62	47	53	2.1
Reclassified Fluent English Proficient (RFEP)	234	105	45	129	55	346	57	62	63	50	53	2.2
English-Learner Students	3,770	680	18	3,090	82	331	47	53	55	41	44	2.0
Unknown	864	333	39	531	61	346	57	63	62	52	51	2.1
Economically Disadvantaged												
No	1,017	334	33	683	67	340	54	58	59	45	50	2.1
Yes	4,780	1,127	24	3,653	76	334	50	55	57	42	46	2.0
Unknown	1,369	483	35	886	65	343	57	60	61	50	50	2.1
Special Education Program Participation												
Students Receiving Services	847	112	13	735	87	323	46	47	50	36	40	1.8
Students Not Receiving Services	6,319	1,832	29	4,487	71	339	53	58	59	46	48	2.1

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

[3] As of 2007, only Grade twelve and Adult Education students may take the July administration of the CAHSEE.

Table 8.A.6: Demographic Summary for All Examinees, Mathematics—July 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	6,910	1,867	27	5,043	73	339	52	51	48	43	37
Grade[3]											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-
Twelfth	5,906	1,505	25	4,401	75	338	52	50	47	42	36
Adult Education	1,004	362	36	642	64	346	58	54	53	49	39
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	3,109	830	27	2,279	73	338	51	51	46	42	35
Female	3,789	1,034	27	2,755	73	340	54	50	49	43	38
Unknown	12	3	25	9	75	346	58	53	57	44	36
Race/Ethnicity											
American Indian or Alaska Native	26	8	31	18	69	341	54	52	49	42	37
Asian	222	96	43	126	57	354	55	59	56	50	49
Pacific Islander	61	12	20	49	80	335	46	50	44	39	37
Filipino	100	31	31	69	69	341	53	53	49	42	41
Hispanic or Latino	4,722	1,178	25	3,544	75	338	52	50	47	42	36
African American	1,034	266	26	768	74	338	52	50	47	41	35
White (not of Hispanic origin)	565	229	41	336	59	348	60	56	53	48	38
Two or More Races	180	47	26	133	74	339	53	50	47	44	37
Language Fluency											
English-Only Students	2,817	832	30	1,985	70	340	54	52	48	43	36
Initially Fluent English Proficient (IFEP)	312	110	35	202	65	343	56	53	52	44	37
Reclassified Fluent English Proficient (RFEP)	495	158	32	337	68	343	57	52	51	45	38
English-Learner Students	2,364	443	19	1,921	81	335	47	48	44	40	36
Unknown	922	324	35	598	65	346	57	54	52	48	38
Economically Disadvantaged											
No	1,172	352	30	820	70	341	54	52	49	43	37
Yes	4,262	1,039	24	3,223	76	338	51	50	46	42	36
Unknown	1,476	476	32	1,000	68	343	55	53	51	46	37
Special Education Program Participation											
Students Receiving Services	753	100	13	653	87	328	42	45	39	34	33
Students Not Receiving Services	6,157	1,767	29	4,390	71	341	54	51	49	44	37

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1.

[3] As of 2007, only Grade twelve and Adult Education students may take the July administration of the CAHSEE.

Table 8.A.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA— July 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		7,166	6,809	95	223	3	134	2	357	5
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	6,195	5,956	96	160	3	79	1	239	4
	Adult Education	971	853	88	63	6	55	6	118	12
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	3,790	3,595	95	131	3	64	2	195	5
	Female	3,369	3,208	95	91	3	70	2	161	5
	Unknown	7	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	21	17	81	1	5	3	14	4	19
	Asian	708	676	95	20	3	12	2	32	5
	Pacific Islander	57	56	98	1	2	0	0	1	2
	Filipino	126	119	94	4	3	3	2	7	6
	Hispanic or Latino	4,859	4,700	97	110	2	49	1	159	3
	African American	687	647	94	35	5	5	1	40	6
	White (not of Hispanic origin)	536	442	82	43	8	51	10	94	18
	Two or More Races	172	152	88	9	5	11	6	20	12
Language Fluency	English-Only Students	2,087	1,900	91	112	5	75	4	187	9
	Initially Fluent English Proficient (IFEP)	211	195	92	10	5	6	3	16	8
	Reclassified Fluent English Proficient (RFEP)	234	221	94	11	5	2	1	13	6
	English-Learner Students	3,770	3,732	99	31	1	7	0	38	1
	Unknown	864	761	88	59	7	44	5	103	12
Economically Disadvantaged	No	1,017	942	93	42	4	33	3	75	7
	Yes	4,780	4,636	97	101	2	43	1	144	3
	Unknown	1,369	1,231	90	80	6	58	4	138	10
Special Education Program Participation	Receiving Services	847	834	98	10	1	3	0	13	2
	Not Receiving Services	6,319	5,975	95	213	3	131	2	344	5

[1] Results for groups with fewer than 11 students are not reported.

[2] As of 2007, only Grade twelve and Adult Education students may take the July administration of the CAHSEE.

Table 8.A.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—July 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		6,910	6,673	97	190	3	47	1	237	3
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	5,906	5,744	97	133	2	29	0	162	3
	Adult Education	1,004	929	93	57	6	18	2	75	7
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	3,109	2,978	96	103	3	28	1	131	4
	Female	3,789	3,684	97	86	2	19	1	105	3
	Unknown	12	11	92	1	8	0	0	1	8
Race/Ethnicity	American Indian or Alaskan Native	26	24	92	1	4	1	4	2	8
	Asian	222	179	81	23	10	20	9	43	19
	Pacific Islander	61	61	100	0	0	0	0	0	0
	Filipino	100	98	98	1	1	1	1	2	2
	Hispanic or Latino	4,722	4,614	98	97	2	11	0	108	2
	African American	1,034	1,015	98	16	2	3	0	19	2
	White (not of Hispanic origin)	565	511	90	43	8	11	2	54	10
	Two or More Races	180	171	95	9	5	0	0	9	5
Language Fluency	English-Only Students	2,817	2,713	96	87	3	17	1	104	4
	Initially Fluent English Proficient (IFEP)	312	300	96	10	3	2	1	12	4
	Reclassified Fluent English Proficient (RFEP)	495	484	98	10	2	1	0	11	2
	English-Learner Students	2,364	2,326	98	30	1	8	0	38	2
	Unknown	922	850	92	53	6	19	2	72	8
Economically Disadvantaged	No	1,172	1,126	96	34	3	12	1	46	4
	Yes	4,262	4,165	98	84	2	13	0	97	2
	Unknown	1,476	1,382	94	72	5	22	1	94	6
Special Education Program Participation	Receiving Services	753	739	98	13	2	1	0	14	2
	Not Receiving Services	6,157	5,934	96	177	3	46	1	223	4

[1] Results for groups with fewer than 11 students are not reported.

[2] As of 2007, only Grade twelve and Adult Education students may take the July administration of the CAHSEE.

Table 8.A.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—July 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	280	297	322	336	350	379	415	337	25	7,166
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	278	297	320	336	350	374	407	335	24	6,195
Adult Education	293	308	328	342	360	407	438	347	29	971
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	276	295	318	336	350	381	411	336	26	3,790
Female	289	302	324	338	350	379	419	338	24	3,369
Unknown	-	-	-	-	-	-	-	-	-	7
Race/Ethnicity										
American Indian or Alaska Native	285	295	320	340	352	407	423	346	35	21
Asian	276	293	318	336	348	376	407	334	26	708
Pacific Islander	293	299	322	336	350	372	381	335	20	57
Filipino	287	301	324	335	348	381	411	337	25	126
Hispanic or Latino	283	299	320	336	348	374	404	335	23	4,859
African American	280	299	324	340	356	381	401	340	25	687
White (not of Hispanic origin)	287	299	326	344	365	419	445	349	35	536
Two or More Races	275	291	320	342	358	407	428	342	33	172
Language Fluency										
English-Only Students	278	297	324	340	356	392	423	342	29	2,087
Initially Fluent English Proficient (IFEP)	280	306	328	344	358	384	419	344	26	211
Reclassified Fluent English Proficient (RFEP)	287	306	332	346	360	381	401	346	22	234
English-Learner Students	280	297	318	332	344	363	381	331	20	3,770
Unknown	285	304	328	342	360	404	438	346	30	864
Economically Disadvantaged										
No	278	297	324	338	354	389	423	340	27	1,017
Yes	283	297	320	334	348	372	401	334	23	4,780
Unknown	276	301	326	340	358	398	433	343	29	1,369
Special Education Program Participation										
Students Receiving Services	275	285	306	322	338	363	392	323	24	847
Students Not Receiving Services	285	301	324	338	352	381	419	339	25	6,319

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation.

[3] Results for groups with fewer than 11 students are not reported.

Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	296	308	327	339	350	372	408	339	21	6,910
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	296	308	325	337	350	368	404	338	20	5,906
Adult Education	304	317	332	343	353	392	437	346	23	1,004
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	291	304	323	337	350	376	415	338	23	3,109
Female	299	313	328	339	350	370	404	340	19	3,789
Unknown	323	323	330	342	356	401	401	346	22	12
Race/Ethnicity										
American Indian or Alaska Native	277	304	328	338	350	395	430	341	29	26
Asian	296	308	328	344	364	437	450	354	37	222
Pacific Islander	296	308	323	336	344	368	378	335	18	61
Filipino	303	315	328	342	352	366	420	341	19	100
Hispanic or Latino	296	308	327	337	348	368	398	338	19	4,722
African American	294	308	327	339	350	366	385	338	19	1,034
White (not of Hispanic origin)	296	313	332	344	357	404	437	348	26	565
Two or More Races	304	311	326	337	351	379	398	339	20	180
Language Fluency										
English-Only Students	294	308	327	339	352	374	408	340	22	2,817
Initially Fluent English Proficient (IFEP)	304	313	330	343	355	374	404	343	21	312
Reclassified Fluent English Proficient (RFEP)	304	319	332	343	352	368	392	343	16	495
English-Learner Students	294	306	323	336	346	362	390	335	19	2,364
Unknown	306	315	332	343	353	395	437	346	24	922
Economically Disadvantaged										
No	291	308	328	339	352	372	425	341	22	1,172
Yes	294	308	325	337	348	368	395	338	19	4,262
Unknown	299	311	330	343	353	390	437	343	23	1,476
Special Education Program Participation										
Students Receiving Services	277	299	313	327	339	362	401	328	21	753
Students Not Receiving Services	299	311	328	339	352	372	411	341	20	6,157

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation.

[3] Results for groups with fewer than 11 students are not reported.

Appendix 8.B: Frequency Distributions and Demographic Summaries—October 2011

Table 8.B.1: Frequency Distributions, ELA—October 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	145	0	145	99
440-449	82	0	227	99
430-439	180	1	407	99
420-429	257	1	664	98
410-419	325	1	989	97
400-409	548	2	1,537	96
390-399	646	2	2,183	94
380-389	1,065	3	3,248	91
370-379	1,578	4	4,826	87
360-369	3,328	9	8,154	78
350-359[1]	5,031	14	13,185	64
340-349	5,717	16	18,902	48
330-339	5,296	15	24,198	33
320-329	4,035	11	28,233	22
310-319	3,067	8	31,300	14
300-309	2,233	6	33,533	8
290-299	1,539	4	35,072	4
280-289	765	2	35,837	1
270-279	506	1	36,343	0

[1] Passing Score = 350

Table 8.B.2: Frequency Distributions, Mathematics—October 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	240	1	240	99
440-449	86	0	326	99
430-439	213	1	539	98
420-429	223	1	762	98
410-419	243	1	1,005	97
400-409	628	2	1,633	95
390-399	458	1	2,091	94
380-389	993	3	3,084	91
370-379	1,645	5	4,729	87
360-369	3,001	8	7,730	78
350-359[1]	5,811	16	13,541	62
340-349	5,780	16	19,321	46
330-339	6,497	18	25,818	28
320-329	4,477	13	30,295	15
310-319	3,428	10	33,723	5
300-309	1,401	4	35,124	2
290-299	347	1	35,471	1
280-289	103	0	35,574	0
270-279	106	0	35,680	0

[1] Passing Score = 350

Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	145	0	145	99
440-449	82	0	227	99
430-439	180	1	407	99
420-429	257	1	664	98
410-419	325	1	989	97
403-409[1]	354	1	1,343	96
390-402	840	2	2,183	94
380-389[2]	1,065	3	3,248	91
370-379	1,578	4	4,826	87
360-369	3,328	9	8,154	78
350-359	5,031	14	13,185	64
340-349	5,717	16	18,902	48
330-339	5,296	15	24,198	33
320-329	4,035	11	28,233	22
310-319	3,067	8	31,300	14
300-309	2,233	6	33,533	8
290-299	1,539	4	35,072	4
280-289	765	2	35,837	1
270-279	506	1	36,343	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	240	1	240	99
440-449	86	0	326	99
430-439	213	1	539	98
422-429[1]	223	1	762	98
410-421	243	1	1,005	97
400-409	628	2	1,633	95
390-399	458	1	2,091	94
380-389[2]	993	3	3,084	91
370-379	1,645	5	4,729	87
360-369	3,001	8	7,730	78
350-359	5,811	16	13,541	62
340-349	5,780	16	19,321	46
330-339	6,497	18	25,818	28
320-329	4,477	13	30,295	15
310-319	3,428	10	33,723	5
300-309	1,401	4	35,124	2
290-299	347	1	35,471	1
280-289	103	0	35,574	0
270-279	106	0	35,680	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	36,343	13,185	36	23,158	64	341	55	60	59	47	55	2.1
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	11,969	4,898	41	7,071	59	344	56	61	61	49	56	2.1
Twelfth	22,665	7,535	33	15,130	67	339	54	59	58	46	54	2.1
Adult Education	1,709	752	44	957	56	349	62	65	63	53	58	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	20,908	7,399	35	13,509	65	339	55	58	58	46	53	2.1
Female	15,400	5,774	37	9,626	63	344	54	61	60	48	57	2.2
Unknown	35	12	34	23	66	335	55	57	60	41	48	1.9
Race/Ethnicity												
American Indian or Alaska Native	321	139	43	182	57	344	60	61	61	49	55	2.1
Asian	2,591	817	32	1,774	68	339	50	58	55	48	58	2.0
Pacific Islander	240	85	35	155	65	340	52	58	58	48	55	2.1
Filipino	489	226	46	263	54	349	56	64	63	52	60	2.2
Hispanic or Latino	22,703	7,367	32	15,336	68	338	54	59	58	45	53	2.1
African American	4,126	1,424	35	2,702	65	338	53	57	58	45	52	2.1
White (not of Hispanic origin)	4,782	2,604	54	2,178	46	356	65	65	68	56	62	2.3
Two or More Races	1,091	523	48	568	52	351	62	63	64	53	59	2.2
Language Fluency												
English-Only Students	15,142	6,635	44	8,507	56	346	58	61	63	50	57	2.2
Initially Fluent English Proficient (IFEP)	1,088	566	52	522	48	352	60	64	65	52	61	2.3
Reclassified Fluent English Proficient (RFEP)	1,724	1,011	59	713	41	353	61	67	67	54	63	2.3
English-Learner Students	16,416	4,085	25	12,331	75	333	50	56	54	43	51	2.0
Unknown	1,973	888	45	1,085	55	349	61	65	63	52	58	2.2
Economically Disadvantaged												
No	7,584	3,734	49	3,850	51	352	61	64	65	53	61	2.2
Yes	23,879	7,413	31	16,466	69	337	52	58	57	45	53	2.1
Unknown	4,880	2,038	42	2,842	58	346	58	62	61	50	57	2.1
Special Education Program Participation												
Students Receiving Services	7,757	1,347	17	6,410	83	325	47	49	49	39	46	1.9
Students Not Receiving Services	28,586	11,838	41	16,748	59	346	57	63	62	50	57	2.2

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.B.6: Demographic Summary for All Examinees, Mathematics—October 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	35,680	13,541	38	22,139	62	345	53	52	51	45	39
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	11,680	4,864	42	6,816	58	346	54	54	52	46	40
Twelfth	22,182	7,933	36	14,249	64	344	53	51	50	44	39
Adult Education	1,818	744	41	1074	59	347	58	53	53	48	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	17,942	6,599	37	11,343	63	344	52	53	50	45	37
Female	17,690	6,925	39	10,765	61	346	55	51	52	45	41
Unknown	48	17	35	31	65	340	51	49	47	45	34
Race/Ethnicity											
American Indian or Alaska Native	332	137	41	195	59	346	56	53	52	47	38
Asian	1,239	761	61	478	39	368	60	64	64	60	56
Pacific Islander	234	86	37	148	63	344	53	51	52	44	40
Filipino	441	218	49	223	51	353	56	55	58	51	45
Hispanic or Latino	22,291	7,716	35	14,575	65	342	52	50	50	43	38
African American	4,901	1,506	31	3,395	69	339	51	48	48	41	36
White (not of Hispanic origin)	5,036	2,577	51	2,459	49	355	61	59	57	51	42
Two or More Races	1,206	540	45	666	55	350	58	57	54	48	41
Language Fluency											
English-Only Students	17,340	6,874	40	10,466	60	346	55	53	51	45	39
Initially Fluent English Proficient (IFEP)	1,359	639	47	720	53	352	59	57	55	48	42
Reclassified Fluent English Proficient (RFEP)	2,622	1,372	52	1,250	48	350	60	54	56	48	42
English-Learner Students	12,232	3,727	30	8,505	70	340	48	48	49	43	39
Unknown	2,127	929	44	1,198	56	349	58	55	54	49	39
Economically Disadvantaged											
No	7,952	3,856	48	4,096	52	353	59	57	56	50	44
Yes	22,709	7,589	33	15,120	67	341	51	49	49	43	37
Unknown	5,019	2,096	42	2,923	58	348	56	54	53	48	40
Special Education Program Participation											
Students Receiving Services	6,417	1,078	17	5,339	83	329	42	43	40	36	32
Students Not Receiving Services	29,263	12,463	43	16,800	57	348	56	54	54	47	41

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 8.B.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—October 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		36,343	33,095	91	1,905	5	1343	4	3,248	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	11,969	10,718	90	760	6	491	4	1,251	10
	Twelfth	22,665	20,925	92	983	4	757	3	1,740	8
	Adult Education	1,709	1,452	85	162	9	95	6	257	15
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	20,908	19,154	92	1,098	5	656	3	1,754	8
	Female	15,400	13,909	90	805	5	686	4	1,491	10
	Unknown	35	32	91	2	6	1	3	3	9
Race/Ethnicity	American Indian or Alaskan Native	321	283	88	29	9	9	3	38	12
	Asian	2,591	2,411	93	82	3	98	4	180	7
	Pacific Islander	240	226	94	11	5	3	1	14	6
	Filipino	489	421	86	28	6	40	8	68	14
	Hispanic or Latino	22,703	21,448	94	,888	4	367	2	1,255	6
	African American	4,126	3,839	93	197	5	90	2	287	7
	White (not of Hispanic origin)	4,782	3,581	75	563	12	638	13	1,201	25
	Two or More Races	1,091	886	81	107	10	98	9	205	19
Language Fluency	English-Only Students	15,142	12,929	85	1,195	8	1,018	7	2,213	15
	Initially Fluent English Proficient (IFEP)	1,088	897	82	108	10	83	8	191	18
	Reclassified Fluent English Proficient (RFEP)	1,724	1,509	88	162	9	53	3	215	12
	English-Learner Students	16,416	16,114	98	234	1	68	0	302	2
	Unknown	1,973	1,646	83	206	10	121	6	327	17
Economically Disadvantaged	No	7,584	6,124	81	686	9	774	10	1,460	19
	Yes	23,879	22,785	95	825	3	269	1	1,094	5
	Unknown	4,880	4,186	86	394	8	300	6	694	14
Special Education Program Participation	Receiving Services	7,757	7,623	98	107	1	27	0	134	2
	Not Receiving Services	28,586	25,472	89	1,798	6	1,316	5	3,114	11

[1] Results for groups with fewer than 11 students are not reported.

