California High School Exit Examination

Technical Report

July 2010–May 2011 Administrations

Prepared by

Educational Testing Service

for

the California Department of Education

March 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 2010–11 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 2010 29
APPENDIX 2.B: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—OCTOBER 2010 39
APPENDIX 2.C: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—NOVEMBER 2010 51
APPENDIX 2.D: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—DECEMBER 2010 66
APPENDIX 2.E: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—FEBRUARY 2011 73
APPENDIX 2.F: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MARCH 2011 87
APPENDIX 2.G: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MAY 2011 103

CHAPTER 3: ITEM DEVELOPMENT **116**

1.0 Item Specifications *116*

2.0 Prepare Item Development Plan *118*
3.0 Train Item Writers *118*
4.0 Order Items Based on Needs Assessment *119*
5.0 Internal Item Reviews (Educational Testing Service) *119*
6.0 External Item Reviews (California Educators and California Department of Education) *121*
7.0 Item Banking *124*
8.0 Create Field Test Sets *125*
9.0 Field-Test Items *126*
10.0 Review and Process Field Test Data *127*
11.0 Create Operational Forms with California Department of Education Review *128*
12.0 Administer Operational Test *128*

CHAPTER 4: TEST DEVELOPMENT 129

TEST LENGTH 129
RULES FOR ITEM SELECTION 129
Test Blueprint *129*
Content Rules and Item Selection *130*
Psychometric Criteria *130*
Rules for Item Sequence and Layout *135*

CHAPTER 5: TEST ADMINISTRATION 136

TEST SECURITY AND CONFIDENTIALITY 136
Educational Testing Service Office of Testing Integrity *136*
Test Development *136*
Item Review *137*
Item Bank *137*
Transfer of Forms and Items to the California Department of Education *138*
Printing *138*
Test Administration *138*
Test Delivery *138*
Processing Security *140*
Data Management *140*
Transfer of Files via Secure Data Exchange *140*
Scoring and Analysis *140*
Reporting and Posting Results *141*
Student Confidentiality *141*
Data Security *141*
PROCEDURES TO MAINTAIN STANDARDIZATION 142
Test Administrators *142*
CAHSEE Directions for Administration *143*
CAHSEE District and Test Site Coordinator's Manual *143*
CAHSEE Online *144*
Test Booklets *144*
STUDENTS WITH DISABILITIES 144
Identification of Students with Test Variations, Accommodations or Modifications *145*
Scoring *145*
DEMOGRAPHIC DATA CORRECTIONS 145
TESTING IRREGULARITIES 145
TEST ADMINISTRATION INCIDENTS 146

CHAPTER 6: ANALYSES 147

SAMPLES USED FOR THE ANALYSES 148
CLASSICAL ITEM ANALYSES 149

Summary of Item Statistics *152*
Procedures for Documenting Items That Fail to Meet the Desired Psychometric Criteria *153*
DIFFERENTIAL ITEM FUNCTIONING ANALYSES 153
ITEM RESPONSE THEORY ANALYSES (CALIBRATION, SCALING AND EQUATING) 156
Measurement Model *156*
Item Calibration and Scaling *156*
True-Score Equating *157*
Equating Braille, Large-Print and Audio CD Forms *158*
Raw-Score to Scale-Score Conversion Tables and Conditional Standard Errors of Measurement *158*
Scaling Field-Test Items *159*
Item Response Theory Model-Data Fit Analyses *159*
Summaries of Scaled Item Response Theory B-Values *160*
RELIABILITY ANALYSES 160
Test Reliabilities and Standard Errors of Measurement *160*
Strand Intercorrelations, Reliabilities and Standard Errors of Measurement *162*
Subgroup Reliabilities and Standard Errors of Measurement *162*
Writing Prompt and Rater Agreement Summary *163*
Decision Classification Analyses *164*
VALIDITY EVIDENCE 165
Test Purpose *166*
Constructs to Be Measured *166*
Intended Test Population *166*
Scores Generated and the Interpretations and Uses of These Scores *167*
Evidence Based on Content *167*
Evidence Based on Relations to Other Variables *169*
Evidence Based on Response Processes *172*
Evidence Based on Internal Structure *172*
Additional Validity Evidence *174*
Conclusions *175*
APPENDIX 6.A: CAHSEE ITEM REVIEW—DESCRIPTION AND EXAMPLES OF CLASSIFICATION CATEGORIES 176
APPENDIX 6.B: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—JULY 2010 179
APPENDIX 6.C: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—OCTOBER 2010 184
APPENDIX 6.D: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—NOVEMBER 2010 189
APPENDIX 6.E: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—DECEMBER 2010 194
APPENDIX 6.F: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—FEBRUARY 2011 199
APPENDIX 6.G: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MARCH 2011 204
APPENDIX 6.H: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MAY 2011 209
APPENDIX 6.I: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—FEBRUARY 2011 214
APPENDIX 6.J: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—MARCH 2011 218
APPENDIX 6.K: INTERCORRELATIONS, RELIABILITY ESTIMATES AND STANDARD ERRORS OF MEASUREMENT 224
APPENDIX 6.L: RATER AGREEMENT ANALYSES 233
APPENDIX 6.M: GENERALIZABILITY ANALYSES 240
APPENDIX 6.N: DECISION CLASSIFICATION RELIABILITY ANALYSES 242
APPENDIX 6.O: SCORING TABLES FOR OPERATIONAL AND SPECIAL TEST VERSIONS 256
APPENDIX 6.P: STANDARD ERRORS OF THETA BASED ON WEIGHTED RAW SCORES 276

CHAPTER 7: PERFORMANCE STANDARDS 277

BACKGROUND AND PROCEDURES 277
CAHSEE PASSING SCORES 278
RESULTS 279

CHAPTER 8: SCORING AND REPORTING 281

PROCEDURES FOR MAINTAINING AND RETRIEVING INDIVIDUAL SCORES 281

Scoring and Reporting Specifications *281*
Scanning and Scoring *282*
TYPES OF SCORES AND STRAND SCORES 283
Raw Score *283*
Strand Score *283*
Scale Score *283*
Passing Scores *283*
Score for Writing Applications *283*
SCORE VERIFICATION PROCEDURES 284
Scoring Key Verification Process *284*
Monitoring and Quality Control of Writing Scoring *284*
Quality Control in Raw-to-Scale Score Conversions *287*
OVERVIEW OF SCORE AGGREGATION PROCEDURES 287
Individual Scores *288*
Group Scores *288*
REPORTS TO BE PRODUCED AND SCORES FOR EACH REPORT 289
Types of Score Reports *289*
Score Report Contents *290*
Score Report Applications *291*
CRITERIA FOR INTERPRETING TEST SCORES 291
CRITERIA FOR INTERPRETING SCORE REPORTS 291
APPENDIX 8.A: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—JULY 2010 292
APPENDIX 8.B: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—OCTOBER 2010 300
APPENDIX 8.C: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—NOVEMBER 2010 308
APPENDIX 8.D: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—DECEMBER 2010 316
APPENDIX 8.E: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—FEBRUARY 2011 324
APPENDIX 8.F: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—MARCH 2011 332
APPENDIX 8.G: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—MAY 2011 340

CHAPTER 9: QUALITY CONTROL PROCEDURES 348

QUALITY CONTROL OF ITEM DEVELOPMENT 348
Item Specifications *348*
Item Writers *348*
Internal Contractor Reviews *348*
Content Expert Reviews *349*
Statewide Pupil Assessment Review Panel Review *349*
Data Review of Field Tested Items *350*
QUALITY CONTROL OF THE ITEM BANK 350
QUALITY CONTROL OF TEST MATERIALS 351
Collecting Test Materials *351*
Processing Test Materials *351*
QUALITY CONTROL OF SCANNING 351
Post-Scanning Edits *352*
QUALITY CONTROL OF IMAGE EDITING 352
QUALITY CONTROL OF ANSWER DOCUMENT PROCESSING AND SCORING 353
Processing of Answer Documents *353*
Scoring and Reporting Specifications *353*
Matching Information on CAHSEE Answer Documents *353*
Matching Multiple-Choice and Writing Scores for English-Language Arts *353*
Storing Answer Documents *353*
QUALITY CONTROL OF PSYCHOMETRIC PROCESSES 354
Scoring Key Verification Process *354*
Quality Control of Item Analyses, Differential Item Functioning and Equating Process *354*

Score Verification Process *355*
Offloads to Test Development *355*
Independent Evaluation of the CAHSEE Program *355*
QUALITY CONTROL OF REPORTING 355
Excluding Student Scores from Summary Reports *356*

CHAPTER 10: HISTORICAL COMPARISONS 357

EXAMINEE PERFORMANCE 357
TEST CHARACTERISTICS 357
APPENDIX 10.A—HISTORICAL COMPARISONS ON STUDENT PERFORMANCE 358
APPENDIX 10.B—HISTORICAL COMPARISONS ON TEST CHARACTERISTICS 364

CHAPTER 11: ALTERNATIVE MEANS PILOT STUDY 366

INTRODUCTION 366
CAHSEE Alternative Means Background *366*
CONCEPTUAL FRAMEWORK 370
Proof of Concept *370*
Collection of Evidence *370*
METHODS 371
Recruitment *371*
Matrix Sampling of Strands *371*
Directions for Administration *372*
Evaluation Procedures *372*
Survey Methodology *373*
Focus Groups *374*
RESULTS 374
Quantitative Data *374*
Qualitative Data *383*
SUMMARY OF FINDINGS 387
Collection of Evidence *387*
Scoring Procedures *388*
Passing Rates *389*
Timeline *389*
RECOMMENDATIONS 389