Table 8.B.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—October 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		35,680	32,596	91	2,322	7	762	2	3,084	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	11,680	10,516	90	911	8	253	2	1,164	10
	Twelfth	22,182	20,435	92	1,256	6	491	2	1,747	8
	Adult Education	1,818	1,645	90	155	9	18	1	173	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	17,942	16,245	91	1,271	7	426	2	1,697	9
	Female	17,690	16,307	92	1,047	6	336	2	1,383	8
	Unknown	48	44	92	4	8	0	0	4	8
Race/Ethnicity	American Indian or Alaskan Native	332	301	91	26	8	5	2	31	9
	Asian	1,239	856	69	212	17	171	14	383	31
	Pacific Islander	234	214	91	18	8	2	1	20	9
	Filipino	441	364	83	57	13	20	5	77	17
	Hispanic or Latino	22,291	21,205	95	941	4	145	1	1,086	5
	African American	4,901	4,670	95	204	4	27	1	231	5
	White (not of Hispanic	5,036	3,978	79	717	14	341	7	1,058	21
	Two or More Races	1,206	1,008	84	147	12	51	4	198	16
Language Fluency	English-Only Students	17,340	15,468	89	1,388	8	484	3	1,872	11
	Initially Fluent English Proficient (IFEP)	1,359	1,163	86	116	9	80	6	196	14
	Reclassified Fluent English Proficient (RFEP)	2,622	2,450	93	149	6	23	1	172	7
	English-Learner Students	12,232	11,654	95	446	4	132	1	578	5
	Unknown	2,127	1,861	87	223	10	43	2	266	13
Economically Disadvantaged	No	7,952	6,547	82	956	12	449	6	1,405	18
	Yes	22,709	21,701	96	858	4	150	1	1,008	4
	Unknown	5,019	4,348	87	508	10	163	3	671	13
Special Education Program Participation	Receiving Services	6,417	6,304	98	107	2	6	0	113	2
	Not Receiving Services	29,263	26,292	90	2,215	8	756	3	2,971	10

[1] Results for groups with fewer than 11 students are not reported.

Table 8.B.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—October 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	277	295	322	340	358	395	433	341	30	36,343
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	277	295	324	344	360	399	433	344	31	11,969
Twelfth	275	293	320	338	354	392	433	339	29	22,665
Adult Education	287	307	330	346	366	405	427	349	30	1,709
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	291	319	340	356	392	427	339	30	20,908
Female	283	301	326	342	358	399	439	344	29	15,400
Unknown	277	279	307	338	362	392	433	335	36	35
Race/Ethnicity										
American Indian or Alaska Native	279	299	322	344	362	392	427	344	30	321
Asian	275	291	320	338	354	392	433	339	30	2,591
Pacific Islander	275	291	324	342	356	381	405	340	27	240
Filipino	283	297	330	346	362	413	450	349	33	489
Hispanic or Latino	277	295	322	338	354	381	413	338	26	22,703
African American	275	289	319	340	356	386	418	338	29	4,126
White (not of Hispanic origin)	277	295	330	352	381	427	450	356	39	4,782
Two or More Races	277	295	326	347	371	418	450	351	37	1,091
Language Fluency										
English-Only Students	275	293	324	344	364	413	446	346	34	15,142
Initially Fluent English Proficient (IFEP)	281	299	332	350	366	413	446	352	33	1,088
Reclassified Fluent English Proficient (RFEP)	287	313	340	354	366	395	422	353	25	1,724
English-Learner Students	277	293	319	334	347	369	389	333	23	16,416
Unknown	277	299	328	346	366	409	433	349	32	1,973
Economically Disadvantaged										
No	279	297	328	347	371	422	450	352	36	7,584
Yes	275	293	320	338	352	378	405	337	26	23,879
Unknown	275	295	324	342	364	409	439	346	33	4,880
Special Education Program Participation										
Students Receiving Services	275	285	305	324	342	366	386	325	25	7,757
Students Not Receiving Services	281	301	328	344	360	402	439	346	30	28,586

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	295	308	326	343	357	394	438	345	27	35,680
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	295	308	328	344	360	397	438	346	27	11,680
Twelfth	295	308	326	341	355	394	438	344	26	22,182
Adult Education	303	314	332	344	358	394	417	347	24	1,818
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	292	306	324	341	357	400	438	344	28	17,942
Female	301	312	330	343	357	392	438	346	25	17,690
Unknown	292	306	320	338	355	384	417	340	26	48
Race/Ethnicity										
American Indian or Alaska Native	303	312	328	344	358	394	432	346	25	332
Asian	306	317	339	357	394	450	450	368	39	1,239
Pacific Islander	301	308	330	343	355	392	413	344	24	234
Filipino	301	310	332	348	366	417	446	353	31	441
Hispanic or Latino	295	308	326	341	355	378	409	342	22	22,291
African American	289	306	323	339	351	378	409	339	23	4,901
White (not of Hispanic origin)	295	308	330	350	372	426	450	355	35	5,036
Two or More Races	298	308	328	346	364	417	450	350	32	1,206
Language Fluency										
English-Only Students	295	308	326	343	358	403	446	346	29	17,340
Initially Fluent English Proficient (IFEP)	301	310	332	346	362	426	450	352	32	1,359
Reclassified Fluent English Proficient (RFEP)	303	319	339	350	360	384	417	350	21	2,622
English-Learner Students	295	308	324	339	351	378	422	340	23	12,232
Unknown	298	312	332	344	362	403	438	349	27	2,127
Economically Disadvantaged										
No	298	310	332	348	366	422	450	353	33	7,952
Yes	295	308	326	341	353	378	409	341	22	22,709
Unknown	298	310	328	344	362	409	446	348	30	5,019
Special Education Program Participation										
Students Receiving Services	283	301	314	326	343	364	387	329	21	6,417
Students Not Receiving Services	301	312	332	344	358	400	446	348	27	29,263

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Appendix 8.C: Frequency Distributions and Demographic Summaries—November 2011

Table 8.C.1: Frequency Distributions, ELA—November 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	399	0	399	99
440-449	547	1	946	99
430-439	747	1	1,693	98
420-429	945	1	2,638	97
410-419	1,107	1	3,745	96
400-409	1,869	2	5,614	94
390-399	2,206	2	7,820	92
380-389	2,893	3	10,713	89
370-379	5,471	6	16,184	83
360-369	10,139	11	26,323	72
350-359[1]	10,263	11	36,586	61
340-349	13,489	14	50,075	47
330-339	14,184	15	64,259	31
320-329	9,016	10	73,275	22
310-319	7,264	8	80,539	14
300-309	6,589	7	87,128	7
290-299	3,518	4	90,646	3
280-289	1,816	2	92,462	1
270-279	1,139	1	93,601	0

[1] Passing Score = 350

Table 8.C.2: Frequency Distributions, Mathematics—November 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	672	1	672	99
440-449	324	0	996	99
430-439	673	1	1,669	98
420-429	678	1	2,347	97
410-419	722	1	3,069	97
400-409	1,204	1	4,273	95
390-399	1,919	2	6,192	93
380-389	2,602	3	8,794	91
370-379	5,248	6	14,042	85
360-369	8,758	9	22,800	76
350-359[1]	16,041	17	38,841	58
340-349	15,111	16	53,952	42
330-339	15,875	17	69,827	25
320-329	10,421	11	80,248	14
310-319	7,950	9	88,198	6
300-309	3,507	4	91,705	2
290-299	1,223	1	92,928	0
280-289	226	0	93,154	0
270-279	214	0	93,368	0

[1] Passing Score = 350

Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	399	0	399	99
440-449	547	1	946	99
430-439	747	1	1,693	98
420-429	945	1	2,638	97
410-419	1,107	1	3,745	96
403-409[1]	1,227	1	4,972	95
390-402	2,848	3	7,820	92
380-389[2]	2,893	3	10,713	89
370-379	5,471	6	16,184	83
360-369	10,139	11	26,323	72
350-359	10,263	11	36,586	61
340-349	13,489	14	50,075	47
330-339	14,184	15	64,259	31
320-329	9,016	10	73,275	22
310-319	7,264	8	80,539	14
300-309	6,589	7	87,128	7
290-299	3,518	4	90,646	3
280-289	1,816	2	92,462	1
270-279	1,139	1	93,601	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	672	1	672	99
440-449	324	0	996	99
430-439	673	1	1,669	98
422-429[1]	345	0	2,014	98
410-421	1,055	1	3,069	97
400-409	1,204	1	4,273	95
390-399	1,919	2	6,192	93
380-389[2]	2,602	3	8,794	91
370-379	5,248	6	14,042	85
360-369	8,758	9	22,800	76
350-359	16,041	17	38,841	58
340-349	15,111	16	53,952	42
330-339	15,875	17	69,827	25
320-329	10,421	11	80,248	14
310-319	7,950	9	88,198	6
300-309	3,507	4	91,705	2
290-299	1,223	1	92,928	0
280-289	226	0	93,154	0
270-279	214	0	93,368	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	93,601	36,586	39	57,015	61	344	68	58	62	49	56	2.0
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	59,976	25,374	42	34,602	58	346	69	59	63	50	57	2.1
Twelfth	30,774	10,011	33	20,763	67	340	65	55	59	47	54	2.0
Adult Education	2,851	1,201	42	1,650	58	349	70	61	65	56	57	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	54,457	19,619	36	34,838	64	341	67	56	61	47	54	2.0
Female	39,064	16,944	43	22,120	57	348	69	59	64	53	58	2.1
Unknown	80	23	29	57	71	334	62	53	58	45	48	1.8
Race/Ethnicity												
American Indian or Alaska Native	661	269	41	392	59	345	69	59	63	48	56	2.0
Asian	5,404	2,140	40	3,264	60	346	63	57	63	52	61	2.0
Pacific Islander	654	274	42	380	58	348	70	59	64	51	59	2.1
Filipino	1,609	819	51	790	49	355	72	61	67	57	62	2.2
Hispanic or Latino	60,336	21,442	36	38,894	64	341	67	56	61	48	54	2.0
African American	9,155	3,277	36	5,878	64	340	67	56	60	46	53	2.0
White (not of Hispanic origin)	13,845	7,450	54	6,395	46	357	74	64	68	56	64	2.2
Two or More Races	1,937	915	47	1,022	53	352	72	62	66	52	60	2.1
Language Fluency												
English-Only Students	38,761	17,963	46	20,798	54	350	72	61	65	51	59	2.1
Initially Fluent English Proficient (IFEP)	3,532	1,980	56	1,552	44	358	76	65	69	56	64	2.2
Reclassified Fluent English Proficient (RFEP)	6,892	4,418	64	2,474	36	358	79	67	70	58	65	2.2
English-Learner Students	41,006	10,670	26	30,336	74	334	62	52	57	45	51	1.9
Unknown	3,410	1,555	46	1,855	54	351	71	62	66	56	58	2.1
Economically Disadvantaged												
No	20,626	10,955	53	9,671	47	357	74	64	68	56	63	2.2
Yes	63,535	21,517	34	42,018	66	339	66	55	60	47	53	2.0
Unknown	9,440	4,114	44	5,326	56	348	69	60	64	53	58	2.1
Special Education Program Participation												
Students Receiving Services	21,084	3,728	18	17,356	82	326	57	47	52	38	45	1.8
Students Not Receiving Services	72,517	32,858	45	39,659	55	349	71	61	65	53	59	2.1

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.C.6: Demographic Summary for All Examinees, Mathematics—November 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics [2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	93,368	38,841	42	54,527	58	347	58	54	53	48	40
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	59,084	25,987	44	33,097	56	348	59	54	53	48	41
Twelfth	31,282	11,617	37	19,665	63	344	57	52	51	47	39
Adult Education	3,002	1,237	41	1,765	59	348	62	52	54	51	39
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	47,631	19,558	41	28,073	59	346	58	54	52	48	39
Female	45,645	19,251	42	26,394	58	347	59	53	53	48	42
Unknown	92	32	35	60	65	341	54	52	46	46	37
Race/Ethnicity											
American Indian or Alaska Native	735	299	41	436	59	345	58	53	51	48	38
Asian	3,188	1,984	62	1,204	38	369	64	66	65	59	57
Pacific Islander	643	286	44	357	56	349	58	56	56	48	42
Filipino	1,482	829	56	653	44	357	63	59	60	53	48
Hispanic or Latino	59,959	22,575	38	37,384	62	343	56	51	51	46	39
African American	10,760	3,791	35	6,969	65	341	56	51	49	44	38
White (not of Hispanic origin)	14,463	8,062	56	6,401	44	357	66	61	59	54	45
Two or More Races	2,138	1,015	47	1,123	53	351	62	57	55	50	42
Language Fluency											
English-Only Students	44,390	20,043	45	24,347	55	349	61	55	54	49	41
Initially Fluent English Proficient (IFEP)	4,294	2,331	54	1,963	46	356	65	58	58	53	45
Reclassified Fluent English Proficient (RFEP)	9,193	5,178	56	4,015	44	353	66	56	58	51	43
English-Learner Students	31,868	9,659	30	22,209	70	340	51	49	48	45	39
Unknown	3,623	1,630	45	1,993	55	351	62	55	56	52	41
Economically Disadvantaged											
No	22,154	11,807	53	10,347	47	356	64	59	58	53	45
Yes	61,346	22,629	37	38,717	63	343	56	51	50	46	39
Unknown	9,868	4,405	45	5,463	55	349	60	55	54	50	41
Special Education Program Participation											
Students Receiving Services	18,266	3,409	19	14,857	81	331	46	46	40	39	33
Students Not Receiving Services	75,102	35,432	47	39,670	53	350	61	55	56	50	42

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 8.C.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—November 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		93,601	82,888	89	5,741	6	4,972	5	10,713	11
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	59,976	52,595	88	4,088	7	3,293	5	7,381	12
	Twelfth	30,774	27,874	91	1,403	5	1,497	5	2,900	9
	Adult Education	2,851	2,419	85	250	9	182	6	432	15
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	54,457	49,017	90	3,111	6	2,329	4	5,440	10
	Female	39,064	33,797	87	2,627	7	2,640	7	5,267	13
	Unknown	80	74	93	3	4	3	4	6	8
Race/Ethnicity	American Indian or Alaskan Native	661	576	87	41	6	44	7	85	13
	Asian	5,404	4,682	87	301	6	421	8	722	13
	Pacific Islander	654	568	87	50	8	36	6	86	13
	Filipino	1,609	1,292	80	127	8	190	12	317	20
	Hispanic or Latino	60,336	55,739	92	2,946	5	1,651	3	4,597	8
	African American	9,155	8,340	91	493	5	322	4	815	9
	White (not of Hispanic origin)	13,845	10,206	74	1,561	11	2,078	15	3,639	26
	Two or More Races	1,937	1,485	77	222	11	230	12	452	23
Language Fluency	English-Only Students	38,761	31,757	82	3,365	9	3,639	9	7,004	18
	Initially Fluent English Proficient (IFEP)	3,532	2,706	77	384	11	442	13	826	23
	Reclassified Fluent English Proficient (RFEP)	6,892	5,702	83	821	12	369	5	1,190	17
	English-Learner Students	41,006	39,943	97	836	2	227	1	1,063	3
	Unknown	3,410	2,780	82	335	10	295	9	630	18
Economically Disadvantaged	No	20,626	15,732	76	2,111	10	2,783	13	4,894	24
	Yes	63,535	59,286	93	2,808	4	1,441	2	4,249	7
	Unknown	9,440	7,870	83	822	9	748	8	1,570	17
Special Education Program Participation	Receiving Services	21,084	20,664	98	300	1	120	1	420	2
	Not Receiving Services	72,517	62,224	86	5,441	8	4,852	7	10,293	14