REFERENCES 393

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 2010)....3
Table E.3.2: CAHSEE Summary Statistics—English-Language Arts (November and December 2010, February 2011)....4
Table E.3.3: CAHSEE Summary Statistics—English-Language Arts (March and May 2011)....5
Table E.4.1: CAHSEE Summary Statistics—Mathematics (July and October 2010)....6
Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2010, February 2011)....7
Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2011)....8
Table 1.1: Testing Date for Each Administration by Content: 2010–11....11
Table 2.1: Items Included in the 2010–11 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: 2010–11....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 2010....29
Table 2.A.2: Summary Statistics by Testing Variations and Disability—Mathematics, July 2010....30
Table 2.A.3: Demographic Summary for All Examinees by Testing Variations—ELA, July 2010....31
Table 2.A.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, July 2010....32
Table 2.A.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, July 2010....33
Table 2.A.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, July 2010....34
Table 2.A.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, July 2010....35
Table 2.A.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, July 2010....36
Table 2.A.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, July 2010....37
Table 2.A.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, July 2010....38
Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2010....39
Table 2.B.2: Summary Statistics by Testing Variations and Disability—Mathematics, October 2010....40
Table 2.B.3: Demographic Summary for All Examinees by Testing Variations—ELA, October 2010....41
Table 2.B.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, October 2010....42
Table 2.B.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, October 2010....43
Table 2.B.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, October 2010....44
Table 2.B.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, October 2010....45
Table 2.B.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, October 2010....47
Table 2.B.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, October 2010....49
Table 2.B.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, October 2010....50
Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2010....51
Table 2.C.2: Summary Statistics by Testing Variations and Disability—Mathematics, November 2010....52
Table 2.C.3: Demographic Summary for All Examinees by Testing Variations—ELA, November 2010....53
Table 2.C.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, November 2010....54
Table 2.C.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, November 2010....55
Table 2.C.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, November 2009....56
Table 2.C.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, November 2010....57
Table 2.C.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, November 2010....60
Table 2.C.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, November 2010....62
Table 2.C.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, November 2010....64
Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2010....66
Table 2.D.2: Summary Statistics by Testing Variations and Disability—Mathematics, December 2010....66
Table 2.D.3: Demographic Summary for All Examinees by Testing Variations—ELA, December 2010....67
Table 2.D.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, December 2010....68
Table 2.D.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, December 2010....69

Table 2.D.6: Percentiles of Scale Scores for Students with Testing Variations— Mathematics, December 2010 70
Table 2.D.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, December 2010 71
Table 2.D.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, December 2010 71
Table 2.D.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, December 2010 72
Table 2.D.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, December 2010 72
Table 2.E.1: Summary Statistics by Testing Variations and Disability—ELA, February 2011 73
Table 2.E.2: Summary Statistics by Testing Variations and Disability—Mathematics, February 2011.....74
Table 2.E.3: Demographic Summary for All Examinees by Testing Variations—ELA, February 2011.....75
Table 2.E.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, February 2011 76
Table 2.E.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, February 2011 77
Table 2.E.6: Percentiles of Scale Scores for Students with Testing Variations — Mathematics, February 2011.....78
Table 2.E.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, February 2011 79
Table 2.E.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, February 2011.....81
Table 2.E.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, February 2011.....83
Table 2.E.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, February 2011 85
Table 2.F.1: Summary Statistics by Testing Variations and Disability—ELA, March 2011.....87
Table 2.F.2: Summary Statistics by Testing Variations and Disability—Mathematics, March 2011 88
Table 2.F.3: Demographic Summary for All Examinees by Testing Variations—ELA, March 2011 89
Table 2.F.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, March 2011.....90
Table 2.F.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, March 2011.....91
Table 2.F.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, March 2011 92
Table 2.F.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, March 2011.93
Table 2.F.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, March 2011 96
Table 2.F.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, March 2011 99
Table 2.F.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, March 2011.....101
Table 2.G.1: Summary Statistics by Testing Variations and Disability—ELA, May 2011 103
Table 2.G.2: Summary Statistics by Testing Variations and Disability—Mathematics, May 2011.....104
Table 2.G.3: Demographic Summary for All Examinees by Testing Variations—ELA, May 2011.....105
Table 2.G.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, May 2011 106
Table 2.G.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, May 2011 107
Table 2.G.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, May 2011 108
Table 2.G.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, May 2011 .109
Table 2.G.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, May 2011 111
Table 2.G.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, May 2011 113
Table 2.G.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, May 2011 115
Table 3.1 Field-Test Percentages for the CAHSEE 118
Table 3.2 Summary of Items and Forms Presented in 2011 CAHSEE 127
Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items.....133
Table 4.2: Number of Items and Mean B-Value Ranges by Content Strand for ELA 133
Table 4.3: Difficulty (B) and Discrimination (R-bis) Specifications for Mathematics 134
Table 4.4: Number of Items and Mean B-Value Ranges by Content Strand for Mathematics.....134
Table 6.1: Listing of Summary Tables for Items 148
Table 6.2: Flagging Criteria for Classical Item Analyses 153
Table 6.3: DIF Categories 155
Table 6.4: Field-Test Items Recommended for Future Forms.....160
Table 6.B.1: Summary of Operational Item Statistics—ELA 179
Table 6.B.2: Summary of Operational Item Statistics—Mathematics 179
Table 6.B.3: IRT Model Data Fit Distribution of Operational Items—ELA 180

Table 6.B.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 180
Table 6.B.5: Operational Items Containing Significant DIF 180
Table 6.B.6: Distribution of Operational Item DIF Classifications—ELA 181
Table 6.B.7: Distribution of Operational Item DIF Classifications—Mathematics 182
Table 6.B.8: Listing of CR Item Statistics—ELA 183
Table 6.C.1: Summary of Operational Item Statistics—ELA 184
Table 6.C.2: Summary of Operational Item Statistics—Mathematics 184
Table 6.C.3: IRT Model Data Fit Distribution of Operational Items—ELA 185
Table 6.C.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 185
Table 6.C.5: Operational Items Containing Significant DIF 185
Table 6.C.6: Distribution of Operational Item DIF Classifications—ELA 186
Table 6.C.7: Distribution of Operational Item DIF Classifications—Mathematics 187
Table 6.C.8: Listing of CR Item Statistics—ELA 188
Table 6.D.1: Summary of Operational Item Statistics—ELA 189
Table 6.D.2: Summary of Operational Item Statistics—Mathematics 189
Table 6.D.3: IRT Model Data Fit Distribution of Operational Items—ELA 190
Table 6.D.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 190
Table 6.D.5: Operational Items Containing Significant DIF 190
Table 6.D.6: Distribution of Operational Item DIF Classifications—ELA 191
Table 6.D.7: Distribution of Operational Item DIF Classifications—Mathematics 192
Table 6.D.8: Listing of CR Item Statistics—ELA 193
Table 6.E.1: Summary of Operational Item Statistics—ELA 194
Table 6.E.2: Summary of Operational Item Statistics—Mathematics 194
Table 6.E.3: IRT Model Data Fit Distribution of Operational Items—ELA 195
Table 6.E.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 195
Table 6.E.5: Operational Items Containing Significant DIF 195
Table 6.E.6: Distribution of Operational Item DIF Classifications—ELA 196
Table 6.E.7: Distribution of Operational Item DIF Classifications—Mathematics 197
Table 6.E.8: Listing of CR Item Statistics—ELA 198
Table 6.F.1: Summary of Operational Item Statistics—ELA 199
Table 6.F.2: Summary of Operational Item Statistics—Mathematics 199
Table 6.F.3: IRT Model Data Fit Distribution of Operational Items—ELA 200
Table 6.F.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 200
Table 6.F.5: Operational Items Containing Significant DIF 200
Table 6.F.6: Distribution of Operational Item DIF Classifications—ELA 201
Table 6.F.7: Distribution of Operational Item DIF Classifications—Mathematics 202
Table 6.F.8: Listing of CR Item Statistics—ELA 203
Table 6.G.1: Summary of Operational Item Statistics—ELA 204
Table 6.G.2: Summary of Operational Item Statistics—Mathematics 204
Table 6.G.3: IRT Model Data Fit Distribution of Operational Items—ELA 205
Table 6.G.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 205
Table 6.G.5: Operational Items Containing Significant DIF 205
Table 6.G.6: Distribution of Operational Item DIF Classifications—ELA 206
Table 6.G.7: Distribution of Operational Item DIF Classifications—Mathematics 207
Table 6.G.8: Listing of CR Item Statistics—ELA 208
Table 6.H.1: Summary of Operational Item Statistics—ELA 209
Table 6.H.2: Summary of Operational Item Statistics—Mathematics 209
Table 6.H.3: IRT Model Data Fit Distribution of Operational Items—ELA 210
Table 6.H.4: IRT Model Data Fit Distribution of Operational Items—Mathematics 210
Table 6.H.5: Operational Items Containing Significant DIF 210
Table 6.H.6: Distribution of Operational Item DIF Classifications—ELA 211
Table 6.H.7: Distribution of Operational Item DIF Classifications—Mathematics 212
Table 6.H.8: Listing of CR Item Statistics—ELA 213
Table 6.I.1: Summary of Field-Test Item Statistics—ELA 214
Table 6.I.2: Summary of Field-Test Item Statistics—Mathematics 214
Table 6.I.3: IRT Model Data Fit Distribution of Field-Test Items—ELA 215
Table 6.I.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics 215
Table 6.I.5: Field-Test Items Containing Significant DIF 215
Table 6.I.6: Distribution of Field-Test Item DIF Classifications—ELA 216
Table 6.I.7: Distribution of Field-Test Item DIF Classifications—Mathematics 217
Table 6.J.1: Summary of Field-Test Item Statistics—ELA 218
Table 6.J.2: Summary of Field-Test Item Statistics—Mathematics 218