[1] Results for groups with fewer than 11 students are not reported

Table 8.C.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—November 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		93,368	84,574	91	6,780	7	2,014	2	8,794	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	59,084	53,139	90	4,626	8	1,319	2	5,945	10
	Twelfth	31,282	28,697	92	1,943	6	642	2	2,585	8
	Adult Education	3,002	2,738	91	211	7	53	2	264	9
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	47,631	42,714	90	3,725	8	1,192	3	4,917	10
	Female	45,645	41,775	92	3,048	7	822	2	3,870	8
	Unknown	92	85	92	7	8	0	0	7	8
Race/Ethnicity	American Indian or Alaskan Native	735	665	90	55	7	15	2	70	10
	Asian	3,188	2,156	68	534	17	498	16	1,032	32
	Pacific Islander	643	567	88	66	10	10	2	76	12
	Filipino	1,482	1,209	82	196	13	77	5	273	18
	Hispanic or Latino	59,959	56,565	94	2,965	5	429	1	3,394	6
	African American	10,760	10,200	95	489	5	71	1	560	5
	White (not of Hispanic origin)	14,463	11,397	79	2,243	16	823	6	3,066	21
	Two or More Races	2,138	1,815	85	232	11	91	4	323	15
Language Fluency	English-Only Students	44,390	38,924	88	4,195	9	1,271	3	5,466	12
	Initially Fluent English Proficient (IFEP)	4,294	3,574	83	493	11	227	5	720	17
	Reclassified Fluent English Proficient (RFEP)	9,193	8,406	91	688	7	99	1	787	9
	English-Learner Students	31,868	30,533	96	1,049	3	286	1	1,335	4
	Unknown	3,623	3,137	87	355	10	131	4	486	13
Economically Disadvantaged	No	22,154	17,935	81	3,005	14	1,214	5	4,219	19
	Yes	61,346	58,047	95	2,801	5	498	1	3,299	5
	Unknown	9,868	8,592	87	974	10	302	3	1,276	13
Special Education Program Participation	Receiving Services	18,266	17,964	98	278	2	24	0	302	2
	Not Receiving Services	75,102	66,610	89	6,502	9	1,990	3	8,492	11

[1] Results for groups with fewer than 11 students are not reported

Table 8.C.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—November 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	294	322	343	362	403	442	344	32	93,601
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	278	296	324	345	364	403	442	346	32	59,976
Twelfth	275	292	319	337	356	400	436	340	32	30,774
Adult Education	284	306	330	345	367	407	436	349	30	2,851
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	292	320	341	360	396	436	341	32	54,457
Female	282	300	328	345	364	411	442	348	32	39,064
Unknown	278	290	305	330	357	393	425	334	33	80
Race/Ethnicity										
American Indian or Alaska Native	278	296	320	343	364	407	436	345	34	661
Asian	280	296	324	343	362	415	449	346	34	5,404
Pacific Islander	278	300	328	345	364	407	442	348	31	654
Filipino	286	304	332	351	374	425	449	355	35	1,609
Hispanic or Latino	278	296	322	341	358	387	420	341	28	60,336
African American	275	290	319	339	360	393	425	340	31	9,155
White (not of Hispanic origin)	275	296	330	353	382	430	450	357	40	13,845
Two or More Races	275	290	324	347	376	430	450	352	39	1,937
Language Fluency										
English-Only Students	275	294	326	347	369	420	449	350	36	38,761
Initially Fluent English Proficient (IFEP)	284	306	336	353	379	425	450	358	35	3,532
Reclassified Fluent English Proficient (RFEP)	288	315	343	358	374	403	425	358	26	6,892
English-Learner Students	278	294	319	336	351	374	393	334	24	41,006
Unknown	280	300	328	347	371	415	442	351	34	3,410
Economically Disadvantaged										
No	280	300	332	353	379	430	450	357	38	20,626
Yes	275	294	320	339	358	384	415	339	28	63,535
Unknown	275	294	324	345	367	415	442	348	35	9,440
Special Education Program Participation										
Students Receiving Services	275	284	306	326	343	369	390	326	26	21,084
Students Not Receiving Services	282	302	330	347	367	411	442	349	32	72,517

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	296	309	328	344	359	398	445	347	27	93,368
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	296	309	330	346	360	398	445	348	27	59,084
Twelfth	294	306	327	343	357	395	445	344	27	31,282
Adult Education	304	315	334	344	359	393	437	348	24	3,002
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	294	306	327	343	360	401	445	346	29	47,631
Female	299	311	332	344	359	395	437	347	25	45,645
Unknown	285	304	321	338	357	393	420	341	26	92
Race/Ethnicity										
American Indian or Alaska Native	288	304	328	343	359	393	431	345	27	735
Asian	302	317	339	359	398	450	450	369	41	3,188
Pacific Islander	296	311	332	346	362	395	431	349	26	643
Filipino	302	315	337	351	370	425	450	357	30	1,482
Hispanic or Latino	296	309	328	343	357	380	416	343	23	59,959
African American	291	306	327	341	355	380	416	341	24	10,760
White (not of Hispanic origin)	296	311	336	353	374	425	450	357	34	14,463
Two or More Races	291	309	330	348	366	416	450	351	32	2,138
Language Fluency										
English-Only Students	294	309	330	346	362	404	445	349	29	44,390
Initially Fluent English Proficient (IFEP)	299	313	336	351	368	425	450	356	31	4,294
Reclassified Fluent English Proficient (RFEP)	306	321	339	351	362	390	425	353	21	9,193
English-Learner Students	296	306	325	339	351	376	416	340	22	31,868
Unknown	302	315	334	346	362	412	450	351	29	3,623
Economically Disadvantaged										
No	299	313	336	351	370	425	450	356	33	22,154
Yes	296	309	327	343	355	380	416	343	23	61,346
Unknown	296	309	330	346	362	408	445	349	29	9,868
Special Education Program Participation										
Students Receiving Services	285	299	317	330	344	366	387	331	21	18,266
Students Not Receiving Services	302	313	334	348	362	404	445	350	27	75,102

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Appendix 8.D: Frequency Distributions and Demographic Summaries—December 2011

Table 8.D.1: Frequency Distributions, ELA—December 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9	0	9	99
440-449	4	0	13	99
430-439	1	0	14	99
420-429	7	0	21	99
410-419	20	1	41	98
400-409	18	1	59	97
390-399	29	1	88	96
380-389	41	2	129	94
370-379	87	4	216	91
360-369	124	5	340	85
350-359[1]	322	14	662	71
340-349	431	19	1,093	53
330-339	406	18	1,499	35
320-329	299	13	1,798	22
310-319	231	10	2,029	13
300-309	152	7	2,181	6
290-299	76	3	2,257	3
280-289	41	2	2,298	1
270-279	21	1	2,319	0

[1]Passing Score = 350

Table 8.D.2: Frequency Distributions, Mathematics—December 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5	0	5	99
440-449	2	0	7	99
430-439	3	0	10	99
420-429	4	0	14	99
410-419	10	0	24	99
400-409	13	1	37	98
390-399	14	1	51	98
380-389	46	2	97	95
370-379	83	4	180	91
360-369	170	8	350	83
350-359[1]	368	18	718	66
340-349	424	20	1,142	45
330-339	463	22	1,605	23
320-329	268	13	1,873	10
310-319	123	6	1,996	4
300-309	70	3	2,066	1
290-299	19	1	2,085	0
280-289	1	0	2,086	0
270-279	1	0	2,087	0

[1]Passing Score = 350

Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9	0	9	99
440-449	4	0	13	99
430-439	1	0	14	99
420-429	7	0	21	99
410-419	20	1	41	98
403-409[1]	11	0	52	98
390-402	36	2	88	96
380-389[2]	41	2	129	94
370-379	87	4	216	91
360-369	124	5	340	85
350-359	322	14	662	71
340-349	431	19	1,093	53
330-339	406	18	1,499	35
320-329	299	13	1,798	22
310-319	231	10	2,029	13
300-309	152	7	2,181	6
290-299	76	3	2,257	3
280-289	41	2	2,298	1
270-279	21	1	2,319	0

[1] Advanced-Level Cut = 403
[2] Proficient-Level Cut = 380

Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5	0	5	99
440-449	2	0	7	99
430-439	3	0	10	99
422-429[1]	4	0	14	99
410-421	10	0	24	99
400-409	13	1	37	98
390-399	14	1	51	98
380-389[2]	46	2	97	95
370-379	83	4	180	91
360-369	170	8	350	83
350-359	368	18	718	66
340-349	424	20	1,142	45
330-339	463	22	1,605	23
320-329	268	13	1,873	10
310-319	123	6	1,996	4
300-309	70	3	2,066	1
290-299	19	1	2,085	0
280-289	1	0	2,086	0
270-279	1	0	2,087	0

[1] Advanced-Level Cut = 422
[2] Proficient-Level Cut = 380

Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	2,319	662	29	1,657	71	339	55	56	59	45	55	2.0
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	219	82	37	137	63	344	56	58	63	47	59	2.1
Twelfth	1,667	398	24	1,269	76	335	52	54	56	43	53	2.0
Adult Education	433	182	42	251	58	350	63	64	64	51	59	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	1,148	321	28	827	72	337	55	56	57	44	54	2.0
Female	1,168	340	29	828	71	340	54	57	60	45	55	2.1
Unknown	3	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	5	-	-	-	-	-	-	-	-	-	-	-
Asian	149	24	16	125	84	331	49	50	55	44	51	1.9
Pacific Islander	21	6	29	15	71	337	56	49	55	41	56	2.1
Filipino	35	5	14	30	86	335	52	53	58	43	53	2.0
Hispanic or Latino	1,640	455	28	1,185	72	338	55	56	59	44	54	2.0
African American	232	68	29	164	71	340	54	57	60	45	55	2.1
White (not of Hispanic origin)	184	78	42	106	58	344	58	59	59	48	58	2.1
Two or More Races	53	23	43	30	57	348	58	65	63	49	61	2.0
Language Fluency												
English-Only Students	539	199	37	340	63	344	56	59	62	46	59	2.1
Initially Fluent English Proficient (IFEP)	38	20	53	18	47	344	62	60	59	48	58	2.1
Reclassified Fluent English Proficient (RFEP)	72	36	50	36	50	350	61	63	67	48	64	2.1
English-Learner Students	1,238	221	18	1,017	82	331	51	51	55	42	51	2.0
Unknown	432	186	43	246	57	350	63	64	65	51	59	2.2
Economically Disadvantaged												
No	276	99	36	177	64	344	57	59	61	47	58	2.1
Yes	1,572	369	23	1,203	77	335	52	54	57	43	53	2.0
Unknown	471	194	41	277	59	348	62	63	63	50	58	2.1
Special Education Program Participation												
Students Receiving Services	248	36	15	212	85	325	47	47	50	38	48	1.9
Students Not Receiving Services	2,071	626	30	1445	70	340	56	57	60	45	56	2.1

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2011

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	2,087	718	34	1,369	66	343	52	46	51	44	36
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	180	85	47	95	53	349	54	51	54	50	40
Twelfth	1,445	443	31	1,002	69	341	51	45	49	42	36
Adult Education	462	190	41	272	59	348	55	49	56	48	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	807	304	38	503	62	344	52	49	52	44	36
Female	1,277	413	32	864	68	342	52	45	51	44	36
Unknown	3	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity											
American Indian or Alaska Native	8	-	-	-	-	-	-	-	-	-	-
Asian	44	24	55	20	45	357	53	53	58	56	53
Pacific Islander	19	7	37	12	63	338	45	44	48	40	36
Filipino	21	10	48	11	52	346	54	45	55	52	37
Hispanic or Latino	1,399	445	32	954	68	342	51	45	51	44	36
African American	345	126	37	219	63	343	53	47	51	44	36
White (not of Hispanic origin)	196	82	42	114	58	348	56	50	54	47	39
Two and More Races	55	20	36	35	64	345	56	49	51	46	36
Language Fluency											
English-Only Students	708	270	38	438	62	344	55	47	52	45	36
Initially Fluent English Proficient (IFEP)	45	20	44	25	56	347	55	51	55	45	36
Reclassified Fluent English Proficient (RFEP)	140	72	51	68	49	350	58	53	57	47	38
English-Learner Students	741	171	23	570	77	337	46	41	46	41	36
Unknown	453	185	41	268	59	349	55	50	56	48	37
Economically Disadvantaged											
No	305	134	44	171	56	348	56	49	55	47	39
Yes	1,301	383	29	918	71	340	50	44	49	42	36
Unknown	481	201	42	280	58	348	56	49	56	48	37
Special Education Program Participation											
Students Receiving Services	176	33	19	143	81	332	43	40	42	37	34
Students Not Receiving Services	1,911	685	36	1,226	64	344	53	47	52	45	37

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 8.D.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA —December 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		2,319	2,190	94	77	3	52	2	129	6
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	219	202	92	12	5	5	2	17	8
	Twelfth	1,667	1,614	97	34	2	19	1	53	3
	Adult Education	433	374	86	31	7	28	6	59	14
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	1,148	1,082	94	44	4	22	2	66	6
	Female	1,168	1,105	95	33	3	30	3	63	5
	Unknown	3	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	5	-	-	-	-	-	-	-	-
	Asian	149	146	98	1	1	2	1	3	2
	Pacific Islander	21	20	95	0	0	1	5	1	5
	Filipino	35	32	91	2	6	1	3	3	9
	Hispanic or Latino	1,640	1,568	96	47	3	25	2	72	4
	African American	232	221	95	7	3	4	2	11	5
	White (not of Hispanic origin)	184	153	83	15	8	16	9	31	17
	Two or More Races	53	46	87	4	8	3	6	7	13
Language Fluency	English-Only Students	539	490	91	28	5	21	4	49	9
	Initially Fluent English Proficient (IFEP)	38	37	97	1	3	0	0	1	3
	Reclassified Fluent English Proficient (RFEP)	72	68	94	4	6	0	0	4	6
	English-Learner Students	1,238	1,223	99	12	1	3	0	15	1
	Unknown	432	372	86	32	7	28	6	60	14
Economically Disadvantaged	No	276	245	89	16	6	15	5	31	11
	Yes	1,572	1,536	98	30	2	6	0	36	2
	Unknown	471	409	87	31	7	31	7	62	13
Special Education Program Participation	Receiving Services	248	246	99	2	1	0	0	2	1
	Not Receiving Services	2,071	1,944	94	75	4	52	3	127	6

[1] Results for groups with fewer than 11 students are not reported.

Table 8.D.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—December 2011

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		2,087	1,990	95	83	4	14	1	97	5
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	180	167	93	9	5	4	2	13	7
	Twelfth	1,445	1,398	97	42	3	5	0	47	3
	Adult Education	462	425	92	32	7	5	1	37	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	807	757	94	40	5	10	1	50	6
	Female	1,277	1,230	96	43	3	4	0	47	4
	Unknown	3	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	8	-	-	-	-	-	-	-	-
	Asian	44	36	82	7	16	1	2	8	18
	Pacific Islander	19	19	100	0	0	0	0	0	0
	Filipino	21	20	95	1	5	0	0	1	5
	Hispanic or Latino	1,399	1,347	96	45	3	7	1	52	4
	African American	345	335	97	10	3	0	0	10	3
	White (not of Hispanic origin)	196	176	90	15	8	5	3	20	10
	Two or More Races	55	50	91	4	7	1	2	5	9
Language Fluency	English-Only Students	708	673	95	27	4	8	1	35	5
	Initially Fluent English Proficient (IFEP)	45	41	91	4	9	0	0	4	9
	Reclassified Fluent English Proficient (RFEP)	140	135	96	5	4	0	0	5	4
	English-Learner Students	741	727	98	13	2	1	0	14	2
	Unknown	453	414	91	34	8	5	1	39	9
Economically Disadvantaged	No	305	281	92	19	6	5	2	24	8
	Yes	1,301	1,264	97	32	2	5	0	37	3
	Unknown	481	445	93	32	7	4	1	36	7
Special Education Program Participation	Receiving Services	176	173	98	3	2	0	0	3	2
	Not Receiving Services	1,911	1,817	95	80	4	14	1	94	5

[1] Results for groups with fewer than 11 students are not reported.

Table 8.D.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—December 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	284	296	323	339	353	384	418	339	26	2,319
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	284	304	329	343	357	392	410	344	26	219
Twelfth	276	293	319	337	349	371	406	335	24	1,667
Adult Education	296	307	331	345	366	406	443	350	30	433
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	276	293	321	337	353	384	423	337	27	1,148
Female	286	300	325	339	353	384	418	340	25	1,168
Unknown	-	-	-	-	-	-	-	-	-	3
Race/Ethnicity										
American Indian or Alaska Native	-	-	-	-	-	-	-	-	-	5
Asian	289	296	315	333	343	364	450	331	24	149
Pacific Islander	291	309	319	331	351	364	450	337	32	21
Filipino	275	291	313	335	345	402	429	335	30	35
Hispanic or Latino	284	298	323	339	351	378	410	338	25	1,640
African American	286	304	325	339	353	378	406	340	23	232
White (not of Hispanic origin)	275	291	314	345	362	414	450	344	37	184
Two or More Races	275	286	331	347	359	414	450	348	33	53
Language Fluency										
English-Only Students	286	302	325	343	357	395	443	344	28	539
Initially Fluent English Proficient (IFEP)	284	286	333	351	359	378	384	344	25	38
Reclassified Fluent English Proficient (RFEP)	309	319	337	350	359	381	392	350	18	72
English-Learner Students	275	291	317	333	347	364	384	331	22	1,238
Unknown	293	307	332	347	364	406	443	350	30	432
Economically Disadvantaged										
No	275	293	325	343	357	406	450	344	32	276
Yes	281	296	321	337	349	371	392	335	23	1,572
Unknown	291	304	329	345	364	410	443	348	30	471
Special Education Program Participation										
Students Receiving Services	275	286	311	325	341	362	373	325	23	248
Students Not Receiving Services	284	300	325	339	353	386	423	340	26	2,071

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	299	311	330	341	353	378	411	343	21	2,087
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	299	311	335	348	362	383	428	349	24	180
Twelfth	299	309	328	341	351	372	399	341	20	1,445
Adult Education	307	318	333	344	359	389	434	348	23	462
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	299	309	330	343	355	382	424	344	23	807
Female	302	314	330	341	351	376	405	342	19	1,277
Unknown	-	-	-	-	-	-	-	-	-	3
Race/Ethnicity										
American Indian or Alaska Native	-	-	-	-	-	-	-	-	-	8
Asian	316	322	340	350	366	411	450	357	28	44
Pacific Islander	304	304	322	337	355	376	376	338	20	19
Filipino	302	311	339	346	355	378	380	346	20	21
Hispanic or Latino	302	311	330	341	351	374	405	342	20	1,399
African American	299	311	332	343	355	374	391	343	19	345
White (not of Hispanic origin)	296	311	333	344	361	399	448	348	26	196
Two or More Races	299	314	328	343	355	405	434	345	26	55
Language Fluency										
English-Only Students	299	311	332	343	355	378	424	344	22	708
Initially Fluent English Proficient (IFEP)	296	322	332	346	359	384	411	347	21	45
Reclassified Fluent English Proficient (RFEP)	304	324	341	350	360	378	389	350	16	140
English-Learner Students	299	309	324	337	348	364	391	337	18	741
Unknown	307	320	333	344	359	391	434	349	23	453
Economically Disadvantaged										
No	302	311	333	346	360	384	440	348	24	305
Yes	299	309	328	339	351	370	399	340	19	1,301
Unknown	307	318	333	344	359	389	415	348	22	481
Special Education Program Participation										
Students Receiving Services	296	302	316	332	344	370	399	332	21	176
Students Not Receiving Services	302	314	332	343	353	378	411	344	21	1,911

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

Appendix 8.E: Frequency Distributions and Demographic Summaries—February 2012

Table 8.E.1: Frequency Distributions, ELA—February 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	4,418	3	4,418	97
440-449	2,125	1	6,543	96
430-439	5,590	4	12,133	92
420-429	7,012	4	19,145	88
410-419	12,020	8	31,165	80
400-409	8,254	5	39,419	75
390-399	16,183	10	55,602	65
380-389	11,339	7	66,941	57
370-379	13,273	8	80,214	49
360-369	14,166	9	94,380	40
350-359[1]	13,190	8	107,570	32
340-349	12,135	8	119,705	24
330-339	10,555	7	130,260	17
320-329	8,515	5	138,775	12
310-319	6,864	4	145,639	7
300-309	5,140	3	150,,779	4
290-299	3,505	2	154,284	2
280-289	1,793	1	156,077	1
270-279	1,195	1	157,272	0

[1] Passing Score = 350

Table 8.E.2: Frequency Distributions, Mathematics—February 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9,313	6	9,313	94
440-449	3,453	2	12,766	92
430-439	6,898	4	19,664	87
420-429	6,902	4	26,566	83
410-419	6,541	4	33,107	79
400-409	9,506	6	42,613	72
390-399	12,030	8	54,643	65
380-389	11,366	7	66,009	57
370-379	13,099	9	79,108	49
360-369	12,483	8	91,591	41
350-359[1]	15,432	10	107,023	31
340-349	13,060	8	120,083	22
330-339	13,516	9	133,599	13
320-329	8,808	6	142,407	8
310-319	6,833	4	149,240	3
300-309	3,331	2	152,571	1
290-299	1,181	1	153,752	0
280-289	228	0	153,980	0
270-279	201	0	154,181	0

[1] Passing Score = 350

Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	4,418	3	4,418	97
440-449	2,125	1	6,543	96
430-439	5,590	4	12,133	92
420-429	7,012	4	19,145	88
410-419	12,020	8	31,165	80
403-409[1]	8,254	5	39,419	75
390-402	16,183	10	55,602	65
380-389[2]	11,339	7	66,941	57
370-379	13,273	8	80,214	49
360-369	14,166	9	94,380	40
350-359	13,190	8	107,570	32
340-349	12,135	8	119,705	24
330-339	10,555	7	130,260	17
320-329	8,515	5	138,775	12
310-319	6,864	4	145,639	7
300-309	5,140	3	150,779	4
290-299	3,505	2	154,284	2
280-289	1,793	1	156,077	1
270-279	1,195	1	157,272	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9,313	6	9,313	94
440-449	3,453	2	12,766	92
430-439	6,898	4	19,664	87
422-429[1]	3,460	2	23,124	85
410-421	9,983	6	33,107	79
400-409	9,506	6	42,613	72
390-399	12,030	8	54,643	65
380-389[2]	11,366	7	66,009	57
370-379	13,099	9	79,108	49
360-369	12,483	8	91,591	41
350-359	15,432	10	107,023	31
340-349	13,060	8	120,083	22
330-339	13,516	9	133,599	13
320-329	8,808	6	142,407	8
310-319	6,833	4	149,240	3
300-309	3,331	2	152,571	1
290-299	1,181	1	153,752	0
280-289	228	0	153,980	0
270-279	201	0	154,181	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.E.5: Demographic Summary for All Examinees, ELA February 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	157,272	107,570	68	49,702	32	371	76	70	73	64	67	2.3
Grade[3]												
Tenth	113,554	95,332	84	18,222	16	384	83	78	79	71	74	2.5
Eleventh	17,695	5,487	31	12,208	69	337	60	52	57	46	50	2.0
Twelfth	23,196	5,617	24	17,579	76	332	57	49	54	43	48	1.9
Adult Education	2,827	1,134	40	1,693	60	346	65	60	61	52	56	2.0
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	83,309	53,097	64	30,212	36	365	74	68	70	62	64	2.2
Female	73,872	54,442	74	19,430	26	377	78	73	75	67	71	2.5
Unknown	91	31	34	60	66	340	62	53	59	46	51	2.0
Race/Ethnicity												
American Indian or Alaska Native	1,435	1,018	71	417	29	368	78	70	74	62	66	2.2
Asian	14,920	11,313	76	3,607	24	383	79	76	76	71	74	2.5
Pacific Islander	1,082	719	66	363	34	365	75	67	71	60	65	2.3
Filipino	4,765	4,004	84	761	16	387	83	78	79	72	76	2.6
Hispanic or Latino	71,958	40,979	57	30,979	43	357	70	64	67	57	61	2.2
African American	13,178	7,298	55	5,880	45	355	70	62	67	55	59	2.2
White (not of Hispanic origin)	45,246	38,568	85	6,678	15	389	85	80	81	74	76	2.5
Two or More Races	4,688	3,671	78	1,017	22	383	82	76	78	69	72	2.5
Language Fluency												
English-Only Students	89,840	68,919	77	20,921	23	379	81	75	77	68	71	2.4
Initially Fluent English Proficient (IFEP)	8,510	7,295	86	1,215	14	390	84	80	81	73	77	2.6
Reclassified Fluent English Proficient (RFEP)	22,773	20,549	90	2,224	10	385	83	80	80	72	76	2.5
English-Learner Students	32,452	9,017	28	23,435	72	334	57	51	55	45	50	1.9
Unknown	3,697	1,790	48	1,907	52	353	68	63	64	55	59	2.1
Economically Disadvantaged												
No	60,597	52,139	86	8,458	14	391	85	80	82	74	77	2.6
Yes	85,186	48,361	57	36,825	43	357	70	64	67	57	61	2.2
Unknown	11,489	7,070	62	4,419	38	365	74	68	70	61	64	2.3
Special Education Program Participation												
Students Receiving Services	20,198	5,223	26	14,975	74	331	56	48	54	43	45	1.9
Students Not Receiving Services	137,074	102,347	75	34,727	25	377	79	74	76	67	71	2.4

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.E.6: Demographic Summary for All Examinees, Mathematics—February 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics [2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	154,181	107,023	69	47,158	31	375	72	67	68	64	59
Grade											
Tenth	113,147	94,391	83	18,756	17	387	77	74	75	71	66
Eleventh	16,544	5,326	32	11,218	68	341	56	48	49	44	39
Twelfth	21,577	6,142	28	15,435	72	339	56	46	47	43	38
Adult Education	2,913	1,164	40	1,749	60	347	62	50	53	51	38
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	78,417	53,705	68	24,712	32	374	71	67	67	64	57
Female	75,659	53,280	70	22,379	30	375	73	66	69	64	60
Unknown	105	38	36	67	64	344	58	50	50	46	41
Race/Ethnicity											
American Indian or Alaska Native	1,444	949	66	495	34	368	70	63	65	61	54
Asian	13,145	11,892	90	1,253	10	407	82	83	84	80	78
Pacific Islander	1,060	735	69	325	31	371	69	66	67	61	58
Filipino	4,685	4,069	87	616	13	393	78	76	79	74	71
Hispanic or Latino	69,929	40,963	59	28,966	41	361	67	60	61	57	51
African American	14,029	7,009	50	7,020	50	354	63	56	56	51	47
White (not of Hispanic origin)	45,112	37,852	84	7,260	16	390	79	75	77	73	67
Two or More Races	4,777	3,554	74	1,223	26	382	75	71	72	68	62
Language Fluency											
English-Only Students	91,873	67,238	73	24,635	27	378	74	69	70	66	60
Initially Fluent English Proficient (IFEP)	8,742	7,253	83	1,489	17	393	79	76	77	74	69
Reclassified Fluent English Proficient (RFEP)	23,550	20,405	87	3,145	13	389	79	75	77	73	67
English-Learner Students	26,257	10,328	39	15,929	61	346	57	51	52	47	44
Unknown	3,759	1,799	48	1,960	52	353	64	55	57	54	43
Economically Disadvantaged											
No	60,930	51,396	84	9,534	16	392	79	76	78	73	68
Yes	81,677	48,614	60	33,063	40	363	67	61	62	57	52
Unknown	11,574	7,013	61	4,561	39	367	69	62	64	60	53
Special Education Program Participation											
Students Receiving Services	17,912	4,980	28	12,932	72	337	52	47	46	41	38
Students Not Receiving Services	136,269	102,043	75	34,226	25	380	74	69	71	67	61

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 8.E.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—February 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		157,272	90,331	57	27,522	17	39,419	25	66,941	43
Grade	Tenth	113,554	49,136	43	25,995	23	38,423	34	64,418	57
	Eleventh	17,695	16,553	94	669	4	473	3	1,142	6
	Twelfth	23,196	22,192	96	615	3	389	2	1,004	4
	Adult Education	2,827	2,450	87	243	9	134	5	377	13
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	83,309	52,015	62	14,030	17	17,264	21	31,294	38
	Female	73,872	38,238	52	13,485	18	22,149	30	35,634	48
	Unknown	91	78	86	7	8	6	7	13	14
Race/Ethnicity	American Indian or Alaskan Native	1,435	873	61	289	20	273	19	562	39
	Asian	14,920	6,527	44	2,508	17	5,885	39	8,393	56
	Pacific Islander	1,082	697	64	199	18	186	17	385	36
	Filipino	4,765	1,867	39	1,069	22	1,829	38	2,898	61
	Hispanic or Latino	71,958	52,383	73	10,392	14	9,183	13	19,575	27
	African American	13,178	9,763	74	1,847	14	1,568	12	3,415	26
	White (not of Hispanic origin)	45,246	16,151	36	10,283	23	18,812	42	29,095	64
	Two or More Races	4,688	2,070	44	935	20	1,683	36	2,618	56
Language Fluency	English-Only Students	89,840	42,844	48	18,081	20	28,915	32	46,996	52
	Initially Fluent English Proficient (IFEP)	8,510	3,157	37	1,876	22	3,477	41	5,353	63
	Reclassified Fluent English Proficient (RFEP)	22,773	10,179	45	6,186	27	6,408	28	12,594	55
	English-Learner Students	32,452	31,236	96	956	3	260	1	1,216	4
	Unknown	3,697	2,915	79	423	11	359	10	782	21
Economically Disadvantaged	No	60,597	21,036	35	13,326	22	26,235	43	39,561	65
	Yes	85,186	61,989	73	12,400	15	10,797	13	23,197	27
	Unknown	11,489	7,306	64	1,796	16	2,387	21	4,183	36
Special Education Program Participation	Receiving Services	20,198	18,662	92	954	5	582	3	1,536	8
	Not Receiving Services	137,074	71,669	52	26,568	19	38,837	28	65,405	48

[1] Results for groups with fewer than 11 students are not reported.

Table 8.E.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— February 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		154,181	88,172	57	42,885	28	23,124	15	66,009	43
Grade	Tenth	113,147	49,243	44	41,133	36	22,771	20	63,904	56
	Eleventh	16,544	15,563	94	792	5	189	1	981	6
	Twelfth	21,577	20,679	96	756	4	142	1	898	4
	Adult Education	2,913	2,687	92	204	7	22	1	226	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	78,417	44,783	57	21,498	27	12,136	15	33,634	43
	Female	75,659	43,296	57	21,376	28	10,987	15	32,363	43
	Unknown	105	93	89	11	10	1	1	12	11
Race/Ethnicity	American Indian or Alaskan Native	1,444	900	62	414	29	130	9	544	38
	Asian	13,145	3,184	24	4,404	34	5,557	42	9,961	76
	Pacific Islander	1,060	664	63	300	28	96	9	396	37
	Filipino	4,685	1,696	36	1,858	40	1,131	24	2,989	64
	Hispanic or Latino	69,929	50,599	72	14,908	21	4,422	6	19,330	28
	African American	14,029	11,136	79	2,340	17	553	4	2,893	21
	White (not of Hispanic origin)	45,112	17,669	39	17,168	38	10,275	23	27,443	61
	Two or More Races	4,777	2,324	49	1,493	31	960	20	2,453	51
Language Fluency	English-Only Students	91,873	48,112	52	28,440	31	15,321	17	43,761	48
	Initially Fluent English Proficient (IFEP)	8,742	3,455	40	2,832	32	2,455	28	5,287	60
	Reclassified Fluent English Proficient (RFEP)	23,550	9,998	42	8,877	38	4,675	20	13,552	58
	English-Learner Students	26,257	23,454	89	2,274	9	529	2	2,803	11
	Unknown	3,759	3,153	84	462	12	144	4	606	16
Economically Disadvantaged	No	60,930	23,022	38	22,393	37	15,515	25	37,908	62
	Yes	81,677	57,461	70	17,963	22	6,253	8	24,216	30
	Unknown	11,574	7,689	66	2,529	22	1,356	12	3,885	34
Special Education Program Participation	Receiving Services	17,912	16,381	91	1,245	7	286	2	1,531	9
	Not Receiving Services	136,269	71,791	53	41,640	31	22,838	17	64,478	47

[1] Results for groups with fewer than 11 students are not reported.

Table 8.E.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—February 2012

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	283	303	340	372	403	439	450	371	41	157,272
Grade[4]										
Tenth	293	318	362	387	410	445	450	384	36	113,554
Eleventh	275	291	316	336	354	387	423	337	29	17,695
Twelfth	275	289	314	332	348	377	410	332	27	23,196
Adult Education	285	305	326	342	362	399	428	346	29	2,827
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	279	299	334	365	396	433	450	365	41	83,309
Female	291	311	348	379	406	445	450	377	40	73,872
Unknown	275	281	320	338	358	406	410	340	33	91
Race/Ethnicity										
American Indian or Alaska Native	276	301	344	369	393	428	450	368	38	1,435
Asian	285	307	350	387	418	450	450	383	44	14,920
Pacific Islander	279	305	340	365	393	428	445	365	37	1,082
Filipino	297	320	362	390	414	445	450	387	37	4,765
Hispanic or Latino	281	299	332	356	385	418	445	357	37	71,958
African American	275	295	328	354	382	418	445	355	37	13,178
White (not of Hispanic origin)	289	316	367	393	418	450	450	389	38	45,246
Two or More Races	285	309	354	387	414	450	450	383	41	4,688
Language Fluency										
English-Only Students	285	307	352	382	410	445	450	379	40	89,840
Initially Fluent English Proficient (IFEP)	297	322	367	393	418	450	450	390	37	8,510
Reclassified Fluent English Proficient (RFEP)	309	338	365	385	403	433	450	385	29	22,773
English-Learner Students	275	293	316	334	352	377	396	334	26	32,452
Unknown	283	303	328	348	374	414	450	353	34	3,697
Economically Disadvantaged										
No	293	320	369	396	418	450	450	391	37	60,597
Yes	279	297	330	356	385	418	445	357	37	85,186
Unknown	279	301	334	362	396	433	450	365	41	11,489
Special Education Program Participation										
Students Receiving Services	275	285	307	328	350	390	423	331	32	20,198
Students Not Receiving Services	291	313	348	379	406	439	450	377	39	137,074

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Table 8.E.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—February 2012

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	298	315	343	370	405	450	450	375	40	154,181
Grade[4]										
Tenth	306	325	361	386	416	450	450	387	38	113,147
Eleventh	293	306	325	339	353	383	426	341	25	16,544
Twelfth	290	303	325	339	352	377	412	339	23	21,577
Adult Education	301	315	334	345	357	388	416	347	22	2,913
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	296	310	341	370	405	450	450	374	42	78,417
Female	303	319	345	370	405	450	450	375	39	75,659
Unknown	301	306	326	341	355	399	412	344	28	105
Race/Ethnicity										
American Indian or Alaska Native	296	310	339	366	393	438	450	368	38	1,444
Asian	312	334	381	412	446	450	450	407	38	13,145
Pacific Islander	301	317	343	368	396	438	450	371	36	1,060
Filipino	310	332	366	393	421	450	450	393	37	4,685
Hispanic or Latino	298	310	336	355	383	431	450	361	35	69,929
African American	293	306	330	348	372	416	450	354	34	14,029
White (not of Hispanic origin)	303	323	363	391	421	450	450	390	39	45,112
Two or More Races	301	317	348	381	416	450	450	382	42	4,777
Language Fluency										
English-Only Students	298	315	346	377	409	450	450	378	41	91,873
Initially Fluent English Proficient (IFEP)	306	326	361	393	426	450	450	393	41	8,742
Reclassified Fluent English Proficient (RFEP)	317	334	363	386	416	450	450	389	35	23,550
English-Learner Students	296	306	326	343	359	399	438	346	28	26,257
Unknown	298	315	334	348	366	416	450	353	30	3,759
Economically Disadvantaged										
No	306	325	363	393	426	450	450	392	39	60,930
Yes	296	310	336	357	386	438	450	363	37	81,677
Unknown	298	312	337	359	393	446	450	367	40	11,574
Special Education Program Participation										
Students Receiving Services	287	298	317	332	352	393	431	337	29	17,912
Students Not Receiving Services	306	321	348	377	409	450	450	380	39	136,269