Table 6.J.3: IRT Model Data Fit Distribution of Field-Test Items—ELA....219
Table 6.J.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics....219
Table 6.J.5: Field-Test Items Containing Significant DIF....220
Table 6.J.6: Distribution of Field-Test Item DIF Classifications—ELA....222
Table 6.J.7: Distribution of Field-Test Item DIF Classifications—Mathematics....223
Table 6.K.1: Intercorrelations and Reliability Estimates by Section—July 2010....224
Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2010....225
Table 6.K.3: Intercorrelations and Reliability Estimates by Section—November 2010....226
Table 6.K.4: Intercorrelations and Reliability Estimates by Section—December 2010....227
Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 2011....228
Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 2011....229
Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 2011....230
Table 6.K.8: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEM) of Subgroups for ELA—February 2011....231
Table 6.K.9: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—February 2011....231
Table 6.K.10: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEM) of Subgroups for ELA—March 2011....232
Table 6.K.11: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—March 2011....232
Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2010....233
Table 6.L.2: Summary Statistics for the ELA Essay Item—July 2010....233
Table 6.L.3: Agreement of First and Second Ratings on the ELA Essay Item—October 2010....234
Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2010....234
Table 6.L.5: Agreement of First and Second Ratings on the ELA Essay Item—November 2010....235
Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2010....235
Table 6.L.7: Agreement of First and Second Ratings on the ELA Essay Item—December 2010....236
Table 6.L.8: Summary Statistics for the ELA Essay Item—December 2010....236
Table 6.L.9: Agreement of First and Second Ratings on the ELA Essay Item—February 2011....237
Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2011....237
Table 6.L.11: Agreement of First and Second Ratings on the ELA Essay Item—March 2011....238
Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2011....238
Table 6.L.13: Agreement of First and Second Ratings on the ELA Essay Item—May 2011....239
Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2011....239
Table 6.M.1: Generalizability Results—July 2010....240
Table 6.M.2: Generalizability Results—October 2010....240
Table 6.M.3: Generalizability Results—November 2010....240
Table 6.M.4: Generalizability Results—December 2010....240
Table 6.M.5: Generalizability Results—February 2011....241
Table 6.M.6: Generalizability Results—March 2011....241
Table 6.M.7: Generalizability Results—May 2011....241
Table 6.N.1: ESEA Reliability Classifications—July 2010....242
Table 6.N.2: Pass/Not Pass Classifications—July 2010....243
Table 6.N.3: ESEA Reliability Classifications—October 2010....244
Table 6.N.4: Pass/Not Pass Classifications—October 2010....245
Table 6.N.5: ESEA Reliability Classifications—November 2010....246
Table 6.N.6: Pass/Not Pass Classifications—November 2010....247
Table 6.N.7: ESEA Reliability Classifications—December 2010....248
Table 6.N.8: Pass/Not Pass Classifications—December 2010....249
Table 6.N.9: ESEA Reliability Classifications—February 2011....250
Table 6.N.10: Pass/Not Pass Classifications—February 2011....251
Table 6.N.11: ESEA Reliability Classifications—March 2011....252
Table 6.N.12: Pass/Not Pass Classifications—March 2011....253
Table 6.N.13: ESEA Reliability Classifications—May 2011....254
Table 6.N.14: Pass/Not Pass Classifications—May 2011....255
Table 6.O.1: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, July 2010....256
Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2010....257
Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2010....258
Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2010....259
Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2010....260
Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2010....261
Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2010....262
Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2010....263

Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2011 264
Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2011 .. 265
Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2011 266
Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2011 267
Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2011 268
Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2011 269
Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA— July, October and November 2010 270
Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July, October and November 2010 271
Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—December 2010 and February 2011 272
Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— December 2010 and February 2011 273
Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—March and May 2011 274
Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—March and May 2011 275
Table 7.1: Passing Scores on the February 2004 CAHSEE 278
Table 7.2: Summary of Cut Scores and Passing Rates Based on Equating Samples 280
Table 8.1: Listing of Frequency Distribution and Demographic Summary Tables 289
Table 8.2: Types of CAHSEE Reports 290
Table 8.A.1: Frequency Distributions, ELA—July 2010 292
Table 8.A.2: Frequency Distributions, Mathematics—July 2010 292
Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2010 293
Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2010 293
Table 8.A.5: Demographic Summary for All Examinees ELA—July 2010 294
Table 8.A.6: Demographic Summary for All Examinees, Mathematics—July 2010 295
Table 8.A.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA— July 2010 296
Table 8.A.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—July 2010 297
Table 8.A.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—July 2010 298
Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2010 . 299
Table 8.B.1: Frequency Distributions, ELA—October 2010 300
Table 8.B.2: Frequency Distributions, Mathematics—October 2010 300
Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2010 301
Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2010 301
Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2010 302
Table 8.B.6: Demographic Summary for All Examinees, Mathematics—October 2010 303
Table 8.B.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—October 2010 304
Table 8.B.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—October 2010 305
Table 8.B.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—October 2010 306
Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2010 307
Table 8.C.1: Frequency Distributions, ELA—November 2010 308
Table 8.C.2: Frequency Distributions, Mathematics—November 2010 308
Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2010 309
Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2010 309
Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2010 310
Table 8.C.6: Demographic Summary for All Examinees, Mathematics—November 2010 311
Table 8.C.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—November 2010 312
Table 8.C.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—November 2010 313
Table 8.C.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—November 2010 314
Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2010 315
Table 8.D.1: Frequency Distributions, ELA—December 2010 316
Table 8.D.2: Frequency Distributions, Mathematics—December 2010 316
Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2010 317
Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2010 317
Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2010 318
Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2010 319
Table 8.D.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA —December 2010 320
Table 8.D.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—December 2010 321
Table 8.D.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—December 2010 322
Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2010 323
Table 8.E.1: Frequency Distributions, ELA—February 2011 324

Table 8.E.2: Frequency Distributions, Mathematics—February 2010 324
Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2011 325
Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2011 325
Table 8.E.5: Demographic Summary for All Examinees, ELA February 2011 326
Table 8.E.6: Demographic Summary for All Examinees, Mathematics—February 2011 327
Table 8.E.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—February 2011 328
Table 8.E.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— February 2011 329
Table 8.E.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—February 2011 330
Table 8.E.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—February 2011 331
Table 8.F.1: Frequency Distributions, ELA—March 2011 332
Table 8.F.2: Frequency Distributions, Mathematics—March 2011 332
Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2011 333
Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2011 333
Table 8.F.6: Demographic Summary for All Examinees, Mathematics—March 2011 335
Table 8.F.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—March 2011 336
Table 8.F.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—March 2011 337
Table 8.F.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—March 2011 338
Table 8.F.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—March 2011 339
Table 8.G.1: Frequency Distributions, ELA—May 2011 340
Table 8.G.2: Frequency Distributions, Mathematics—May 2011 340
Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2011 341
Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2011 341
Table 8.G.6: Demographic Summary for All Examinees, Mathematics—May 2011 343
Table 8.G.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—May 2011 344
Table 8.G.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— May 2011 345
Table 8.G.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—May 2011 346
Table 8.G.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—May 2011 347
Table 10.A.1: Number of Examinees Tested, Scale Score Means, and Standard Deviations of CAHSEE across 2008–09, 2009–10 and 2010–11 358
Table 10.A.2: Percentage of Students Passing Each Content Area across 2008–09, 2009–10 and 2010–11 358
Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced across 2008–09, 2009–10 and 2010–11 ..359
Table 10.A.4: Scale Score Distributions across 2008, 2009 and 2010 for ELA (July to December) 360
Table 10.A.5: Scale Score Distributions across 2009, 2010 and 2011 for ELA (February to May) 361
Table 10.A.6: Scale Score Distributions across 2008, 2009 and 2010 for Mathematics (July to December) 362
Table 10.A.7: Scale Score Distributions across 2009, 2010 and 2011 for Mathematics (February to May) 363
Table 10.B.1: Average Proportion Correct of Operational Test Items across 2008–09, 2009–10 and 2010–11 364
Table 10.B.2: Average IRT b-values of Operational Test Items across 2008–09, 2009–10 and 2010–11 364
Table 10.B.3 Average Point-Biserial Correlation of Operational Test Items across 2008–09, 2009–10 and 2010–11 365
Table 10.B.4 Reliabilities and Standard Errors of Measurement (SEM) of Operational Test Forms across 2008–09, 2009–10 and 2010–11 365
Table 11.1 Student Demographic Characteristics by Content Area 376
Table 11.2. Special Education Considerations by Content Area 377
Table 11.3 Teacher Reported “Most Recent” CAHSEE Test Results: Descriptive Statistics 378
Table 11.4 Teacher Reported Number of CAHSEE Attempts by Content Area 378
Table 11.5 Teacher Reported CAHSEE Pass Status by Content Area and Student Sample 378
Table 11.6. Task Difficulty by Work Sample and Strand for ELA 380
Table 11.7. Task Difficulty by Work Sample and Strand for Mathematics 380
Table 11.8. Correlations and Standard Errors between Rater Scores and CAHSEE Strand Scores 381
Table 11.9. Summary Statistics of Rater Agreement for ELA 382
Table 11.10. Summary Statistics of Rater Agreement for Mathematics 383

List of Figures

Figure 3.1 The ETS Item Development Process for the CAHSEE 116
Figure 4.1: Example of Test Characteristic Curves 132
Figure 6.1: CAHSEE Item Analysis — Mathematics Item 152
Figure 6.2: CAHSEE Item Analysis — ELA Item 152
Figure 6.3: Samples of CAHSEE Item-Fit Rating Categories 178
Figure 11.1 Flowchart representing the Proposed Two-Tiered CAHSEE Alternative Means Process 368
Figure 11.2: CAHSEE Alternative Means Work Sample Evaluation Rubric 372

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 2010, and February, March, and May 2011.

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 2010, and February, March, and May 2011 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 that 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 2010	13,106	3,505	4,884	4,717
October 2010	60,416	24,297	18,897	17,222
November 2010	140,049	60,723	41,893	37,433
December 2010	3,914	924	1,481	1,509
February 2011	178,902	131,500	23,410	23,992
March 2011	432,383	371,516	30,545	30,322
May 2011	57,771	20,185	18,920	18,666

[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 2010	8,389	2,290 (27)	8,222	2,130 (26)
October 2010	43,194	16,053 (37)	41,519	13,015 (31)
November 2010	102,616	40,719 (40)	98,156	34,053 (35)
December 2010	2,405	863 (36)	2,433	702 (29)
February 2011	154,910	103,797 (67)	155,492	105,524 (68)
March 2011	402,061	309,354 (77)	401,838	310,589 (77)
May 2011	39,105	13,348 (34)	38,851	13,857 (36)

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 2010–11 CAHSEE administrations included in this report are presented in Tables E.3 and E.4 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 2010)

Administration	July 2010	October 2010
Scale Score Information		
Number of Examinees	8,389	43,194
Mean	337	342
SD[1]	25	29
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	336	342
Raw Score Information		
Number of Examinees	8,389	43,194
Mean	47.91	50.17
SD	12.11	13.65
Possible Range	0-90	0-90
Obtained Range	1-89	0-90
Median	48	51
Test Information		
Reliability	0.87	0.89
Raw Score Standard Error of Measurement	4.37	4.43
Mean Omits	0.59	0.68
SD Omits	4.76	5.07
Percent Responding to:	89	88
All Items	97	97
All Items - 1 Item	98	98
All Items - 2 Items	99	98
All Items - 3 Items	99	99
All Items - 4 Items	99	99
All Items - 5 Items	99	99
Item Information[2]		
Number of Items	72	72
Mean Observed Average Item Score (AIS)	0.54	0.56
Equated Mean Rasch B-Value	-0.07	-0.05
Mean R-biserial	0.40	0.45
SD R-biserial	0.11	0.08
Examinee Information		
ELA-Only Examinees		
Number of Examinees	4,884	18,897
Mean Scale Score	337	343
SD Scale Score	21	24
Median Scale Score	338	344
ELA and Mathematics Examinees		
Number of Examinees	3,505	24,297
Mean Scale Score	337	342
SD Scale Score	30	33
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 2010, February 2011)