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Appendix 8.F: Frequency Distributions and Demographic Summaries—March 2012

Table 8.F.1: Frequency Distributions, ELA—March 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	12,321	3	12,321	97
440-449	6,634	2	18,955	95
430-439	17,291	4	36,246	91
420-429	21,600	6	57,846	85
410-419	24,517	6	82,363	79
400-409	38,878	10	121,241	69
390-399	38,878	10	160,119	59
380-389	35,307	9	195,426	50
370-379	40,633	10	236,059	39
360-369	32,708	8	268,767	31
350-359[1]	32,185	8	300,952	22
340-349	20,296	5	321,248	17
330-339	19,484	5	340,732	12
320-329	14,897	4	355,629	8
310-319	11,117	3	366,746	6
300-309	8,468	2	375,214	3
290-299	6,427	2	381,641	2
280-289	3,821	1	385,462	1
270-279	2,642	1	388,104	0

[1] Passing Score = 350

Table 8.F.2: Frequency Distributions, Mathematics—March 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	26,965	7	26,965	93
440-449	20,015	5	46,980	88
430-439	10,041	3	57,021	85
420-429	20,063	5	77,084	80
410-419	29,883	8	106,967	72
400-409	28,511	7	135,478	65
390-399	36,138	9	171,616	55
380-389	33,616	9	205,232	46
370-379	36,969	10	242,201	37
360-369	32,189	8	274,390	28
350-359[1]	27,803	7	302,193	21
340-349	29,090	8	331,283	14
330-339	19,067	5	350,350	9
320-329	17,547	5	367,897	4
310-319	8,491	2	376,388	2
300-309	4,908	1	381,296	1
290-299	1,548	0	382,844	0
280-289	286	0	383,130	0
270-279	266	0	383,396	0

[1] Passing Score = 350

Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	12,321	3	12,321	97
440-449	6,634	2	18,955	95
430-439	17,291	4	36,246	91
420-429	21,600	6	57,846	85
410-419	24,517	6	82,363	79
403-409[1]	25,735	7	108,098	72
390-402	52,021	13	160,119	59
380-389[2]	35,307	9	195,426	50
370-379	40,633	10	236,059	39
360-369	32,708	8	268,767	31
350-359	32,185	8	300,952	22
340-349	20,296	5	321,248	17
330-339	19,484	5	340,732	12
320-329	14,897	4	355,629	8
310-319	11,117	3	366,746	6
300-309	8,468	2	375,214	3
290-299	6,427	2	381,641	2
280-289	3,821	1	385,462	1
270-279	2,642	1	388,104	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	26,965	7	26,965	93
440-449	20,015	5	46,980	88
430-439	10,041	3	57,021	85
422-429[1]	20,063	5	77,084	80
410-421	29,883	8	106,967	72
400-409	28,511	7	135,478	65
390-399	36,138	9	171,616	55
380-389[2]	33,616	9	205,232	46
370-379	36,969	10	242,201	37
360-369	32,189	8	274,390	28
350-359	27,803	7	302,193	21
340-349	29,090	8	331,283	14
330-339	19,067	5	350,350	9
320-329	17,547	5	367,897	4
310-319	8,491	2	376,388	2
300-309	4,908	1	381,296	1
290-299	1,548	0	382,844	0
280-289	286	0	383,130	0
270-279	266	0	383,396	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.F.5: Demographic Summary for All Examinees, ELA—March 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	388,104	300,952	78	87,152	22	378	79	75	80	69	75	2.4
Grade												
Tenth	344,852	287,591	83	57,261	17	384	82	77	83	72	77	2.4
Eleventh	24,881	8,416	34	16,465	66	338	62	55	63	46	54	2.0
Twelfth	15,757	3,789	24	11,968	76	330	57	51	57	43	51	1.9
Adult Education	2,614	1,156	44	1,458	56	348	69	63	67	54	57	2.0
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	200,455	146,657	73	53,798	27	373	78	72	78	66	72	2.3
Female	187,505	154,216	82	33,289	18	384	81	77	83	72	78	2.5
Unknown	144	79	55	65	45	354	72	64	71	54	62	2.0
Race/Ethnicity												
American Indian or Alaska Native	2,504	1,894	76	610	24	375	81	73	80	66	72	2.3
Asian	33,137	29,072	88	4,065	12	400	84	82	85	79	84	2.7
Pacific Islander	2,093	1,619	77	474	23	377	79	74	80	68	76	2.4
Filipino	10,107	9,033	89	1,074	11	395	83	81	85	77	83	2.6
Hispanic or Latino	214,603	153,358	71	61,245	29	369	75	71	77	65	71	2.3
African American	25,955	17,340	67	8,615	33	365	76	68	75	61	68	2.2
White (not of Hispanic origin)	92,245	82,510	89	9,735	11	395	88	82	87	76	82	2.6
Two or More Races	7,460	6,126	82	1,334	18	386	83	77	83	71	77	2.5
Language Fluency												
English-Only Students	198,191	165,174	83	33,017	17	386	84	78	83	72	78	2.5
Initially Fluent English Proficient (IFEP)	26,589	24,310	91	2,279	9	397	86	82	87	78	83	2.6
Reclassified Fluent English Proficient (RFEP)	87,289	81,279	93	6,010	7	388	83	80	85	76	81	2.5
English-Learner Students	71,643	27,845	39	43,798	61	340	61	57	64	49	56	2.0
Unknown	4,392	2,344	53	2,048	47	355	71	65	70	57	62	2.1
Economically Disadvantaged												
No	133,657	120,699	90	12,958	10	397	87	82	87	77	82	2.6
Yes	226,129	158,506	70	67,623	30	367	75	70	77	64	70	2.3
Unknown	28,318	21,747	77	6,571	23	380	80	75	81	69	75	2.4
Special Education Program Participation												
Students Receiving Services	39,782	12,914	32	26,868	68	335	60	51	61	45	52	1.9
Students Not Receiving Services	348,322	288,038	83	60,284	17	384	82	77	83	72	77	2.4

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.F.6: Demographic Summary for All Examinees, Mathematics—March 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics[2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	383,396	302,193	79	81,203	21	384	74	70	74	69	64
Grade											
Tenth	343,298	289,296	84	54,002	16	389	76	73	76	72	67
Eleventh	22,637	7,871	35	14,766	65	344	54	49	52	46	39
Twelfth	14,729	3,949	27	10,780	73	339	52	47	48	43	36
Adult Education	2,732	1,077	39	1,655	61	349	59	51	54	52	39
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	194,182	151,306	78	42,876	22	384	73	71	74	69	64
Female	189,062	150,809	80	38,253	20	384	75	70	74	70	65
Unknown	152	78	51	74	49	355	61	56	59	52	44
Race/Ethnicity											
American Indian or Alaska Native	2,527	1,917	76	610	24	377	72	68	71	65	59
Asian	31,412	29,824	95	1,588	5	418	85	85	87	87	84
Pacific Islander	2,079	1,663	80	416	20	382	73	70	73	69	64
Filipino	9,936	9,075	91	861	9	402	80	78	82	80	76
Hispanic or Latino	211,879	155,038	73	56,841	27	375	70	66	69	64	59
African American	26,509	16,597	63	9,912	37	365	65	61	64	58	53
White (not of Hispanic origin)	91,605	82,070	90	9,535	10	398	81	77	81	77	72
Two or More Races	7,449	6,009	81	1,440	19	388	76	72	75	71	65
Language Fluency											
English-Only Students	198,955	162,622	82	36,333	18	388	76	72	76	71	66
Initially Fluent English Proficient (IFEP)	26,824	24,100	90	2,724	10	401	82	78	81	78	73
Reclassified Fluent English Proficient (RFEP)	88,335	80,913	92	7,422	8	395	80	76	79	76	71
English-Learner Students	64,805	32,283	50	32,522	50	353	57	55	58	52	48
Unknown	4,477	2,275	51	2,202	49	357	62	56	59	55	45
Economically Disadvantaged											
No	133,446	119,957	90	13,489	10	400	81	77	81	78	73
Yes	221,638	160,700	73	60,938	27	375	70	66	69	64	59
Unknown	28,312	21,536	76	6,776	24	384	74	70	73	69	63
Special Education Program Participation											
Students Receiving Services	35,864	12,552	35	23,312	65	344	52	49	51	45	41
Students Not Receiving Services	347,532	289,641	83	57,891	17	388	76	72	76	72	67

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1— Algebra 1

Table 8.F.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—March 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		388,104	192,678	50	87,328	23	108,098	28	195,426	50
Grade	Tenth	344,852	152,239	44	85,485	25	107,128	31	192,613	56
	Eleventh	24,881	23,161	93	1,133	5	587	2	1,720	7
	Twelfth	15,757	15,074	96	445	3	238	2	683	4
	Adult Education	2,614	2,204	84	265	10	145	6	410	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	200,455	110,720	55	43,113	22	46,622	23	89,735	45
	Female	187,505	81,849	44	44,196	24	61,460	33	105,656	56
	Unknown	144	109	76	19	13	16	11	35	24
Race/Ethnicity	American Indian or Alaskan Native	2,504	1,289	51	596	24	619	25	1,215	49
	Asian	33,137	9,091	27	6,473	20	17,573	53	24,046	73
	Pacific Islander	2,093	1,077	51	500	24	516	25	1,016	49
	Filipino	10,107	3,060	30	2,554	25	4,493	44	7,047	70
	Hispanic or Latino	214,603	131,000	61	46,943	22	36,660	17	83,603	39
	African American	25,955	16,652	64	5,106	20	4,197	16	9,303	36
	White (not of Hispanic origin)	92,245	27,466	30	23,511	25	41,268	45	64,779	70
	Two or More Races	7,460	3,043	41	1,645	22	2,772	37	4,417	59
Language Fluency	English-Only Students	198,191	80,346	41	48,093	24	69,752	35	117,845	59
	Initially Fluent English Proficient (IFEP)	26,589	7,824	29	6,759	25	12,006	45	18,765	71
	Reclassified Fluent English Proficient (RFEP)	87,289	34,626	40	27,757	32	24,906	29	52,663	60
	English-Learner Students	71,643	66,569	93	4,119	6	955	1	5,074	7
	Unknown	4,392	3,313	75	600	14	479	11	1,079	25
Economically Disadvantaged	No	133,657	38,543	29	33,453	25	61,661	46	95,114	71
	Yes	226,129	140,912	62	47,742	21	37,475	17	85,217	38
	Unknown	28,318	13,223	47	6,133	22	8,962	32	15,095	53
Special Education Program	Receiving Services	39,782	35,974	90	2,555	6	1,253	3	3,808	10
	Not Receiving Services	348,322	156,704	45	84,773	24	106,845	31	191,618	55

[1] Results for groups with fewer than 11 students are not reported.

Table 8.F.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—March 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		383,396	178,164	46	128,148	33	77,084	20	205,232	54
Grade	Tenth	343,298	140,493	41	126,136	37	76,669	22	202,805	59
	Eleventh	22,637	21,110	93	1,245	5	282	1	1,527	7
	Twelfth	14,729	14,083	96	551	4	95	1	646	4
	Adult Education	2,732	2,478	91	216	8	38	1	254	9
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	194,182	88,827	46	63,589	33	41,766	22	105,355	54
	Female	189,062	89,217	47	64,536	34	35,309	19	99,845	53
	Unknown	152	120	79	23	15	9	6	32	21
Race/Ethnicity	American Indian or Alaskan Native	2,527	1,328	53	877	35	322	13	1,199	47
	Asian	31,412	4,645	15	9,188	29	17,579	56	26,767	85
	Pacific Islander	2,079	996	48	743	36	340	16	1,083	52
	Filipino	9,936	2,554	26	4,050	41	3,332	34	7,382	74
	Hispanic or Latino	211,879	120,640	57	66,861	32	24,378	12	91,239	43
	African American	26,509	17,688	67	6,778	26	2,043	8	8,821	33
	White (not of Hispanic origin)	91,605	27,179	30	37,161	41	27,265	30	64,426	70
	Two or More Races	7,449	3,134	42	2,490	33	1,825	24	4,315	58
Language Fluency	English-Only Students	198,955	82,799	42	71,436	36	44,720	22	116,156	58
	Initially Fluent English Proficient (IFEP)	26,824	7,833	29	9,410	35	9,581	36	18,991	71
	Reclassified Fluent English Proficient (RFEP)	88,335	30,418	34	37,163	42	20,754	23	57,917	66
	English-Learner Students	64,805	53,583	83	9,426	15	1,796	3	11,222	17
	Unknown	4,477	3,531	79	713	16	233	5	946	21
Economically Disadvantaged	No	133,446	38,650	29	51,148	38	43,648	33	94,796	71
	Yes	221,638	126,203	57	68,248	31	27,187	12	95,435	43
	Unknown	28,312	13,311	47	8,752	31	6,249	22	15,001	53
Special Education Program Participation	Receiving Services	35,864	31,263	87	3,635	10	966	3	4,601	13
	Not Receiving Services	347,532	146,901	42	124,513	36	76,118	22	200,631	58

[1] Results for groups with fewer than 11 students are not reported.

Table 8.F.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—March 2012

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	285	307	352	382	405	439	450	378	39	388,104
Grade[4]										
Tenth	289	317	361	385	409	446	450	384	37	344,852
Eleventh	275	289	319	337	355	388	423	338	30	24,881
Twelfth	275	285	311	331	348	377	413	330	29	15,757
Adult Education	281	301	327	344	366	405	433	348	31	2,614
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	279	301	346	377	401	439	450	373	40	200,455
Female	291	317	359	388	413	446	450	384	38	187,505
Unknown	275	287	329	355	379	418	439	354	39	144
Race/Ethnicity										
American Indian or Alaska Native	279	307	350	379	401	439	450	375	39	2,504
Asian	293	321	377	405	433	450	450	400	40	33,137
Pacific Islander	287	315	352	379	401	439	450	377	36	2,093
Filipino	297	331	374	398	423	450	450	395	35	10,107
Hispanic or Latino	283	303	346	371	394	428	450	369	36	214,603
African American	275	295	339	366	391	428	450	365	39	25,955
White (not of Hispanic origin)	293	327	377	398	423	450	450	395	36	92,245
Two or More Races	279	307	361	391	418	450	450	386	41	7,460
Language Fluency										
English-Only Students	285	313	364	391	413	446	450	386	38	198,191
Initially Fluent English Proficient (IFEP)	303	335	377	398	423	450	450	397	34	26,589
Reclassified Fluent English Proficient (RFEP)	317	344	371	388	405	439	450	388	28	87,289
English-Learner Students	275	291	321	342	359	385	405	340	29	71,643
Unknown	275	297	329	352	379	423	446	355	37	4,392
Economically Disadvantaged										
No	297	331	377	401	423	450	450	397	35	133,657
Yes	281	301	344	369	394	428	450	367	37	226,129
Unknown	281	305	352	385	409	446	450	380	41	28,318
Special Education Program Participation										
Students Receiving Services	275	283	309	333	357	394	428	335	34	39,782
Students Not Receiving Services	293	319	359	385	409	446	450	384	36	348,322

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Table 8.F.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—March 2012

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	304	322	355	383	414	450	450	384	39	383,396
Grade[4]										
Tenth	309	327	362	387	418	450	450	389	37	343,298
Eleventh	294	309	327	342	357	385	428	344	24	22,637
Twelfth	294	307	324	339	351	376	410	339	22	14,729
Adult Education	304	318	335	346	358	395	428	349	23	2,732
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	302	318	355	385	414	450	450	384	40	194,182
Female	309	326	357	383	410	450	450	384	37	189,062
Unknown	296	307	330	351	372	423	448	355	34	152
Race/Ethnicity										
American Indian or Alaska Native	302	318	351	376	401	448	450	377	36	2,527
Asian	322	349	398	428	450	450	450	418	34	31,412
Pacific Islander	304	324	357	383	407	448	450	382	36	2,079
Filipino	318	339	378	404	428	450	450	402	35	9,936
Hispanic or Latino	304	320	348	372	398	440	450	375	36	211,879
African American	296	311	339	362	390	433	450	365	36	26,509
White (not of Hispanic origin)	311	335	374	401	428	450	450	398	36	91,605
Two or More Races	302	320	358	387	418	450	450	388	40	7,449
Language Fluency										
English-Only Students	304	322	360	387	418	450	450	388	39	198,955
Initially Fluent English Proficient (IFEP)	313	337	374	404	433	450	450	401	37	26,824
Reclassified Fluent English Proficient (RFEP)	324	342	370	392	418	450	450	395	33	88,335
English-Learner Students	299	311	333	349	370	407	440	353	29	64,805
Unknown	299	313	337	351	372	423	450	357	32	4,477
Economically Disadvantaged										
No	313	335	374	401	433	450	450	400	36	133,446
Yes	302	318	348	372	401	448	450	375	36	221,638
Unknown	304	320	351	383	414	450	450	384	41	28,312
Special Education Program Participation										
Students Receiving Services	291	304	322	339	360	404	448	344	31	35,864
Students Not Receiving Services	311	329	360	387	414	450	450	388	37	347,532

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Appendix 8.G: Frequency Distributions and Demographic Summaries—May 2012

Table 8.G.1: Frequency Distributions, ELA—May 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	219	1	219	99
440-449	120	0	339	99
430-439	129	0	468	99
420-429	402	1	870	98
410-419	742	2	1,612	96
400-409	588	2	2,200	94
390-399	926	2	3,126	92
380-389	1,447	4	4,573	88
370-379	1,568	4	6,141	84
360-369	1,900	5	8,041	79
350-359[1]	3,328	9	11,369	71
340-349	4,591	12	15,960	59
330-339	6,240	16	22,200	43
320-329	4,749	12	26,949	31
310-319	4,602	12	31,551	19
300-309	3,165	8	34,716	11
290-299	2,295	6	37,011	6
280-289	1,228	3	38,239	2
275-279	927	2	39,166	0