Administration	November 2010	December 2010	February 2011
Scale Score Information			
Number of Examinees	102,616	2,405	154,910
Mean	345	342	370
SD	30	25	41
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	343	342	370
Raw Score Information			
Number of Examinees	102,616	2,405	154,910
Mean	51.58	50.34	61.24
SD	13.92	11.74	16.64
Possible Range	0-90	0-90	0-90
Obtained Range	0-90	2-90	0-90
Median	52	51	64
Test Information			
Reliability	0.90	0.86	0.94
Raw Score Standard Error of Measurement	4.34	4.32	4.10
Mean Omits	0.59	0.33	0.38
SD Omits	4.66	3.12	3.80
Percent Responding to:			
All Items	89	90	93
All Items - 1 Item	97	98	98
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	99	99
Item Information[1]			
Number of Items	72	72	72
Mean Observed Average Item Score (AIS)	0.58	0.57	0.70
Equated Mean Rasch B-Value	-0.06	-0.07	-0.04
Mean R-biserial	0.46	0.40	0.56
SD R-biserial	0.10	0.09	0.09
Examinee Information			
ELA-Only Examinees			
Number of Examinees	41,893	1,481	23,410
Mean Scale Score	345	342	339
SD Scale Score	25	21	27
Median Scale Score	345	342	339
ELA and Mathematics Examinees			
Number of Examinees	60,723	924	131,500
Mean Scale Score	344	341	376
SD Scale Score	33	31	40
Median Scale Score	342	340	377

[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 2011)

Administration	March 2011	May 2011
Scale Score Information		
Number of Examinees	402,061	39,105
Mean	381	340
SD	40	34
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	383	336
Raw Score Information		
Number of Examinees	402,061	39,105
Mean	64.60	48.64
SD	15.04	15.95
Possible Range	0-90	0-90
Obtained Range	0-90	0-90
Median	68	48
Test Information		
Reliability	0.94	0.92
Raw Score Standard Error of Measurement	3.81	4.46
Mean Omits	0.29	0.97
SD Omits	3.06	6.32
Percent Responding to:		
All Items	93	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)	0.75	0.54
Equated Mean Rasch B-Value	-0.06	-0.04
Mean R-biserial	0.55	0.51
SD R-biserial	0.09	0.09
Examinee Information		
ELA-Only Examinees		
Number of Examinees	30,545	18,920
Mean Scale Score	344	340
SD Scale Score	31	31
Median Scale Score	343	338
ELA and Mathematics Examinees		
Number of Examinees	371,516	20,185
Mean Scale Score	384	340
SD Scale Score	39	37
Median Scale Score	386	334

[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 2010)

Administration	July 2010	October 2010
Scale Score Information		
Number of Examinees	8,222	41,519
Mean	338	342
SD	22	26
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	337	339
Raw Score Information		
Number of Examinees	8,222	41,519
Mean	36.59	38.48
Standard Deviation	10.84	13.03
Possible Range	0-80	0-80
Obtained Range	1-79	0-80
Median	36	37
Test Information		
Reliability	0.86	0.90
Raw Score Standard Error of Measurement	4.12	4.10
Mean Omits	0.41	0.54
SD Omits	3.36	3.97
Percent Responding to:		
All Items	87	85
All Items - 1 Item	96	96
All Items - 2 Items	98	98
All Items - 3 Items	99	99
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)	0.45	0.47
Equated Mean Rasch B-Value	-0.24	-0.25
Mean R-biserial	0.37	0.43
SD R-biserial	0.10	0.1
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	4,717	17,222
Mean Scale Score	340	343
SD Scale Score	18	20
Median Scale Score	339	342
ELA and Mathematics Examinees		
Number of Examinees	3,505	24,297
Mean Scale Score	336	341
SD Scale Score	26	30
Median Scale Score	332	335

Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2010, February 2011)

Administration	November 2010	December 2010	February 2011
Scale Score Information			
Number of Examinees	98,156	2,433	155,492
Mean	344	340	373
SD	28	23	40
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	340	339	370
Raw Score Information			
Number of Examinees	98,156	2,433	155,492
Mean	39.68	37.85	51.56
Standard Deviation	13.37	11.50	17.30
Possible Range	0-80	0-80	0-80
Obtained Range	0-80	5-80	0-80
Median	38	37	53
Test Information			
Reliability	0.91	0.87	0.95
Raw Score Standard Error of Measurement	4.06	4.13	3.71
Mean Omits	0.49	0.29	0.31
SD Omits	3.74	2.49	2.84
Percent Responding to:			
All Items	86	89	90
All Items - 1 Item	96	97	97
All Items - 2 Items	98	98	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)	0.49	0.47	0.64
Equated Mean Rasch B-Value	-0.26	-0.25	-0.18
Mean R-biserial	0.45	0.39	0.58
SD R-biserial	0.10	0.10	0.09
Examinee Information			
Mathematics-Only Examinees			
Number of Examinees	37,433	1,509	23,992
Mean Scale Score	345	342	344
SD Scale Score	21	20	22
Median Scale Score	344	341	343
ELA and Mathematics Examinees			
Number of Examinees	60,723	924	131,500
Mean Scale Score	344	338	379
SD Scale Score	32	27	41
Median Scale Score	339	335	378

Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2011)

Administration	March 2011	May 2011
Scale Score Information		
Number of Examinees	401,838	38,851
Mean	382	344
SD	39	28
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	381	341
Raw Score Information		
Number of Examinees	401,838	38,851
Mean	55.23	38.27
Standard Deviation	16.28	13.76
Possible Range	0-80	0-80
Obtained Range	0-80	0-80
Median	58	37
Test Information		
Reliability	0.95	0.91
Raw Score Standard Error of Measurement	3.59	4.08
Mean Omits	0.24	0.68
SD Omits	2.25	4.74
Percent Responding to:		
All Items	90	86
All Items - 1 Item	98	96
All Items - 2 Items	99	98
All Items - 3 Items	99	98
All Items - 4 Items	100	98
All Items - 5 Items	100	99
Item Information		
Number of Items	80	80
Mean Observed Average Item Score (AIS)	0.69	0.47
Equated Mean Rasch B-Value	-0.18	-0.18
Mean R-biserial	0.57	0.45
SD R-biserial	0.08	0.11
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	30,322	18,666
Mean Scale Score	347	345
SD Scale Score	25	25
Median Scale Score	345	343
ELA and Mathematics Examinees		
Number of Examinees	371,516	20,185
Mean Scale Score	385	343
SD Scale Score	39	31
Median Scale Score	384	338

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 for another three years in July 2004, and again in August 2008. In 2009–10, 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 2010 and February, March, and May 2011.

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 encourages 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 2010–11 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: 2010–11

Administration	English-Language Arts (Tuesday)	Mathematics (Wednesday)
July 2010	July 27, 2010	July 28, 2010
October 2010	October 5, 2010	October 6, 2010
November 2010	November 9, 2010	November 10, 2010
December 2010	December 4, 2010[1]	December 12, 2010[1]
February 2011	February 1, 2011	February 2, 2011
March 2011	March 8, 2011	March 9, 2011
May 2011	May 10, 2011	May 11, 2011

[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 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 2010–11 School Year

There were no new or unusual changes in examination content and test administration in the 2010–11 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 was 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 ETS 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 2010 and the three 2011 administrations in the 2010–11 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 2010–11 administrations.
- Chapter 5 presents the processes involved in the actual 2010–11 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 2010–11 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.

- Chapter 11 describes the pilot study conducted to investigate alternative means by which eligible SWDs may demonstrate the same level of academic achievement in the content standards in ELA and mathematics required for passing the CAHSEE.

Each chapter contains summary tables in the body of the text. Extended appendices that give more detailed information are provided at the end 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 2010–11 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 7 field-test items for a total of 80 items. The mathematics test form consists 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 2010-11 testing cycle, one mathematics field-test form was administered for the February and March administrations due to the state budget cuts. Additional ELA item development took place to meet the requirements for operationally-ready items for the item bank. Due to the timing of item development, one field-test form was administered in February and 76 field-test forms were administered in March. 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 2010–11 Administrations

Administration	Operational Items	Field-Test Forms	Number of Field-Test Items[1]
July 2010			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
October 2010			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
November 2010			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
December 2010			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
February 2011			
English–language arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
March 2011			
English–language arts	72 MC, 1 CR	76	493
Mathematics	80 MC	1	12
May 2011			
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 2010-11 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 are 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, 2010a), 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 2010b) 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 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 13 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 modification 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, scribe, spell checker, grammar checker or an English dictionary.

Table 2.4: CAHSEE Modification/Accommodation Table: 2010–11

A. Student marks in test booklet (other than responses) -*Cannot be marked for CAHSEE*		M. Test administered at home or in hospital by test examiner. Accommodation	HH
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	N. Dictionary. ELA Modification; Math Modification	DI(E) DM(M)
C. Responses dictated [orally, or in Manually Coded English or American Sign Language] to a scribe for selected-response items (MC questions). Accommodation	OR	O. Manually Coded English or American Sign Language to present test questions. ELA Modification; Math Accommodation	SL
D. Word processing software with spell and grammar check tools turned off for the essay responses (writing portion of the test) - *Cannot be marked on the math portion of box number 13, ELA side only.* ELA Accommodation	SO	P. Test questions read aloud to student or audio presentation (CD). ELA Modification; Math Accommodation	OP
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 marked on the math portion of box 13, ELA side only.* ELA Accommodation	EO	Q. Calculators on mathematics test. - *Cannot be marked for the ELA portion of box 13, math side only* Math Modification	CA
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 marked on the math portion of box number 13, ELA side only.* ELA Accommodation	AN	R. Arithmetic table on mathematics test. - *Cannot be marked for the ELA portion of box 13, math side only.* Math Modification	AT
G. Braille transcriptions provided by the test contractor. ELA Accommodation	BV	S. Math manipulatives on mathematics test. -*Cannot be marked for the ELA portion of box 13, math side only.* Math Modification	MM
H. Large print versions. Test items enlarged if font larger than required on large print version. Accommodation	LV	T. Word processing software with spell and grammar check tools enabled on the essay responses writing portion of the test. - *Cannot be marked for the math portion of box 13, ELA side only.* ELA Modification	SC
I. Extra time on test within a testing day. -*Cannot be marked for CAHSEE*		U. Essay responses dictated orally, in Manually Coded English, or in American Sign Language to a scribe, [audio recorder, or speech-to-text converter] (scribe provides spelling, grammar, and language conventions). -*Cannot be marked for the math portion of box 13, ELA side only.* ELA Modification	ER
J. Test over more than one day for a test or test part to be administered in a single sitting. Accommodation	TD	V. Assistive device that interferes with the independent work of the student on the MC and/or essay responses. Modification	AD
K. Supervised breaks within a section of the test. Accommodation	SB	W. Unlisted Modification. Modification	UM
L. Administration of the test at the most beneficial time of day to the student. Accommodation	BT	X. Unlisted Accommodation. Accommodation	UA

Note: The shaded sections are not applicable to 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 2010–11 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 with Specific Learning Disability across the seven administrations ranged from 326 (December 2010) to 30,956 (March 2011) for ELA and 291 (December 2010) to 30,882 (March 2011) for mathematics. For students who tested with accommodations or modifications, the rates of achieving a score of 350 or higher ranged from 16 to 26 percent for ELA and from 15 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 percent 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 Braille version or tested with the large print version, 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 speakers and reclassified fluent English proficient outperformed other categories in most but not all 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 2010–11 items were calibrated and placed on the reference scale using a set of linking items selected from the 2010 forms and re-administered in 2011. 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 2010

Table 2.A.1: Summary Statistics by Testing Variations and Disability—ELA, July 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	90	328	32	19
Modifications	363	328	23	19
All	453	328	25	19
Disability				
Autism	30	325	23	20
Deaf	-	-	-	-
Deaf-Blindness	-	-	-	-
Emotional Disturbance	82	325	34	15
Hard of Hearing	17	326	20	12
Mental Retardation	19	300	20	0
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	68	331	26	26
Specific Learning Disability	926	325	24	16
Speech or Language Impairment	49	331	19	14
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 less than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.2: Summary Statistics by Testing Variations and Disability—Mathematics, July 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	84	331	24	23
Modifications	500	330	20	15
All	584	330	21	16
Disability				
Autism	26	329	20	8
Deaf	-	-	-	-
Deaf-Blindness	-	-	-	-
Emotional Disturbance	83	322	24	13
Hard of Hearing	12	332	33	25
Mental Retardation	20	306	10	0
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	66	335	20	24
Specific Learning Disability	858	328	19	13
Speech or Language Impairment	45	331	21	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 less than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.A.3: Demographic Summary for All Examinees by Testing Variations—ELA, July 2010

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	924	168	18	756	82	327	59	49	51	38	46	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 Responses (EO)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	37	8	22	29	78	333	60	54	52	39	47	2.2
Supervised Breaks (SB)	111	28	25	83	75	331	63	53	53	39	45	1.9
Beneficial Time (BT)	19	4	21	15	79	327	68	54	54	44	35	1.5
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	100	17	17	83	83	328	65	49	52	33	46	2.0
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	287	53	18	234	82	328	60	50	53	39	45	1.9
Spell Checker or Grammar Checker (SC)	27	6	22	21	78	327	53	52	54	38	44	1.9
Essay Responses (ER)	19	7	37	12	63	343	71	55	56	47	53	2.6
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	13	1	8	12	92	318	52	50	40	26	45	1.9

[1] Results for groups with less 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 2010

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	902	132	15	770	85	328	43	46	40	35	30
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)	24	2	8	22	92	326	45	51	38	31	26
Supervised Breaks (SB)	91	18	20	73	80	332	49	48	43	37	31
Beneficial Time (BT)	14	8	57	6	43	349	59	66	54	43	39
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	168	30	18	138	82	332	49	49	41	37	31
Calculator (CA)	495	76	15	419	85	330	44	49	40	36	30
Arithmetic Table (AT)	35	12	34	23	66	338	54	53	46	41	35
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	19	4	21	15	79	330	47	46	44	36	29

[1] Results for groups with less 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 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	287	310	326	342	367	390	327	25	924
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)	275	293	322	336	348	365	374	333	21	37
Supervised Breaks (SB)	275	283	312	326	350	387	413	331	30	111
Beneficial Time (BT)	275	275	306	328	348	367	367	327	26	19
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	276	294	312	329	342	361	368	328	20	100
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	275	295	310	326	342	367	387	328	24	287
Spell Checker or Grammar Checker (SC)	289	297	310	326	340	362	371	327	22	27
Essay Responses (ER)	304	304	324	340	358	402	402	343	23	19
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	280	280	310	324	328	352	352	318	17	13

[1] SD — Standard Deviation

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

Table 2.A.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, July 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	284	298	314	324	341	366	388	328	21	902
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)	284	293	314	329	342	353	359	326	20	24
Supervised Breaks (SB)	290	298	317	328	344	376	399	332	23	91
Beneficial Time (BT)	293	293	317	363	381	399	399	349	35	14
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	290	301	317	328	343	376	390	332	22	168
Calculator (CA)	284	301	317	328	341	366	393	330	20	495
Arithmetic Table (AT)	303	303	312	330	361	390	393	338	27	35
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	276	276	317	328	344	399	399	330	26	19

[1] SD — Standard Deviation

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

Table 2.A.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, July 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Dictionary	Specific Learning Disability	78	329	19	19
Essay Reponses (ER)	Specific Learning Disability	14	342	20	43
IEP or Section 504 Plan	Autism	23	323	23	17
	Emotional Disturbance	50	326	36	18
	Hard of Hearing	12	324	23	8
	Mental Retardation	14	306	19	0
	Other Health Impairment	48	333	24	29
	Specific Learning Disability	644	326	24	17
	Speech or Language Impairment	25	332	21	12
Oral Presentation	Other Health Impairment	14	329	25	7
	Specific Learning Disability	210	328	22	20
Spell Checker Or Grammar Checker	Specific Learning Disability	20	326	19	15
Supervised Breaks	Emotional Disturbance	12	333	36	25
	Specific Learning Disability	82	332	30	27
Test Over More Than One Day	Specific Learning Disability	23	335	21	26

[1] Results for groups with less 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 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	22	341	30	41
Calculator	Autism	13	333	22	15
	Emotional Disturbance	16	319	18	6
	Other Health Impairment	28	334	22	21
	Specific Learning Disability	355	330	19	14
	Speech or Language Impairment	16	330	26	19
IEP or Section 504 Plan	Autism	23	331	19	9
	Emotional Disturbance	57	319	25	7
	Mental Retardation	13	309	7	0
	Other Health Impairment	48	335	21	23
	Specific Learning Disability	625	328	20	14
	Speech or Language Impairment	22	333	24	18
Oral Presentation	Specific Learning Disability	118	332	20	17
Supervised Breaks	Emotional Disturbance	12	317	18	8
	Specific Learning Disability	65	334	22	20
Test Over More Than One Day	Specific Learning Disability	11	335	13	9
Unlisted Accommodation	Specific Learning Disability	16	332	27	25

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	English Only	16	327	28	25
Dictionary	English Only	48	327	22	17
	English Learner	47	328	19	17
Essay Reponses (ER)	English Only	13	343	22	31
IEP or Section 504 Plan	English Only	507	330	26	22
	Initially Fluent English Proficient	33	330	27	27
	Reclassified Fluency	26	336	30	31
	English Learner	349	322	23	12
Oral Presentation	English Only	159	332	25	23
	English Learner	111	323	22	13
Spell Checker Or Grammar Checker	English Only	19	331	20	21
Supervised Breaks	English Only	65	332	30	28
	English Learner	42	329	31	21
Test Over More Than One Day	English Only	27	335	22	22

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	25	337	25	28
Beneficial Time	English Only	11	346	37	55
Calculator	English Only	292	331	20	17
	Initially Fluent English Proficient	27	331	27	19
	English Learner	166	327	19	12
IEP or Section 504 Plan	English Only	526	329	21	16
	Initially Fluent English Proficient	42	328	26	17
	Reclassified Fluency	26	334	19	23
	English Learner	301	326	19	12
Oral Presentation	English Only	94	333	24	19
	English Learner	63	328	17	11
Supervised Breaks	English Only	57	333	25	25
	English Learner	26	333	19	15
Test Over More Than One Day	English Only	20	325	21	10
Unlisted Accommodation	English Only	13	332	13	15

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

[2] SD — Standard Deviation

Appendix 2.B: Results of Testing Variations and Disability Analyses—October 2010

Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,479	326	25	19
Modifications	2,903	330	25	21
All	4,382	328	25	21
Disability				
Unknown	-	-	-	-
Autism	258	322	24	16
Deaf	136	314	21	4
Deaf-Blindness	-	-	-	-
Emotional Disturbance	718	328	32	25
Hard of Hearing	124	326	23	15
Mental Retardation	315	307	20	3
Multiple Disability	26	323	27	23
Orthopedic Impairment	93	325	19	10
Other Health Impairment	649	332	26	23
Specific Learning Disability	8,129	328	23	19
Speech or Language Impairment	428	328	23	20
Traumatic Brain Injury	38	324	24	11
Visual Impairment	36	336	32	31

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

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

[2] SD — Standard Deviation

Table 2.B.2: Summary Statistics by Testing Variations and Disability—Mathematics, October 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,197	326	20	12
Modifications	4,496	331	20	16
All	5,693	330	20	15
Disability				
Unknown	-	-	-	-
Autism	257	328	21	15
Deaf	108	327	19	10
Deaf-Blindness	-	-	-	-
Emotional Disturbance	778	328	25	17
Hard of Hearing	96	330	19	13
Mental Retardation	301	315	17	3
Multiple Disability	23	321	20	9
Orthopedic Impairment	98	330	19	15
Other Health Impairment	720	331	23	17
Specific Learning Disability	7,752	329	20	14
Speech or Language Impairment	387	329	20	14
Traumatic Brain Injury	37	333	23	11
Visual Impairment	40	332	27	15