[1] Passing Score = 350

Table 8.G.2: Frequency Distributions, Mathematics—May 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	365	1	365	99
440-449	118	0	483	99
430-439	310	1	793	98
420-429	338	1	1,131	97
410-419	388	1	1,519	96
400-409	618	2	2,137	94
390-399	899	2	3,036	92
380-389	1,054	3	4,090	89
370-379	1,677	5	5,767	84
360-369	2,360	6	8,127	78
350-359[1]	4,426	12	12,553	65
340-349	5,848	16	18,401	49
330-339	5,195	14	23,596	35
320-329	5,761	16	29,357	19
310-319	3,469	10	32,826	10
300-309	2,669	7	35,495	2
290-299	542	1	36,037	1
280-289	199	1	36,236	0
270-279	132	0	36,368	0

[1] Passing Score = 350

Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	219	1	219	99
440-449	120	0	339	99
430-439	129	0	468	99
420-429	402	1	870	98
410-419	742	2	1,612	96
403-409	289	1	1,901	95
390-402	1,225	3	3,126	92
380-389	1,447	4	4,573	88
370-379	1,568	4	6,141	84
360-369	1,900	5	8,041	79
350-359	3,328	9	11,369	71
340-349	4,591	12	15,960	59
330-339	6,240	16	22,200	43
320-329	4,749	12	26,949	31
310-319	4,602	12	31,551	19
300-309	3,165	8	34,716	11
290-299	2,295	6	37,011	6
280-289	1,228	3	38,239	2
270-279	927	2	39,166	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2012

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	365	1	365	99
440-449	118	0	483	99
430-439	310	1	793	98
422-429	162	0	955	97
410-421	564	2	1,519	96
400-409	618	2	2,137	94
390-399	899	2	3,036	92
380-389	1,054	3	4,090	89
370-379	1,677	5	5,767	84
360-369	2,360	6	8,127	78
350-359	4,426	12	12,553	65
340-349	5,848	16	18,401	49
330-339	5,195	14	23,596	35
320-329	5,761	16	29,357	19
310-319	3,469	10	32,826	10
300-309	2,669	7	35,495	2
290-299	542	1	36,037	1
280-289	199	1	36,236	0
270-279	132	0	36,368	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.G.5: Demographic Summary for All Examinees, ELA—May 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Reading [2] Avg. Percent Correct			Writing [2] Avg. Percent Correct		Writing Applications Mean Score
							RW	RC	RL	WS	WC	Essay
Total Examinees	39,166	11,369	29	27,797	71	337	62	54	57	49	52	2.1
Grade												
Tenth	9,318	5,328	57	3,990	43	358	73	65	69	59	63	2.3
Eleventh	13,350	3,306	25	10,044	75	334	60	53	56	46	50	2.0
Twelfth	14,062	1,861	13	12,201	87	326	55	48	50	42	45	2.0
Adult Education	2,431	872	36	1,559	64	345	67	63	61	56	55	2.1
Unknown	5	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	22,430	5,898	26	16,532	74	334	60	52	56	47	49	2.0
Female	16,656	5,451	33	11,205	67	342	64	57	60	51	55	2.2
Unknown	80	20	25	60	75	334	62	54	57	47	46	2.0
Race/Ethnicity												
American Indian or Alaska Native	340	116	34	224	66	339	63	55	60	49	53	2.0
Asian	1,823	448	25	1,375	75	337	57	55	55	48	54	2.1
Pacific Islander	292	89	30	203	70	338	58	54	60	48	53	2.1
Filipino	452	153	34	299	66	346	64	60	61	55	56	2.3
Hispanic or Latino	24,422	5,982	24	18,440	76	334	60	53	55	47	50	2.1
African American	4,029	948	24	3,081	76	330	59	50	55	43	46	2.0
White (not of Hispanic origin)	6,653	3,154	47	3,499	53	353	70	62	66	56	60	2.3
Two or More Races	1,155	479	41	676	59	346	66	59	63	52	56	2.2
Language Fluency												
English-Only Students	17,286	6,572	38	10,714	62	343	66	57	61	51	55	2.2
Initially Fluent English Proficient (IFEP)	1,215	539	44	676	56	349	71	61	64	54	58	2.3
Reclassified Fluent English Proficient (RFEP)	2,325	1,308	56	1,017	44	356	73	66	68	59	63	2.4
English-Learner Students	15,592	1,944	12	13,648	88	326	54	48	50	43	46	1.9
Unknown	2,748	1,006	37	1,742	63	344	67	62	61	55	55	2.1
Economically Disadvantaged												
No	7,390	3,377	46	4,013	54	352	69	62	65	56	60	2.3
Yes	26,766	6,201	23	20,565	77	332	59	52	55	46	49	2.0
Unknown	5,010	1,791	36	3,219	64	343	66	59	60	53	54	2.1
Special Education Program Participation												
Students Receiving Services	9,025	904	10	8,121	90	318	50	42	47	37	39	1.8
Students Not Receiving Services	30,141	10,465	35	19,676	65	343	65	58	61	52	55	2.2

[1] Results for groups with fewer than 11 students are not reported.

[2] RW — Word Analysis, RC — Reading Comprehension, RL — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

Table 8.G.6: Demographic Summary for All Examinees, Mathematics—May 2012

	N Tested[1]	N Pass	Percent Pass	N Not Pass	Percent Not Pass	Mean Scale Score	Strands for Mathematics [2] Average Percent Correct				
							PS	NS	AF	MG	A1
Total Examinees	36,368	12,553	35	23,815	65	344	56	50	51	47	36
Grade											
Tenth	9,016	5,283	59	3,733	41	362	65	62	62	57	46
Eleventh	12,217	3,698	30	8,519	70	340	53	48	50	44	34
Twelfth	12,428	2,532	20	9,896	80	334	50	43	45	41	30
Adult Education	2,702	1,037	38	1,665	62	346	59	49	54	52	33
Unknown	5	-	-	-	-	-	-	-	-	-	-
Gender											
Male	18,220	6,159	34	12,061	66	343	54	51	50	47	34
Female	18,059	6,362	35	11,697	65	345	57	49	53	47	37
Unknown	89	32	36	57	64	343	56	50	50	49	34
Race/Ethnicity											
American Indian or Alaska Native	361	140	39	221	61	344	56	50	51	48	34
Asian	884	447	51	437	49	363	60	62	61	57	49
Pacific Islander	249	86	35	163	65	346	55	52	53	47	36
Filipino	418	198	47	220	53	355	59	56	59	54	43
Hispanic or Latino	22,506	6,905	31	15,601	69	341	54	47	50	45	34
African American	4,496	1,089	24	3,407	76	335	51	45	45	41	31
White (not of Hispanic origin)	6,306	3,203	51	3,103	49	358	63	59	59	55	42
Two or More Races	1,148	485	42	663	58	349	59	54	54	50	37
Language Fluency											
English-Only Students	18,280	6,958	38	11,322	62	346	57	52	53	48	36
Initially Fluent English Proficient (IFEP)	1,318	608	46	710	54	353	62	56	57	52	40
Reclassified Fluent English Proficient (RFEP)	2,902	1,567	54	1,335	46	357	66	58	60	54	43
English-Learner Students	10,956	2,289	21	8,667	79	335	48	44	45	41	32
Unknown	2,912	1,131	39	1,781	61	346	58	50	54	51	34
Economically Disadvantaged											
No	7,345	3,463	47	3,882	53	355	62	57	58	53	42
Yes	23,863	7,082	30	16,781	70	340	53	47	49	44	34
Unknown	5,160	2,008	39	3,152	61	347	58	51	54	50	35
Special Education Program Participation											
Students Receiving Services	7,365	907	12	6,458	88	326	42	40	38	35	28
Students Not Receiving Services	29,003	11,646	40	17,357	60	349	59	52	55	50	37

[1] Results for groups with fewer than 11 students are not reported.

[2] PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, A1 — Algebra 1

Table 8.G.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—May 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		39,166	34,593	88	2,672	7	1,901	5	4,573	12
Grade	Tenth	9,318	6,292	68	1,668	18	1,358	15	3,026	32
	Eleventh	13,350	12,466	93	557	4	327	2	884	7
	Twelfth	14,062	13,689	97	255	2	118	1	373	3
	Adult Education	2,431	2,143	88	190	8	98	4	288	12
	Unknown	5	-	-	-	-	-	-	-	-
Gender	Male	22,430	20,254	90	1,322	6	854	4	2,176	10
	Female	16,656	14,269	86	1,343	8	1,044	6	2,387	14
	Unknown	80	70	88	7	9	3	4	10	13
Race/Ethnicity	American Indian or Alaskan Native	340	285	84	39	11	16	5	55	16
	Asian	1,823	1,600	88	93	5	130	7	223	12
	Pacific Islander	292	261	89	19	7	12	4	31	11
	Filipino	452	375	83	38	8	39	9	77	17
	Hispanic or Latino	24,422	22,607	93	1,268	5	547	2	1,815	7
	African American	4,029	3,744	93	199	5	86	2	285	7
	White (not of Hispanic origin)	6,653	4,802	72	894	13	957	14	1,851	28
	Two or More Races	1,155	919	80	122	11	114	10	236	20
Language Fluency	English-Only Students	17,286	14,101	82	1,737	10	1,448	8	3,185	18
	Initially Fluent English Proficient (IFEP)	1,215	961	79	150	12	104	9	254	21
	Reclassified Fluent English Proficient (RFEP)	2,325	1,763	76	376	16	186	8	562	24
	English-Learner Students	15,592	15,382	99	182	1	28	0	210	1
	Unknown	2,748	2,386	87	227	8	135	5	362	13
Economically Disadvantaged	No	7,390	5,435	74	959	13	996	13	1,955	26
	Yes	26,766	24,901	93	1,292	5	573	2	1,865	7
	Unknown	5,010	4,257	85	421	8	332	7	753	15
Special Education Program	Receiving Services	9,025	8,857	98	120	1	48	1	168	2
	Not Receiving Services	30,141	25,736	85	2,552	8	1,853	6	4,405	15

[1] Results for groups with fewer than 11 students are not reported.

Table 8.G.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— May 2012

Subgroup	Group	N Tested[1]	N Below Proficient	Percent Below Proficient	N Proficient	Percent Proficient	N Advanced	Percent Advanced	N Above Proficient	Percent Above Proficient
Total Examinees		36,368	32,278	89	3,135	9	955	3	4,090	11
Grade	Tenth	9,016	6,257	69	2,028	22	731	8	2,759	31
	Eleventh	12,217	11,456	94	616	5	145	1	761	6
	Twelfth	12,428	12,082	97	285	2	61	0	346	3
	Adult Education	2,702	2,480	92	204	8	18	1	222	8
	Unknown	5	-	-	-	-	-	-	-	-
Gender	Male	18,220	16,057	88	1,607	9	556	3	2,163	12
	Female	18,059	16,139	89	1,521	8	399	2	1,920	11
	Unknown	89	82	92	7	8			7	8
Race/Ethnicity	American Indian or Alaskan Native	361	319	88	32	9	10	3	42	12
	Asian	884	626	71	133	15	125	14	258	29
	Pacific Islander	249	218	88	22	9	9	4	31	12
	Filipino	418	336	80	52	12	30	7	82	20
	Hispanic or Latino	22,506	20,813	92	1,423	6	270	1	1,693	8
	African American	4,496	4,276	95	193	4	27	1	220	5
	White (not of Hispanic origin)	6,306	4,713	75	1,155	18	438	7	1,593	25
	Two or More Races	1,148	977	85	125	11	46	4	171	15
Language Fluency	English-Only Students	18,280	15,625	85	1,993	11	662	4	2,655	15
	Initially Fluent English Proficient (IFEP)	1,318	1,081	82	170	13	67	5	237	18
	Reclassified Fluent English Proficient (RFEP)	2,902	2,301	79	466	16	135	5	601	21
	English-Learner Students	10,956	10,641	97	266	2	49	0	315	3
	Unknown	2,912	2,630	90	240	8	42	1	282	10
Economically Disadvantaged	No	7,345	5,651	77	1,178	16	516	7	1,694	23
	Yes	23,863	22,130	93	1,449	6	284	1	1,733	7
	Unknown	5,160	4,497	87	508	10	155	3	663	13
Special Education Program	Receiving Services	7,365	7,224	98	123	2	18	0	141	2
	Not Receiving Services	29,003	25,054	86	3,012	10	937	3	3,949	14

[1] Results for groups with fewer than 11 students are not reported.

Table 8.G.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—May 2012

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	289	313	334	355	402	435	337	34	39,166
Grade[4]										
Tenth	275	291	326	359	389	429	450	358	42	9,318
Eleventh	275	289	313	334	349	386	424	334	29	13,350
Twelfth	275	287	308	326	339	366	402	326	25	14,062
Adult Education	287	306	326	341	359	399	424	345	28	2,431
Unknown	-	-	-	-	-	-	-	-	-	5
Gender										
Male	275	285	310	330	351	399	429	334	33	22,430
Female	279	294	321	337	359	410	442	342	34	16,656
Unknown	275	281	310	334	350	402	442	334	36	80
Race/Ethnicity										
American Indian or Alaska Native	275	285	313	336	364	402	435	339	36	340
Asian	277	291	313	330	349	419	450	337	36	1,823
Pacific Islander	275	289	315	336	353	402	424	338	32	292
Filipino	287	302	323	339	362	419	442	346	34	452
Hispanic or Latino	275	289	313	332	349	389	419	334	29	24,422
African American	275	281	306	328	349	386	419	330	32	4,029
White (not of Hispanic origin)	275	291	321	347	383	429	450	353	42	6,653
Two or More Races	275	285	317	341	371	419	450	346	40	1,155
Language Fluency										
English-Only Students	275	287	315	339	366	419	442	343	39	17,286
Initially Fluent English Proficient (IFEP)	275	294	324	345	373	414	450	349	37	1,215
Reclassified Fluent English Proficient (RFEP)	277	300	334	355	378	414	442	356	34	2,325
English-Learner Students	275	289	310	326	341	362	383	326	23	15,592
Unknown	277	298	324	341	362	402	429	344	31	2,748
Economically Disadvantaged										
No	275	293	323	345	381	429	450	352	41	7,390
Yes	275	287	311	330	349	386	419	332	30	26,766
Unknown	275	291	319	339	362	410	442	343	35	5,010
Special Education Program Participation										
Students Receiving Services	275	279	298	315	334	362	389	318	25	9,025
Students Not Receiving Services	275	294	321	339	362	410	442	343	34	30,141

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Table 8.G.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—May 2012

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	292	305	324	340	358	405	450	344	30	36,368
Grade[4]										
Tenth	289	307	331	359	388	437	450	362	39	9,016
Eleventh	292	305	324	338	354	386	425	340	26	12,217
Twelfth	292	302	318	333	347	369	402	334	22	12,428
Adult Education	302	314	331	344	359	391	416	346	23	2,702
Unknown	-	-	-	-	-	-	-	-	-	5
Gender										
Male	285	302	320	338	358	408	450	343	32	18,220
Female	294	307	326	340	358	402	445	345	29	18,059
Unknown	275	305	324	342	359	391	402	343	26	89
Race/Ethnicity										
American Indian or Alaska Native	282	302	324	340	359	405	450	344	32	361
Asian	297	309	329	351	391	450	450	363	44	884
Pacific Islander	292	305	328	342	358	408	450	346	31	249
Filipino	302	307	331	349	371	431	450	355	35	418
Hispanic or Latino	292	305	322	338	354	391	425	341	26	22,506
African American	285	300	318	333	349	379	412	335	25	4,496
White (not of Hispanic origin)	294	307	329	351	381	431	450	358	37	6,306
Two or More Races	289	305	326	344	365	416	450	349	33	1,148
Language Fluency										
English-Only Students	289	302	324	342	361	412	450	346	33	18,280
Initially Fluent English Proficient (IFEP)	297	309	331	347	369	425	450	353	34	1,318
Reclassified Fluent English Proficient (RFEP)	300	314	337	352	373	420	450	357	32	2,902
English-Learner Students	292	302	320	333	347	371	402	335	22	10,956
Unknown	294	309	329	344	359	396	431	346	26	2,912
Economically Disadvantaged										
No	294	307	329	347	377	431	450	355	37	7,345
Yes	289	302	322	337	352	388	425	340	27	23,863
Unknown	294	307	328	342	361	408	450	347	31	5,160
Special Education Program Participation										
Students Receiving Services	282	297	312	322	338	363	393	326	22	7,365
Students Not Receiving Services	294	309	328	344	361	412	450	349	31	29,003

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with fewer than 11 students are not reported.

[4] Grade ten students can only take the CAHSEE one time in the spring during the February, March, or May administration.

Chapter 9: Quality Control Procedures

ETS implements rigorous quality control procedures throughout the test development, administration, scoring, and reporting processes. As part of this effort, ETS maintains the Office of Professional Standards, which resides in the legal department. The office publishes and maintains the *ETS Standards for Quality and Fairness*, with the purposes of helping design, develop, and deliver technically sound, fair, and useful products and services, and to help the public and auditors evaluate those products and services.

In addition, every department that is involved in the testing cycle designs and implements an independent set of procedures to ensure the quality of its products. ETS established an Office of Quality Assurance, which provides tools and oversight to assist program managers in this endeavor. In the next sections, these procedures are described.

Quality Control of Item Development

The item development process for the CAHSEE is described in detail in Chapter 3. The following sections highlight elements of the process devoted specifically to quality control of item development.

Item Specifications

ETS maintains item development specifications for the CAHSEE and has developed an item development plan to guide the writing of the items for both content areas. Item writing emphasis is determined in consultation with the CDE. Adherence to these specifications ensures the maintenance of quality and consistency of the item development process.

Item Writers

The items for the CAHSEE are written by panels of item writers that have a thorough understanding of the California content standards. The item writers are carefully screened and selected by senior content staff. Only those with strong content and teaching backgrounds are invited to participate in an extensive training program for item writers.

Internal Contractor Reviews

Once items have been written, ETS assessment specialists make sure that each item goes through an internal review process. Every step of this process is designed to produce items that exceed industry standards for quality. It includes three rounds of content reviews, an editorial review, an internal bias and sensitivity review, and a high-level review and approval by a content area director. A carefully designed and monitored workflow and detailed checklists help to ensure that all items meet the specifications at each step of the process.

Content Review

ETS assessment specialists make sure that the test items and related materials comply with ETS written guidelines for clarity, style, accuracy, and appropriateness and with approved item specifications. The artwork and graphics for the items are created during the

internal content review period so assessment specialists can evaluate the correctness and appropriateness of the art early in the item development process. ETS selects visual stimuli that are relevant to the item content and that are easily understood.

Editorial Review

Another step in the ETS internal review process involves a team of specially trained editors who check questions for clarity, correctness, and grade-level appropriateness of language, adherence to style guidelines, and conformity to item-writing best practices. The editorial review also includes cycles of copyediting and proofreading.

Bias and Sensitivity Review

One of the final steps in the internal review process is to have all items and stimuli reviewed for bias and sensitivity. Only staff members who have participated in ETS Fairness Training conduct this bias and sensitivity review. These staff members have been trained to identify and eliminate test questions that contain content that could be construed as offensive to, or biased against, members of specific ethnic, racial, or gender groups.

Assessment Director Review

As a final quality control step, the content area's assessment director or another senior-level content reviewer reads each item before it is presented to the CDE.

Content Expert Reviews

In addition to the content reviews completed by ETS content-area experts and the content staff at the CDE, all CAHSEE items are reviewed by content review committees and bias and sensitivity review committees. The review committees are advisory panels to ETS on areas related to item development for the CAHSEE.

Content Review Meetings for CAHSEE Items

The content review committee reviews the newly developed items prior to field testing and checks for content correctness, content appropriateness, technical quality, and alignment to the California content standards.

Bias and Sensitivity Review Meetings for CAHSEE Items

The Bias and Sensitivity Review Committee reviews newly developed items prior to field testing to screen for potential bias that may affect the performance of a particular group of students.

Statewide Pupil Assessment Review Panel Review

The SPAR panel is responsible for reviewing and approving test items before they are used as operational or field-test items. The SPAR examines the items for intrusiveness into students' personal lives such as student and family beliefs, morality, religion, or sexuality. The SPAR panel representatives ensure that the test items conform to the requirements of *EC* Section 60614. The CR writing tasks are also presented to the SPAR panel for review.

If the SPAR panel rejects specific items and/or CR writing tasks, the items and/or tasks are replaced.

Data Review of Field-Tested Items

Newly developed items are field tested to obtain statistical information about item performance. The information is used to evaluate items that are on operational test forms. The CDE defines the criteria for acceptable or unacceptable item statistics. These criteria ensure that each item (1) has an appropriate level of difficulty for the target population; (2) discriminates well between examinees that differ in ability; and (3) conforms well to the statistical model underlying the measurement of the intended constructs.

Data Review Committee members review and discuss the items that have been flagged for C-level DIF. Some of the items have also been flagged for poor statistics and do not meet the psychometric criteria for item quality.

The panel members also use the results of analyses for DIF to make judgments about the appropriateness of items for various subgroups. The panelists respond to questions such as:

- Is there a content problem within the item?
- Are there any instructional issues that have negatively affected the performance of the item?

The panelists make recommendations about whether to accept or reject each item for inclusion in the CAHSEE item bank.

Quality Control of the Item Bank

After the completion of the analyses, items are placed in the item bank with their statistics. ETS delivers the items to the CDE through the CAHSEE electronic item bank. The item bank database is maintained by a staff of application systems programmers, led by the Item Bank Manager. All processes are logged; all change requests, including item bank updates for item availability status, are tracked; and all output and CAHSEE item bank deliveries undergo quality-control for accuracy.

The quality of the item bank and secure transfer of the CAHSEE item bank to the CDE are critical. The ETS internal item bank database resides on a server within the ETS firewall; access to the SQL, the server database, is strictly controlled by means of system administration. The electronic item banking application includes a login/password system to authorize access to the database or designated portions of the database. In addition, only users authorized to access the specific database are able to use the item bank. Users are authorized by a designated administrator at the CDE and at ETS.

The SFTP is the current method to deliver the CAHSEE electronic item bank to the CDE. All files posted on the SFTP site by the item bank staff are encrypted with a password.

The measures taken for ensuring the accuracy, confidentiality, and security of electronic files are as follows:

- Electronic forms of test content, documentation, and item banks are backed up electronically, with the backup media kept off-site, to prevent loss from a system breakdown or a natural disaster.

- The off-site backup files are kept in secure storage with access limited to authorized personnel only.
- Advanced network security measures are used to prevent unauthorized electronic access to the item bank.

Quality Control of Test Materials

Collecting Test Materials

Once the tests are administered, school districts return scorable materials within five working days and non-scorable materials within ten working days after the last testing day of each test administration period. Districts are provided color-coded labels identifying scorable and non-scorable materials and labels with bar-coded information identifying the school and district. The school districts apply the appropriate labels and number the cartons prior to returning the materials to the processing center. Scorable materials are returned via overnight carrier and non-scorable materials are returned by designated overland carrier.

The use of the color-coded labels streamlines the return process. All scorable materials are delivered to the Pearson scanning and scoring facilities in Iowa City, Iowa. The non-scorable materials, including test booklets, are returned to the Security Processing Department in Pearson's Cedar Rapids, Iowa, facility. ETS and Pearson closely monitor the return of materials. The CAHSEE Support Center at ETS contacts school districts that do not return their materials in a timely manner and works with them to facilitate the return of the test materials.

Processing Test Materials

Upon receipt of the test materials, Pearson uses precise inventory and test processing systems, in addition to quality assurance procedures, to maintain an up-to-date accounting of all the testing materials within their facilities. The materials are removed carefully from the shipping cartons and examined for a number of conditions, including physical damage, shipping errors, and omissions. A visual inspection to compare the number of students recorded on the School and Grade Identification (SGID) sheets with the number of answer documents in the stack is also conducted.

Pearson's image scanning process captures security information electronically and compares scorable material quantities reported on the SGIDs to actual documents scanned. School districts are contacted by phone if there are any missing shipments or if the quantity of materials returned appears to be less than expected.

Quality Control of Scanning

The CAHSEE has multiple administrations each school year, but the answer document remains the same for each administration. As such, there are two scanning quality control initiatives. The first initiative takes place during the development of the scannable form. The second initiative takes place prior to the scanning process for each of the seven administrations.

Before any CAHSEE answer documents are distributed for use, Pearson conducts a complete check of the scanning system using the new document. Pearson creates test decks of approximately 25 answer documents for ELA and mathematics marked to cover response ranges, demographic data, blanks, double marks, and other responses. Fictitious students are created to verify that each marking possibility is processed correctly by the scanning program. The output file generated as a result of this activity is thoroughly checked against each answer document after each stage to verify that the scanner is capturing marks correctly. When the program output is confirmed to match the expected results, a scan program release form is signed, and the scan program is placed in the production environment under configuration management.

For each test administration, Pearson conducts what it calls the “blue dot file.” Early return answer documents, numbering 300 to 500, are scanned and a quality control file is created. The Quality Assurance Office pulls random documents from the blue dot batch and performs one-to-one matches of the documents to the scan files, verifying that the scan program is capturing the data accurately. Not until this check is complete and signed off does full processing begin for that particular administration.

The intensity levels of each scanner are constantly monitored throughout each administration for quality control purposes. Intensity diagnostic sheets are run before and during each batch to verify that the scanner is working properly. In the event that a scanner fails to properly pick up data on the diagnostic sheets, the scanner is recalibrated before it can resume processing student documents.

Documents received in poor condition (torn, folded, or water-stained) that cannot be fed through the high-speed scanners are either scanned using a flatbed scanner or keyed into the system manually.

Post-Scanning Edits

After scanning, there are opportunities for demographic data to be edited:

- After scanning by Pearson online editors.
- After student results are posted by the CAHSEE district coordinators (online demographic data corrections).

Online corrections are limited to those changes that do not change a student's reporting status or score. Corrections may be made throughout the year up until the time ETS prepares annual reports for the CDE. If the nature of the correction is such that the student's reporting status is changed (removal of a test modification, for example), the changes are authorized by the CDE and ETS makes the corrections. Corrected data are used for quarterly and annual reporting and for technical reports.

Quality Control of Image Editing

When ETS receives the blue dot file from Pearson, the MC items are scored in the SKM system and essay images are uploaded to the OSN. The images are reviewed by online scoring leaders to confirm that the images have been saved correctly. Pearson does not begin processing answer documents until they have received this confirmation from ETS.

Quality Control of Answer Document Processing and Scoring

Processing of Answer Documents

Once processing begins, Pearson sends scanned files several times each day to ETS for scoring. The files contain several batches of up to 1,500 records per batch. Within each batch, several records at the beginning, middle, and end are identified for QC. Photocopies are made of the identified answer documents and sent to the ETS resolutions area, where they are compared with the electronic file. Any discrepancies are reported to program management for resolution. The record is put on hold until the discrepancy is resolved. This procedure assures that the scored record matches the physical answer document.

Scoring and Reporting Specifications

ETS develops standardized scoring procedures and specifications to assure testing materials are processed and scored accurately. These documents include:

- General Reporting Specifications
- Form Planner Specifications
- Matching Criteria for MC and Writing Answer Documents

Each document is explained in detail in Chapter 8. The scoring specifications are reviewed and revised by the CDE and ETS each year. After the specifications are finalized, the CDE issues a formal approval of the scoring and reporting specifications.

Matching Information on CAHSEE Answer Documents

Answer documents are designed to produce a single, complete record for each student. The record includes demographic data and scanned responses. The scored responses and the total test scores are computed and merged into the same record. All scores must comply with ETS scoring specifications.

CAHSEE answer documents contain unique numbered lithocodes that are scannable and eye-readable. The lithocodes allow all pages of the document to be linked throughout processing, even after the documents have been separated into single sheets for scanning.

Matching Multiple-Choice and Writing Scores for English-Language Arts

Each student record is assigned a unique ETS identification number. When essay scores are uploaded to the CAHSEE database, they are matched with their associated MC scores.

Storing Answer Documents

After the answer documents have been scanned, they are palletized and placed in the secure storage facilities at Pearson. The materials are stored for one year from the date of the exam. At that time, ETS requests permission to salvage the materials. After receiving CDE approval, the materials are salvaged in a secure manner.

Quality Control of Psychometric Processes

Scoring Key Verification Process

ETS takes various necessary measures to ascertain that the scoring keys are applied to the student responses as expected and the student scores are computed accurately. As described in detail in Chapter 8 (see Scoring Key Verification Process section), various quality control checks are performed before keys are finalized in the SKM system.

Quality Control of Item Analyses, Differential Item Functioning and Equating Process

The psychometric analyses conducted at ETS undergo comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists are employed by members of the team for each of the statistical procedures performed on the CAHSEE. Quality assurance checks also include comparisons of the current year's statistics to those from previous years. The results of preliminary classical item analyses provide a check on scoring keys that are also reviewed by a senior psychometrician. The items that are flagged for questionable statistical attributes are sent to test development staff for their review; their comments are reviewed by the psychometricians before items are approved for inclusion in the equating process.

In addition to the team of psychometricians and data analysts, the results of the equating process are reviewed by a psychometric manager and a senior psychometric advisor. Several pieces of informative analyses are provided to facilitate the process. The CDE also performs a replication of the equating results. The CDE replicates both content areas for the census administrations and may replicate one content area for the non-census administrations.

A few additional checks are performed for each process, as described below:

Calibrations

During the calibration process, checks are made to ascertain that the correct options for the analyses are selected. Checks are also made on the number of items, the number of examinees with valid scores, IRT Rasch item difficulties, standard errors for the Rasch item difficulties, and the match of selected statistics to the results on the same statistics obtained during preliminary item analyses. Psychometricians also perform detailed reviews of plots and statistics to investigate model fit.

Scaling

During the scaling process, checks are made on the number of linking items, their average item difficulty, the number of items dropped during the stability check of the scaling process, Rasch item difficulties, standard errors of the Rasch item difficulty estimates, and the scaling constant.

Scoring Tables

Once the equating activities are complete and raw-to-scale scoring tables are generated, the psychometricians carry out quality control checks on each scoring table. Scoring tables

are checked to verify that all raw scores are included in the tables, that scale scores increase as raw scores increase, and that the cut points for Pass, Proficient, and Advanced levels are correctly identified. As a check on the reasonableness of the cut scores, psychometricians compare passing rates of all students and students in various demographic subgroups from the current administration with passing rates from the same administration in previous years. After all quality control steps are completed and any differences are resolved, a psychometric manager and a senior psychometric advisor inspect the equating process and scoring tables as the final step in quality control.

Score Verification Process

ETS utilizes the raw-to-scale scoring tables to compute scale scores for each student. ETS verifies the scale scores by reviewing longitudinal data for reasonableness. The results are used to look at the trends for the state. The results of the longitudinal analyses are provided to the CDE and jointly discussed. If any anomalies in the results were to arise, they would be investigated further and discussed. After obtaining explanations that satisfy both the CDE and ETS, scores would then be released.

Offloads to Test Development

The statistics based on classical item analyses, DIF analyses, and IRT analyses are provided to test development staff in specially designed Excel spreadsheets called Statistical Offloads. These statistics are used for future test assembly. Before their release, the item statistic offloads are checked by the psychometric staff to make sure they are accurately combined from various analyses.

Independent Evaluation of the CAHSEE Program

HumRRO has functioned as the independent evaluator of the CAHSEE program since January 2000. During this time, HumRRO has analyzed and reported on a wide range of topics. The evaluation reports cover analyses of test results, analyses of questionnaire responses, and other evaluation activities. The annual and biennial evaluation reports may be found on the CDE CAHSEE Independent Evaluation Reports Web page at http://www.cde.ca.gov/ta/tg/hs/evaluations.asp.

Quality Control of Reporting

Quality control of reporting is carried out by two support centers at ETS. The Enterprise Scoring and Reporting (ENSR) and DQS groups work in tandem to monitor quality control of all CAHSEE reports. Quality control procedures are performed for data and production quality. The steps include:

- Data validation and verification of all extract files used for statistical analysis and production of student detail files and LEA summary reports.
- DQS replication of the summary report data to compare them to the data created by the IT group. The quality control check is completed before ENSR produces the summary report files (i.e., pdf files) that are printed and distributed to LEAs.
- Comparison of a sampling of student data on the Individual Student Report against data in the CAHSEE student data base. This comparison includes student name, birth

date, student ID, grade, County District School (CDS) code, and test results. The comparison is completed before reports are distributed.

- Comparison of summary report pdf files created from IT generated data files against summary reports created from DQS replicated summary data.
- Confirmation of the number of printed Individual Student Report impressions against the number of records on the file sent to the printer.

All reports are required to include a single, accurate CDS code, a school name, a district name, and a county name. The CDE Master File, provided monthly by the CDE, is used to validate school identity and authorization to administer the CAHSEE. Reports are not released for distribution or for posting on CAHSEE Online until all quality control processes are completed and quality standards have been met.

Excluding Student Scores from Summary Reports

ETS provides specifications to the CDE that document when to exclude student scores from accountability reporting. These specifications include the logic for handling answer documents that, for example, indicate the student tested but marked no answers, did not complete the test due to medical emergency, or tested using modifications. The methods for handling other anomalies are also covered in the specifications.

Chapter 10: Historical Comparisons

Historical comparisons of the CAHSEE results are routinely performed to identify the trends in examinee performance and test characteristics over time. Such comparisons are performed over a period of the three most recent years of administration: 2009–10, 2010–11, and 2011-12. The indicators of examinee performance include the means and standard deviations of scale scores, the percentages of examinees classified into the Pass, Proficient, and Advanced performance levels, and the observed score distributions. Test characteristics are compared by looking at the mean proportion correct, mean IRT *b*-value, mean point-biserial correlation, and the overall score reliability and SEM for each CAHSEE operational test form.

Examinee Performance

Table 10.A.1 shows the number of examinees assessed and the means and standard deviations of examinees' scale scores in 2009–10, 2010–11, and 2011-12 for the ELA and mathematics tests.

Students taking the CAHSEE are classified into Pass/Not Pass as well as ESEA performance levels: Below Proficient, Proficient, and Advanced. The percentages of students passing each content area are presented in Table 10.A.2. The percentages of students assigned to ESEA performance levels are presented in Table 10.A.3. Although passing the CAHSEE is a requirement for graduation, the SBE established the Proficient level as the desired achievement goal for all students by 2014. This goal for all students is consistent with school growth targets for state accountability and federal requirements under the ESEA.

The distributions of scale scores observed in the three most recent years are shown in Table 10.A.4 through Table 10.A.7. Scale score distributions for the summer and fall administrations are found in Table 10.A.4 for ELA and Table 10.A.6 for mathematics. Scale score distributions for the winter and spring administrations are found in Table 10.A.5 for ELA and Table 10.A.7 for mathematics. For the CAHSEE, a minimum score of 350 is required to pass the exam, and a minimum score of 380 is required to reach the proficient level of performance.

Test Characteristics

The results of the CAHSEE over the past several years indicate that the CAHSEE tests meet the technical criteria established in professional standards for high-stakes tests. Table 10.B.1 and Table 10.B.2 present, respectively, the average proportion correct values and the mean equated IRT *b*-values[12, 13] for the items on the ELA and mathematics tests. The mean proportion correct is affected both by the difficulty of the items and the abilities of the students taking them. The mean equated IRT *b*-values reflect only average item difficulty. The average point-biserial correlations for the items on the ELA and mathematics tests are presented in Table 10.B.3. The reliabilities and SEMs expressed in raw score units appear in Table 10.B.4 for both content areas across administrations and years. Like

[12] These statistics are based on the equating samples.
[13] Comparisons of mean *b*-values should only be made within a given subject test (e.g., ELA or mathematics).

the average proportion correct, point-biserial correlations and reliabilities are affected by both item characteristics and student characteristics.