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

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

[2] SD — Standard Deviation

Table 2.B.3: Demographic Summary for All Examinees by Testing Variations—ELA, October 2010

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	8,823	1,673	19	7,150	81	327	59	48	49	43	45	1.9
Transfer of T/B Responses to A/D (TS)	60	17	28	43	72	334	59	51	53	47	50	2.0
Oral Responses Dictated to a Scribe (OR)	30	7	23	23	77	332	57	52	53	45	49	2.0
Spell Checker/Grammar Checker Off (SO)	58	16	28	42	72	340	66	57	56	49	50	2.2
Essay Responses (EO)	57	10	18	47	82	325	57	47	46	40	44	1.9
Assistive Device No Interference (AN)	44	15	34	29	66	341	74	57	56	45	57	2.1
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	21	4	19	17	81	323	52	47	47	42	43	1.6
Test Over More Than One Day (TD)	412	105	25	307	75	329	60	49	50	43	47	1.9
Supervised Breaks (SB)	1,761	383	22	1,378	78	328	59	48	50	43	45	1.9
Beneficial Time (BT)	328	62	19	266	81	324	56	46	47	39	43	1.8
Tested At Home or Hospital (HH)	21	9	43	12	57	340	70	56	54	50	57	1.9
Dictionary (DI)	1,015	221	22	794	78	329	61	49	51	44	46	1.9
Sign Language (SL)	57	2	4	55	96	318	50	44	41	37	45	1.6
Oral Presentation (OP)	2,313	547	24	1,766	76	332	62	50	53	46	47	1.9
Spell Checker or Grammar Checker (SC)	157	53	34	104	66	337	65	54	56	47	50	2.1
Essay Responses (ER)	45	14	31	31	69	337	61	50	54	43	50	2.4
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	18	7	39	11	61	341	65	56	54	55	53	2.1
Unlisted Accommodation (UA)	299	62	21	237	79	327	60	48	50	42	44	1.9
Writing Only (WO)	142	25	18	117	82	328	57	47	49	44	46	2.0

[1] Results for groups with less 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 2010

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	8,761	1,263	14	7,498	86	329	43	48	41	34	32
Transfer of T/B Responses to A/D (TS)	53	10	19	43	81	334	50	52	43	35	35
Oral Responses Dictated to a Scribe (OR)	15	3	20	12	80	333	47	48	41	43	35
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	20	1	5	19	95	321	35	42	35	32	29
Test Over More Than One Day (TD)	274	55	20	219	80	332	45	52	43	35	32
Supervised Breaks (SB)	1,493	224	15	1,269	85	328	43	47	40	34	32
Beneficial Time (BT)	289	34	12	255	88	325	41	43	39	32	30
Tested At Home or Hospital (HH)	13	5	38	8	62	345	54	59	47	49	39
Dictionary for Math (DM)	233	52	22	181	78	335	49	56	43	37	33
Sign Language (SL)	52	6	12	46	88	334	44	53	48	34	34
Oral Presentation (OP)	1,194	233	20	961	80	333	46	53	42	37	33
Calculator (CA)	4,416	704	16	3,712	84	331	43	52	41	35	32
Arithmetic Table (AT)	430	94	22	336	78	335	47	56	43	38	34
Math Manipulative (MM)	71	26	37	45	63	341	54	57	48	42	36
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	33	11	33	22	67	344	52	62	48	39	45
Unlisted Accommodation (UA)	283	36	13	247	87	327	40	46	39	34	32

[1] Results for groups with less 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 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	275	287	310	327	344	367	387	327	24	8,823
Transfer of T/B Responses to A/D (TS)	287	294	315	329	350	380	390	334	25	60
Oral Responses Dictated to a Scribe (OR)	275	279	312	332	346	379	402	332	29	30
Spell Checker/Grammar Checker Off (SO)	289	297	321	340	352	399	433	340	28	58
Essay Reponses (EO)	281	291	310	323	336	362	369	325	21	57
Assistive Device No Interference (AN)	293	299	329	339	354	374	428	341	24	44
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	275	277	306	317	340	381	396	323	32	21
Test Over More Than One Day (TD)	275	279	312	329	350	369	384	329	27	412
Supervised Breaks (SB)	275	287	308	327	346	371	387	328	26	1,761
Beneficial Time (BT)	275	279	304	323	344	367	384	324	27	328
Tested At Home Or Hospital (HH)	275	304	315	334	362	390	423	340	36	21
Dictionary (DI)	275	289	312	331	346	371	387	329	25	1,015
Sign Language (SL)	275	285	306	315	332	348	410	318	22	57
Oral Presentation (OP)	277	293	314	332	348	374	390	332	25	2,313
Spell Checker Or Grammar Checker (SC)	279	291	321	340	354	376	390	337	24	157
Essay Reponses (ER)	295	301	323	332	352	376	393	337	22	45
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	297	297	319	336	374	396	396	341	33	18
Unlisted Accommodation (UA)	275	283	308	327	344	369	387	327	25	299
Writing Only (WO)	275	291	312	327	344	367	396	328	24	142

[1] SD — Standard Deviation

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

Table 2.B.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, October 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	288	301	315	326	341	364	387	329	20	8,761
Transfer of T/B Responses to A/D (TS)	293	299	319	330	347	375	395	334	21	53
Oral Responses Dictated to a Scribe (OR)	290	290	323	332	342	380	380	333	24	15
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	291	292	308	323	332	353	359	321	19	20
Test Over More Than One Day (TD)	277	296	315	330	346	373	400	332	24	274
Supervised Breaks (SB)	284	299	313	326	341	367	397	328	22	1,493
Beneficial Time (BT)	275	293	310	321	337	364	395	325	22	289
Tested At Home or Hospital (HH)	296	296	321	346	364	430	430	345	36	13
Dictionary for Math (DM)	284	304	319	332	347	373	400	335	23	233
Sign Language (SL)	296	306	323	335	342	367	392	334	17	52
Oral Presentation (OP)	291	304	317	330	346	373	415	333	23	1,194
Calculator (CA)	288	301	317	328	342	366	389	331	20	4,416
Arithmetic Table (AT)	299	308	321	332	347	369	387	335	20	430
Math Manipulatives (MM)	299	306	326	339	360	371	415	341	22	71
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	306	308	317	337	364	424	430	344	32	33
Unlisted Accommodation (UA)	291	301	313	324	337	360	369	327	18	283

[1] SD — Standard Deviation

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

Table 2.B.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, October 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	Specific Learning Disability	30	342	25	33
Beneficial Time	Autism	12	330	17	8
	Emotional Disturbance	65	319	28	15
	Mental Retardation	18	302	21	0
	Other Health Impairment	14	329	33	21
	Specific Learning Disability	174	329	25	24
Dictionary	Autism	22	322	23	14
	Emotional Disturbance	21	333	26	33
	Mental Retardation	39	315	20	5
	Other Health Impairment	43	333	26	28
	Specific Learning Disability	759	330	25	23
	Speech or Language Impairment	31	324	24	16
Essay Reponses (EO)	Specific Learning Disability	22	318	18	14
Essay Reponses (ER)	Specific Learning Disability	25	343	20	36
IEP or Section 504 Plan	Autism	197	323	24	17
	Deaf	72	316	24	7
	Emotional Disturbance	531	327	31	24
	Hard of Hearing	89	325	23	15
	Mental Retardation	256	308	20	4
	Multiple Disability	22	323	27	23
	Orthopedic Impairment	73	324	18	8
	Other Health Impairment	471	332	26	25
	Specific Learning Disability	5,952	328	23	19
	Speech or Language Impairment	293	327	23	18
	Traumatic Brain Injury	32	324	24	9
	Visual Impairment	29	338	33	34
Large Print Version	Visual Impairment	11	329	32	18
Oral Presentation	Autism	47	320	24	11
	Emotional Disturbance	71	331	31	25
	Hard of Hearing	19	337	19	26
	Mental Retardation	88	314	21	6
	Orthopedic Impairment	32	327	20	13
	Other Health Impairment	101	336	27	34
	Specific Learning Disability	1,693	333	24	25
	Speech or Language Impairment	87	331	22	21
Oral Responses Dictated to a Scribe	Specific Learning Disability	13	316	25	0
Sign Language	Deaf	34	316	24	6
	Hard of Hearing	18	322	18	0
Spell Checker Or Grammar Checker	Other Health Impairment	11	341	28	45
	Specific Learning Disability	118	337	24	33
Spell Checker Or Grammar Checker Off	Specific Learning Disability	24	336	27	25
Supervised Breaks	Autism	43	319	27	12
	Emotional Disturbance	153	324	30	22
	Hard of Hearing	17	316	26	12
	Mental Retardation	86	311	22	6

Table 2.B.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Orthopedic Impairment	14	331	18	14
	Other Health Impairment	109	334	27	30
	Specific Learning Disability	1,104	329	25	22
	Speech or Language Impairment	62	330	27	31
Test Over More Than One Day	Autism	14	330	25	29
	Emotional Disturbance	35	322	29	20
	Mental Retardation	13	313	24	15
	Other Health Impairment	21	331	24	19
	Specific Learning Disability	263	330	27	27
	Speech or Language Impairment	23	325	21	22
Transfer of Student T/B Responses to A/D	Specific Learning Disability	28	339	25	36
Unlisted Accommodation	Emotional Disturbance	19	319	26	11
	Mental Retardation	13	301	18	0
	Other Health Impairment	22	334	27	32
	Specific Learning Disability	205	328	25	22
Writing ONLY	Specific Learning Disability	97	330	24	20

[1] Results for groups with less 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 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Emotional Disturbance	15	352	31	47
	Other Health Impairment	29	339	17	28
	Specific Learning Disability	337	334	19	22
	Speech or Language Impairment	13	332	19	15
Beneficial Time	Emotional Disturbance	62	320	23	6
	Mental Retardation	16	313	15	6
	Other Health Impairment	15	326	26	20
	Specific Learning Disability	143	329	21	14
Calculator	Autism	110	329	19	15
	Deaf	17	337	23	18
	Emotional Disturbance	199	331	24	20
	Hard of Hearing	21	335	21	14
	Mental Retardation	133	318	17	5
	Multiple Disability	11	330	22	18
	Orthopedic Impairment	48	333	21	21
	Other Health Impairment	270	332	21	17
	Specific Learning Disability	3,018	331	20	17
	Speech or Language Impairment	159	330	18	9
	Traumatic Brain Injury	12	330	16	0
	Visual Impairment	13	334	24	15
Dictionary for Math	Specific Learning Disability	183	335	22	23
IEP or Section 504 Plan	Autism	201	327	20	14
	Deaf	58	331	21	16
	Emotional Disturbance	583	328	24	16
	Hard of Hearing	70	332	19	11
	Mental Retardation	238	316	17	4
	Multiple Disability	21	322	20	10
	Orthopedic Impairment	80	330	19	15
	Other Health Impairment	529	331	21	16
	Specific Learning Disability	5,796	329	20	15
	Speech or Language Impairment	270	329	18	12
	Traumatic Brain Injury	32	333	22	9
	Visual Impairment	29	332	28	17
Math Manipulatives	Specific Learning Disability	52	342	22	38
Oral Presentation	Autism	22	326	20	18
	Emotional Disturbance	47	347	29	36
	Mental Retardation	48	323	17	8
	Orthopedic Impairment	14	332	16	21
	Other Health Impairment	60	335	25	27
	Specific Learning Disability	831	334	22	20
	Speech or Language Impairment	39	333	20	13

Table 2.B.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Sign Language	Deaf	24	337	20	21
	Hard of Hearing	14	331	18	7
	Specific Learning Disability	11	334	10	0
Supervised Breaks	Autism	39	324	20	13
	Emotional Disturbance	138	325	25	14
	Hard of Hearing	13	336	29	23
	Mental Retardation	68	317	18	9
	Orthopedic Impairment	14	339	22	21
	Other Health Impairment	102	329	22	19
	Specific Learning Disability	932	330	22	16
	Speech or Language Impairment	44	327	21	14
Test Over More Than One Day	Emotional Disturbance	37	326	22	8
	Other Health Impairment	21	329	25	19
	Specific Learning Disability	168	334	24	23
	Speech or Language Impairment	11	337	26	18
Transfer of Student T/B Responses to A/D	Specific Learning Disability	20	337	19	25
Unlisted Accommodation	Emotional Disturbance	23	329	23	22
	Mental Retardation	11	314	12	0
	Other Health Impairment	20	329	22	20
	Specific Learning Disability	184	326	16	11
Unlisted Modification	Specific Learning Disability	25	347	34	40

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	27	341	29	37
Beneficial Time	English Only	217	323	27	18
	English Learner	95	325	25	21
Dictionary	English Only	518	331	27	25
	Initially Fluent English Proficient	21	320	25	19
	Reclassified Fluent English Proficient	36	341	26	39
	English Learner	417	326	21	17
	Unknown	23	326	31	22
Essay Reponses (EO)	English Only	41	325	21	17
Essay Reponses (ER)	English Only	31	340	23	39
IEP or Section 504 Plan	English Only	5,086	329	25	21
	Initially Fluent English Proficient	194	328	24	18
	Reclassified Fluent English Proficient	271	336	26	31
	English Learner	3,143	325	23	14
	Unknown	129	328	26	22
Large Print Version	English Only	12	332	31	25
Oral Presentation	English Only	1,294	334	26	26
	Initially Fluent English Proficient	33	333	25	21
	Reclassified Fluent English Proficient	81	338	25	35
	English Learner	842	329	22	18
	Unknown	63	329	29	27
Oral Responses Dictated to a Scribe	English Only	19	332	34	32
Sign Language	English Only	29	321	18	3
	English Learner	25	310	17	0
Spell Checker Or Grammar Checker	English Only	95	336	27	36
	English Learner	45	337	21	31
Spell Checker Or Grammar Checker Off	English Only	39	341	28	28
	English Learner	15	336	26	20
Supervised Breaks	English Only	1,071	329	27	23
	Initially Fluent English Proficient	35	328	25	20
	Reclassified Fluent English Proficient	81	335	28	32
	English Learner	531	325	24	17
	Unknown	43	330	24	26
Test Over More Than One Day	English Only	270	328	27	25
	Reclassified Fluent English Proficient	15	336	30	40
	English Learner	112	329	26	26
Tested At Home Or Hospital	English Only	12	333	33	33
Transfer of Student T/B Responses to A/D	English Only	46	332	26	24
Unlisted Accommodation	English Only	193	329	25	22
	English Learner	90	320	22	12
Writing ONLY	English Only	77	330	26	22
	English Learner	54	323	20	9

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	272	336	19	24
	Initially Fluent English Proficient	11	334	19	27
	Reclassified Fluent English Proficient	12	345	21	33
	English Learner	121	334	19	18
	Unknown	14	316	17	0
Beneficial Time	English Only	190	325	23	13
	English Learner	83	324	18	8
Calculator	English Only	2,746	331	21	18
	Initially Fluent English Proficient	96	331	19	17
	Reclassified Fluent English Proficient	143	336	20	21
	English Learner	1,353	329	18	12
	Unknown	78	324	17	5
Dictionary for Math	English Only	140	335	24	24
	English Learner	81	336	22	21
IEP or Section 504 Plan	English Only	5,332	329	21	16
	Initially Fluent English Proficient	216	332	21	19
	Reclassified Fluent English Proficient	298	335	19	19
	English Learner	2,763	327	19	11
	Unknown	152	326	20	11
Large Print Version	English Only	12	319	18	0
Math Manipulatives	English Only	42	342	23	40
	Unknown	15	348	14	40
Oral Presentation	English Only	709	334	25	23
	Initially Fluent English Proficient	13	339	18	15
	Reclassified Fluent English Proficient	36	337	18	17
	English Learner	398	331	19	15
	Unknown	38	332	21	18
Oral Responses Dictated to a Scribe	English Only	13	336	24	23
Sign Language	English Only	22	333	21	9
	English Learner	25	334	14	12
Supervised Breaks	English Only	965	329	23	16
	Initially Fluent English Proficient	39	326	18	10
	Reclassified Fluent English Proficient	54	337	24	22
	English Learner	392	327	20	12
	Unknown	43	332	21	19
Test Over More Than One Day	English Only	194	331	24	20
	English Learner	57	333	22	23
	Unknown	11	323	25	9
Transfer of Student T/B Responses to A/D	English Only	45	334	22	20
Unlisted Accommodation	English Only	197	329	18	15
	English Learner	75	321	16	4
Unlisted Modification	English Only	25	347	34	36

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

[2] SD — Standard Deviation

Appendix 2.C: Results of Testing Variations and Disability Analyses—November 2010

Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,488	327	27	19
Modifications	5,834	329	24	19
All	9,322	328	25	19
Disability				
Autism	668	330	29	23
Deaf	262	311	22	6
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,660	330	33	26
Hard of Hearing	272	328	22	17
Mental Retardation	609	305	20	2
Multiple Disability	40	322	27	18
Orthopedic Impairment	185	329	28	23
Other Health Impairment	1,831	334	27	27
Specific Learning Disability	19,075	328	24	18
Speech or Language Impairment	1,234	332	24	22
Traumatic Brain Injury	107	324	22	14
Visual Impairment	83	335	33	28

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

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

[2] SD — Standard Deviation

Table 2.C.2: Summary Statistics by Testing Variations and Disability—Mathematics, November 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,012	327	22	15
Modifications	9,360	332	21	18
All	12,372	331	21	17
Disability				
Autism	674	331	23	19
Deaf	221	325	18	7
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,859	329	26	18
Hard of Hearing	236	332	21	18
Mental Retardation	616	313	17	2
Multiple Disability	34	320	16	6
Orthopedic Impairment	209	329	21	13
Other Health Impairment	1,954	332	23	20
Specific Learning Disability	18,434	329	21	15
Speech or Language Impairment	1,111	332	21	18
Traumatic Brain Injury	98	328	19	13
Visual Impairment	88	333	30	26

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

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

[2] SD — Standard Deviation

Table 2.C.3: Demographic Summary for All Examinees by Testing Variations—ELA, November 2010

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	20,902	3,871	19	17,031	81	328	58	50	53	43	41	1.9
Transfer of T/B Responses to A/D (TS)	104	35	34	69	66	338	63	54	58	53	46	2.0
Oral Responses Dictated to a Scribe (OR)	74	20	27	54	73	334	60	55	57	47	43	1.9
Spell Checker/Grammar Checker Off (SO)	183	51	28	132	72	335	62	54	56	48	45	2.0
Essay Responses (EO)	81	23	28	58	72	335	62	54	56	50	43	2.0
Assistive Device No Interference (AN)	49	12	24	37	76	335	69	53	57	51	46	1.9
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	47	11	23	36	77	334	59	50	56	48	47	1.9
Test Over More Than One Day (TD)	980	221	23	759	77	330	59	50	54	44	42	1.9
Supervised Breaks (SB)	3,770	690	18	3,080	82	327	57	48	52	42	41	1.9
Beneficial Time (BT)	450	75	17	375	83	325	56	47	51	40	39	1.8
Tested At Home or Hospital (HH)	27	9	33	18	67	334	59	51	58	53	43	1.8
Dictionary (DI)	1,723	300	17	1,423	83	328	62	50	53	42	41	1.9
Sign Language (SL)	93	11	12	82	88	318	46	40	49	39	37	1.8
Oral Presentation (OP)	4,995	956	19	4,039	81	330	60	51	55	45	42	1.9
Spell Checker or Grammar Checker (SC)	452	107	24	345	76	333	61	52	57	46	42	2.1
Essay Responses (ER)	76	17	22	59	78	331	60	50	53	46	38	2.2
Assistive Device (AD)	26	9	35	17	65	337	63	55	59	48	40	2.3
Unlisted Modification (UM)	20	5	25	15	75	335	56	55	58	46	43	2.0
Unlisted Accommodation (UA)	434	80	18	354	82	328	60	50	53	44	42	1.9
Writing Only(WO)	184	19	10	165	90	322	52	45	50	40	39	1.8

[1] Results for groups with less 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 2010