Appendix 10.A—Historical Comparisons on Student Performance

Table 10.A.1: Number of Examinees Tested, Scale Score Means, and Standard Deviations of the CAHSEE Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Number of Students (with valid scores)			Scale Score Mean and Standard Deviation (SD)					
					2009–10		2010–11		2011–12	
		2009–10	2010–11	2011–12	Mean	SD	Mean	SD	Mean	SD
English–Language Arts	July	12,912	8,389	7,166	335	25	337	25	337	25
	October	41,570	43,194	36,343	340	31	342	29	341	30
	November	107,342	102,616	93,601	343	30	345	30	344	32
	December	2,754	2,405	2,319	338	28	342	25	339	26
	February	171,323	154,910	157,272	372	41	370	41	371	41
	March	393,121	402,061	388,104	376	40	381	40	378	39
	May	40,769	39,105	39,166	341	33	340	34	337	34
Mathematics	July	12,388	8,222	6,910	343	21	338	22	339	21
	October	40,423	41,519	35,680	344	26	342	26	345	27
	November	103,221	98,156	93,368	348	27	344	28	347	27
	December	2,556	2,433	2,087	342	21	340	23	343	21
	February	168,363	155,492	154,181	373	40	373	40	375	40
	March	388,590	401,838	383,396	381	40	382	39	384	39
	May	38,939	38,851	36,368	341	28	344	28	344	30

Table 10.A.2: Percentage of Students Passing Each Content Area Across 2009–10, 2010–11 and 2011–12

Admin	Percent Passing					
	English–Language Arts			Mathematics		
	2009–10	2010–11	2011–12	2009–10	2010–11	2011–12
July	24	27	27	33	26	27
October	34	37	36	34	31	38
November	39	40	39	42	35	42
December	30	36	29	30	29	34
February	69	67	68	67	68	69
March	74	77	78	75	77	79
May	35	34	29	32	36	35

Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Below Proficient			Proficient			Advanced		
		2009–10	2010–11	2011–12	2009–10	2010–11	2011–12	2009–10	2010–11	2011–12
English–Language Arts	July	95	95	95	3	4	3	2	1	2
	October	91	92	91	5	5	5	4	3	4
	November	91	90	89	6	6	6	3	4	5
	December	92	94	94	5	4	4	3	2	2
	February	56	58	57	20	19	18	24	23	25
	March	53	47	50	20	23	22	27	30	28
	May	88	88	88	6	7	7	6	5	5
Mathematics	July	96	96	97	3	3	2	1	1	1
	October	92	93	91	6	5	7	2	2	2
	November	90	91	91	7	7	7	3	2	2
	December	95	95	95	4	4	4	1	1	1
	February	58	59	57	28	26	28	14	15	15
	March	52	50	46	30	31	34	18	19	20
	May	91	90	89	7	8	8	2	2	3

Table 10.A.4: Scale Score Distributions Across 2009, 2010 and 2011 for ELA (July to December)

Scale Score Distribution	July			October			November			December		
	2009	2010	2011	2009	2010	2011	2009	2010	2011	2009	2010	2011
450	15	9	7	237	174	145	327	400	399	12	3	9
440 – 449	20	4	8	229	85	82	186	238	547	3	3	4
430 – 439	12	15	13	122	234	180	526	599	747	16	5	1
420 – 429	35	24	26	320	288	257	649	840	945	14	11	7
410 – 419	53	38	39	537	498	325	1,370	1,003	1,107	15	19	20
400 – 409	102	59	59	380	415	548	1,100	1,828	1,869	28	14	18
390 – 399	134	122	63	744	1,004	646	2,722	2,163	2,206	48	38	29
380 – 389	228	132	142	1,312	1,038	1,065	2,867	3,885	2,893	75	41	41
370 – 379	383	250	211	1,999	2,040	1,578	7,312	5,986	5,471	100	86	87
360 – 369	631	582	495	3,043	4,188	3,328	9,475	11,571	10,139	167	226	124
350 – 359	1,482	1,055	881	5,398	6,089	5,031	15,412	12,206	10,263	361	417	322
340 – 349	2,164	1,486	1,277	6,092	7,004	5,717	16,627	18,756	13,489	403	462	431
330 – 339	2,298	1,519	1,334	5,839	6,284	5,296	14,403	13,294	14,184	467	427	406
320 – 329	1,935	1,205	988	4,923	4,789	4,035	11,190	10,114	9,016	364	283	299
310 – 319	1,612	849	744	3,975	4,187	3,067	9,893	7,822	7,264	279	155	231
300 – 309	904	491	449	3,403	2,253	2,233	6,120	5,420	6,589	219	83	152
290 – 299	519	306	255	1,577	1,395	1,539	4,024	4,042	3,518	99	80	76
280 – 289	226	140	109	863	723	765	1,702	1,262	1,816	51	30	41
275 – 279	159	103	66	577	506	506	1,437	1,187	1,139	33	22	21

Table 10.A.5: Scale Score Distributions Across 2010, 2011 and 2012 for ELA (February to May)

Scale Score Distribution	February			March			May		
	2010	2011	2012	2010	2011	2012	2010	2011	2012
450	6,750	5,534	4,418	16,863	21,724	12,321	194	229	219
440 – 449	2,735	2,469	2,125	7,069	8,297	6,634	256	130	120
430 – 439	6,873	5,920	5,590	17,570	19,558	17,291	170	304	129
420 – 429	7,819	7,120	7,012	20,527	22,657	21,600	434	435	402
410 – 419	8,523	7,744	12,020	22,493	36,871	24,517	719	517	742
400 – 409	13,006	12,001	8,254	34,700	25,808	38,878	819	924	588
390 – 399	12,874	11,726	16,183	33,524	37,067	38,878	881	973	926
380 – 389	16,087	14,183	11,339	31,876	44,350	35,307	1,445	1,394	1,447
370 – 379	14,445	12,369	13,273	37,678	37,110	40,633	1,841	1,645	1,568
360 – 369	13,545	11,337	14,166	33,007	30,510	32,708	3,214	2,793	1,900
350 – 359	16,190	13,394	13,190	35,279	25,402	32,185	4,464	4,004	3,328
340 – 349	14,355	12,899	12,135	29,868	25,848	20,296	5,224	4,878	4,591
330 – 339	11,657	10,856	10,555	23,624	19,709	19,484	6,093	4,948	6,240
320 – 329	8,858	10,169	8,515	17,163	15,066	14,897	4,216	4,527	4,749
310 – 319	6,798	6,212	6,864	12,151	11,839	11,117	3,734	4,354	4,602
300 – 309	4,899	4,876	5,140	8,616	9,257	8,468	3,274	3,059	3,165
290 – 299	3,197	3,279	3,505	6,563	6,094	6,427	2,085	2,166	2,295
280 – 289	1,662	1,689	1,793	2,675	3,102	3,821	973	1,001	1,228
275 – 279	1,050	1,133	1,195	1,875	1,792	2,642	733	764	927

Table 10.A.6: Scale Score Distributions Across 2009, 2010 and 2011 for Mathematics (July to December)

Scale Score Distribution	July			October			November			December		
	2009	2010	2011	2009	2010	2011	2009	2010	2011	2009	2010	2011
450	31	16	13	218	204	240	940	806	672	6	4	5
440 – 449	6	9	9	95	91	86	349	285	324	1	0	2
430 – 439	14	17	15	239	234	213	750	687	673	5	7	3
420 – 429	24	15	17	228	171	223	778	744	678	11	11	4
410 – 419	63	19	11	382	419	243	1,217	797	722	18	6	10
400 – 409	74	48	46	470	501	628	1,331	1,322	1,204	14	20	13
390 – 399	127	88	61	534	551	458	2,129	2,062	1,919	24	32	14
380 – 389	182	119	65	1,101	1,156	993	2,765	2,582	2,602	57	48	46
370 – 379	448	219	171	1,902	1,338	1,645	5,552	4,795	5,248	61	91	83
360 – 369	1,195	435	393	3,251	3,443	3,001	9,859	8,047	8,758	205	144	170
350 – 359	1,909	1,145	1,066	5,398	4,907	5,811	17,607	11,926	16,041	368	339	368
340 – 349	2,972	1,405	1,375	8,202	7,362	5,780	19,955	17,861	15,111	565	486	424
330 – 339	2,184	1,969	1,625	6,700	7,704	6,497	15,201	17,677	15,875	515	503	463
320 – 329	1,712	1,305	983	6,291	5,783	4,477	11,989	12,603	10,421	448	344	268
310 – 319	1,047	923	667	2,930	4,607	3,428	8,456	9,548	7,950	143	236	123
300 – 309	297	340	275	1,743	2,096	1,401	3,080	4,332	3,507	95	117	70
290 – 299	65	109	82	562	721	347	891	1,382	1,223	13	35	19
280 – 289	15	26	13	99	123	103	129	418	226	4	8	1
275 – 279	23	15	23	78	108	106	243	282	214	3	2	1

Table 10.A.7: Scale Score Distributions Across 2010, 2011 and 2012 for Mathematics (February to May)

Scale Score Distribution	February			March			May		
	2010	2011	2012	2010	2011	2012	2010	2011	2012
450	9,423	8,060	9,313	33,829	26,615	26,965	201	237	365
440 – 449	3,390	3,105	3,453	9,242	19,698	20,015	106	135	118
430 – 439	7,176	6,496	6,898	9,432	9,917	10,041	215	268	310
420 – 429	7,007	6,521	6,902	27,208	19,734	20,063	240	302	338
410 – 419	7,140	9,485	6,541	17,686	29,187	29,883	303	312	388
400 – 409	10,308	9,297	9,506	25,766	28,409	28,511	533	730	618
390 – 399	10,034	11,534	12,030	32,878	35,906	36,138	820	653	899
380 – 389	15,728	10,778	11,366	31,367	33,337	33,616	964	1,343	1,054
370 – 379	11,891	12,497	13,099	36,091	37,533	36,969	1,616	1,762	1,677
360 – 369	14,075	12,138	12,483	39,142	33,684	32,189	2,436	2,817	2,360
350 – 359	16,876	15,613	15,432	29,700	36,569	27,803	5,094	5,298	4,426
340 – 349	17,137	13,576	13,060	32,305	27,102	29,090	5,774	6,895	5,848
330 – 339	12,968	14,497	13,516	22,761	26,739	19,067	7,591	5,870	5,195
320 – 329	11,034	9,904	8,808	18,806	17,556	17,547	5,414	5,274	5,761
310 – 319	8,442	7,417	6,833	14,291	10,417	8,491	4,456	4,210	3,469
300 – 309	3,872	3,237	3,331	5,879	7,323	4,908	2,110	1,890	2,669
290 – 299	1,369	835	1,181	1,458	1,379	1,548	768	619	542
280 – 289	297	265	228	434	412	286	146	114	199
275 – 279	196	237	201	315	321	266	152	122	132

Appendix 10.B—Historical Comparisons on Test Characteristics

Table 10.B.1: Average Proportion Correct of Operational Test Items Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Average *p*-value		
		2009–10	2010–11	2011–12
English–Language Arts	July	0.53	0.54	0.53
	October	0.56	0.56	0.56
	November	0.56	0.58	0.58
	December	0.57	0.57	0.54
	February	0.71	0.70	0.70
	March	0.74	0.75	0.76
	May	0.53	0.54	0.55
Mathematics	July	0.45	0.45	0.46
	October	0.48	0.47	0.48
	November	0.50	0.49	0.51
	December	0.48	0.47	0.46
	February	0.65	0.64	0.66
	March	0.67	0.69	0.71
	May	0.47	0.47	0.48

Table 10.B.2: Average IRT b-values of Operational Test Items Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Average IRT b-value		
		2009–10	2010–11	2011–12
English–Language Arts	July	-0.09	-0.07	-0.01
	October	-0.10	-0.05	-0.05
	November	-0.03	-0.06	-0.09
	December	-0.17	-0.07	-0.04
	February	-0.06	-0.04	-0.01
	March	-0.12	-0.06	-0.16
	May	0.05	-0.04	-0.11
Mathematics	July	-0.10	-0.24	-0.24
	October	-0.21	-0.25	-0.19
	November	-0.17	-0.26	-0.24
	December	-0.27	-0.25	-0.15
	February	-0.28	-0.18	-0.24
	March	-0.16	-0.18	-0.20
	May	-0.24	-0.18	-0.22

Table 10.B.3: Average Point-Biserial Correlation of Operational Test Items Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Average Point-Biserial Correlation		
		2009–10	2010–11	2011–12
English–Language Arts	July	0.31	0.31	0.31
	October	0.36	0.35	0.36
	November	0.36	0.36	0.37
	December	0.34	0.31	0.31
	February	0.45	0.45	0.46
	March	0.43	0.44	0.44
	May	0.39	0.40	0.40
Mathematics	July	0.28	0.28	0.27
	October	0.33	0.33	0.34
	November	0.34	0.35	0.34
	December	0.28	0.30	0.28
	February	0.46	0.47	0.46
	March	0.46	0.46	0.45
	May	0.34	0.35	0.37

Table 10.B.4: Reliabilities and Standard Errors of Measurement (SEMs) of Operational Test Forms Across 2009–10, 2010–11 and 2011–12

Subject	Admin	Reliability			SEM		
		2009–10	2010–11	2011–12	2009–10	2010–11	2011–12
English–Language Arts	July	0.87	0.87	0.87	4.38	4.37	4.40
	October	0.91	0.89	0.90	4.35	4.43	4.37
	November	0.90	0.90	0.91	4.38	4.34	4.36
	December	0.89	0.86	0.87	4.42	4.32	4.25
	February	0.92	0.94	0.94	4.38	4.10	4.05
	March	0.93	0.94	0.93	3.91	3.81	3.81
	May	0.92	0.92	0.92	4.49	4.46	4.48
Mathematics	July	0.85	0.86	0.85	4.11	4.12	4.16
	October	0.89	0.90	0.90	4.07	4.10	4.10
	November	0.90	0.91	0.90	4.08	4.06	4.10
	December	0.85	0.87	0.85	4.15	4.13	4.14
	February	0.95	0.95	0.95	3.68	3.71	3.67
	March	0.95	0.95	0.95	3.61	3.59	3.56
	May	0.91	0.91	0.92	4.06	4.08	4.05

References

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

California Department of Education. (2006). *Mathematics Framework for California Public Schools*. Sacramento, CA: California Department of Education.

California Department of Education. (2011a). *California High School Exit Examination District and Test Site Coordinator's Manual (July 2010–May 2011)*. Sacramento, CA: California Department of Education.

California Department of Education. (2011b). *California High School Exit Examination Directions for Administration (July 2010–May 2011).* Sacramento, CA; California Department of Education.

California Department of Education. (2011c). *California Content Standards*. http://www.cde.ca.gov/be/st/ss/.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Cronbach, L. J. (1951). Coefficient alpha and the independent structure of tests. *Psychometrika, 16*, 292-334.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.). *Educational Measurement* (2nd ed.). Washington, DC: American Council on Education.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Drasgow, F. (1988). Polychoric and polyserial correlations. In Kotz L., Johnson N. L. (Eds.), *Encyclopedia of statistical sciences (Vol. 7).* New York: Wiley.

Educational Testing Service. (2002). *ETS Standards for Quality and Fairness*. Office of Testing Integrity, Princeton, NJ: Educational Testing Service.

Feldt, L.S., & Brennan, R. L. (1989). Reliability. In R.L. Linn (Ed.), *Educational Measurement* (3rd ed). Phoenix, AZ: Oryx Press.

Gaffney, T., & Perryman, C. (July, 2009). *A longitudinal look at the factor structure of educational achievement tests.* Paper presented at the International Meeting of the Psychometric Society. University of Cambridge: Cambridge, UK.

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed.). Washington, DC: American Council on Education and National Council on Measurement in Education.

Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (1998, April). *The bookmark standard setting procedure: Methodology and recent implementations.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Lewis. C., & Thayer, D. T. (1996). Unpublished Seminar Notes.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal of Educational Measurement, 32*, 179-197.

Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.) (pp.13–103). New York: Macmillan.

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

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *17*, 159-176.

Muraki, E. and Bock, R. D. (1995). *PARSCALE: Parameter scaling of rating data* (Version 2.2). Chicago, IL: Scientific Software, Inc.

Patrick, R. & Way, D. (March, 2008). *Field testing and equating designs for state educational assessments*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201-210.

Taylor, L., Johnstone, C., & Hardoin, M. M. (20011) 2011 review of CAHSEE test quality. In D. E. Becker, L. L. Wise, M. M. Hardoin, and C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination (CAHSEE): 2011 Evaluation Report.* (FR-11-51). Alexandria, VA: Human Resources Research Organization.

United States Department of Education (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Retrieved September 21, 2009 from http://www.ed.gov/policy/elsec/guid/saaprguidance.pdf.

Webb, N. L. (2005). *Web alignment tool: Training manual*. Madison, WI: Wisconsin Center for Education Research.

Wise, L. L. (2007). A closer look at specific populations. In D. E. Becker & C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination (CAHSEE): 2007 Evaluation Report.* (FR-07-69). Alexandria, VA: Human Resources Research Organization.

Wise, L. L. (2009). A closer look at students who did not pass. In D. E. Becker, L. L. Wise, and C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination (CAHSEE): 2009 Evaluation Report.* (FR-09-65). Alexandria, VA: Human Resources Research Organization.

Wise, L. L., & Rui, N. (2007). Results from the 2006-07 administrations. In D. E. Becker & C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination (CAHSEE): 2007 Evaluation Report.* (FR-07-69). Alexandria, VA: Human Resources Research Organization.