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	20,902	3,325	16	17,577	84	329	43	44	43	38	32
Transfer of T/B Responses to A/D (TS)	95	25	26	70	74	336	46	49	49	40	36
Oral Responses Dictated to a Scribe (OR)	38	11	29	27	71	337	46	54	46	41	36
Braille Version (BV)	18	3	17	15	83	327	40	47	38	35	32
Large Print Version (LV)	53	11	21	42	79	335	46	49	47	40	35
Test Over More Than One Day (TD)	544	102	19	442	81	331	44	45	45	38	32
Supervised Breaks (SB)	3,080	454	15	2,626	85	327	42	42	42	36	31
Beneficial Time (BT)	385	49	13	336	87	324	39	40	41	33	29
Tested At Home or Hospital (HH)	29	15	52	14	48	346	55	53	55	48	40
Dictionary for Math (DM)	278	50	18	228	82	333	44	49	44	40	32
Sign Language (SL)	107	8	7	99	93	327	39	43	41	35	34
Oral Presentation (OP)	2,909	523	18	2,386	82	331	43	46	44	38	33
Calculator (CA)	9,234	1,662	18	7,572	82	332	44	49	44	39	32
Arithmetic Table (AT)	661	112	17	549	83	332	45	48	43	39	32
Math Manipulative (MM)	60	9	15	51	85	332	41	48	45	41	31
Assistive Device (AD)	22	7	32	15	68	342	47	58	50	44	40
Unlisted Modification (UM)	92	21	23	71	77	333	46	50	43	38	35
Unlisted Accommodation (UA)	506	70	14	436	86	329	43	44	44	37	30

[1] Results for groups with less 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 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	288	309	328	345	369	389	328	25	20,902
Transfer of T/B Responses to A/D (TS)	284	290	313	336	358	399	450	338	35	104
Oral Responses Dictated to a Scribe (OR)	275	286	311	336	351	399	422	334	33	74
Spell Checker/Grammar Checker Off (SO)	282	296	313	332	354	376	418	335	28	183
Essay Responses (EO)	282	292	311	336	351	384	409	335	30	81
Assistive Device No Interference (AN)	282	293	317	336	349	369	427	335	25	49
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	282	290	311	330	347	402	446	334	33	47
Test Over More Than One Day (TD)	275	288	311	330	349	374	389	330	26	980
Supervised Breaks (SB)	275	286	307	326	345	369	389	327	26	3,770
Beneficial Time (BT)	275	282	303	324	343	381	402	325	29	450
Tested At Home or Hospital (HH)	275	288	299	324	364	399	413	334	39	27
Dictionary (DI)	275	292	313	328	343	367	384	328	23	1,723
Sign Language (SL)	275	279	299	313	334	367	402	318	27	93
Oral Presentation (OP)	277	292	313	330	345	369	387	330	23	4,995
Spell Checker or Grammar Checker (SC)	284	296	315	332	349	374	387	333	23	452
Essay Responses (ER)	288	294	312	328	345	379	399	331	25	76
Assistive Device (AD)	309	309	320	329	356	374	384	337	22	26
Unlisted Modification (UM)	292	296	309	333	352	397	409	335	31	20
Unlisted Accommodation (UA)	275	288	309	328	345	374	389	328	26	434
Writing Only (WO)	275	286	303	320	340	364	406	322	25	184

[1] SD — Standard Deviation

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

Table 2.C.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, November 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	286	298	314	328	342	368	392	329	22	20,902
Transfer of T/B Responses to A/D (TS)	279	303	316	335	351	398	445	336	28	95
Oral Responses Dictated to a Scribe (OR)	298	300	320	330	355	387	398	337	26	38
Braille Version (BV)	295	295	303	325	339	402	402	327	27	18
Large Print Version (LV)	289	298	322	331	344	401	420	335	26	53
Test Over More Than One Day (TD)	283	298	314	330	344	370	401	331	23	544
Supervised Breaks (SB)	283	295	312	324	340	366	385	327	22	3,080
Beneficial Time (BT)	275	292	310	320	337	360	385	324	22	385
Tested At Home or Hospital (HH)	276	279	312	351	372	395	450	346	39	29
Dictionary for Math (DM)	292	300	320	330	344	376	401	333	22	278
Sign Language (SL)	286	300	314	326	335	366	408	327	21	107
Oral Presentation (OP)	289	300	314	328	344	370	395	331	22	2,909
Calculator (CA)	289	303	316	330	344	368	392	332	21	9,234
Arithmetic Table (AT)	289	300	316	330	344	368	401	332	22	661
Math Manipulatives (MM)	295	303	316	329	346	371	401	332	22	60
Assistive Device (AD)	307	310	330	344	355	370	374	342	18	22
Unlisted Modification (UM)	283	295	318	331	347	378	385	333	24	92
Unlisted Accommodation (UA)	292	300	314	328	340	370	390	329	21	506

[1] SD — Standard Deviation

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

Table 2.C.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, November 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	16	338	23	31
Assistive Device No Interference	Specific Learning Disability	23	343	27	39
Beneficial Time	Emotional Disturbance	108	321	29	14
	Other Health Impairment	30	334	28	23
	Specific Learning Disability	222	324	26	14
	Speech or Language Impairment	19	329	24	21
Dictionary	Autism	33	332	22	21
	Deaf	37	315	21	5
	Emotional Disturbance	43	336	26	37
	Hard of Hearing	17	330	24	18
	Mental Retardation	46	308	17	0
	Orthopedic Impairment	12	334	29	33
	Other Health Impairment	72	336	25	25
	Specific Learning Disability	1,291	329	22	17
	Speech or Language Impairment	64	326	25	22
Essay Reponses (EO)	Orthopedic Impairment	11	333	28	36
	Specific Learning Disability	36	328	28	19
Essay Reponses (ER)	Specific Learning Disability	42	331	24	21
IEP or Section 504 Plan	Autism	509	330	29	23
	Deaf	168	310	22	6
	Emotional Disturbance	1,271	329	33	25
	Hard of Hearing	187	330	22	18
	Mental Retardation	507	304	20	2
	Multiple Disability	29	322	28	21
	Orthopedic Impairment	137	328	28	22
	Other Health Impairment	1,340	333	26	25
	Specific Learning Disability	14,161	328	24	17
	Speech or Language Impairment	820	329	24	18
	Traumatic Brain Injury	84	324	23	13
	Visual Impairment	70	337	35	31
Large Print Version	Visual Impairment	21	336	42	24
Oral Presentation	Autism	128	328	27	21
	Emotional Disturbance	131	326	28	18
	Hard of Hearing	44	334	23	25
	Mental Retardation	177	309	19	2
	Multiple Disability	12	325	29	25
	Orthopedic Impairment	45	336	28	29
	Other Health Impairment	249	334	24	22
	Specific Learning Disability	3,573	331	23	19
	Speech or Language Impairment	233	329	22	19
	Traumatic Brain Injury	26	324	20	12
	Visual Impairment	18	343	27	33
Oral Responses Dictated to a Scribe	Orthopedic Impairment	13	340	28	31
	Specific Learning Disability	31	335	30	29

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Sign Language	Deaf	57	313	25	7
	Hard of Hearing	23	324	22	9
Spell Checker Or Grammar Checker	Autism	18	339	24	33
	Deaf	14	304	10	0
	Emotional Disturbance	15	334	22	20
	Mental Retardation	15	310	18	7
	Other Health Impairment	32	340	26	34
	Specific Learning Disability	272	335	21	24
	Speech or Language Impairment	45	333	25	27
Spell Checker Or Grammar Checker Off	Autism	12	337	33	42
	Emotional Disturbance	12	354	22	50
	Mental Retardation	11	311	21	0
	Other Health Impairment	14	320	20	0
	Specific Learning Disability	103	335	24	27
Supervised Breaks	Autism	106	328	29	25
	Emotional Disturbance	316	325	31	20
	Hard of Hearing	29	331	26	21
	Mental Retardation	110	301	17	2
	Orthopedic Impairment	21	324	36	29
	Other Health Impairment	227	331	26	23
	Specific Learning Disability	2,520	327	24	17
	Speech or Language Impairment	144	328	24	17
	Traumatic Brain Injury	13	319	23	15
Test Over More Than One Day	Autism	23	331	30	26
	Deaf	31	306	17	3
	Emotional Disturbance	66	327	30	24
	Hard of Hearing	18	331	28	17
	Mental Retardation	21	305	22	5
	Other Health Impairment	55	334	27	31
	Specific Learning Disability	667	330	25	22
	Speech or Language Impairment	33	335	26	27
Transfer of Student T/B Responses to A/D	Other Health Impairment	14	336	33	43
	Specific Learning Disability	39	335	25	28
	Visual Impairment	13	344	50	31
Unlisted Accommodation	Autism	14	328	22	21
	Emotional Disturbance	19	334	31	32
	Other Health Impairment	26	340	37	35
	Specific Learning Disability	293	326	24	13
	Speech or Language Impairment	11	320	16	0
Unlisted Modification	Specific Learning Disability	11	336	28	27
Writing ONLY	Deaf	46	304	14	0
	Hard of Hearing	12	319	18	0

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Specific Learning Disability	87	329	25	14

[1] Results for groups with less 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 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Autism	15	326	17	7
	Emotional Disturbance	19	331	30	16
	Mental Retardation	24	316	16	0
	Other Health Impairment	43	334	23	19
	Specific Learning Disability	488	333	21	18
	Speech or Language Impairment	21	326	19	10
Assistive Device	Specific Learning Disability	17	340	19	29
Beneficial Time	Emotional Disturbance	103	321	25	14
	Other Health Impairment	25	325	27	16
	Specific Learning Disability	177	323	19	10
	Speech or Language Impairment	19	320	17	0
Calculator	Autism	264	333	23	21
	Deaf	46	323	12	2
	Emotional Disturbance	411	332	24	21
	Hard of Hearing	65	336	24	23
	Mental Retardation	232	316	17	4
	Multiple Disability	14	324	18	7
	Orthopedic Impairment	88	330	21	15
	Other Health Impairment	639	334	22	22
	Specific Learning Disability	6,309	332	21	18
	Speech or Language Impairment	402	332	20	17
	Traumatic Brain Injury	35	328	16	11
	Visual Impairment	23	336	32	26
Dictionary for Math	Other Health Impairment	11	322	18	0
	Specific Learning Disability	214	333	21	19
	Speech or Language Impairment	24	327	20	13
IEP or Section 504 Plan	Autism	515	331	23	19
	Deaf	146	325	19	5
	Emotional Disturbance	1,444	329	26	19
	Hard of Hearing	168	333	21	17
	Mental Retardation	515	313	16	2
	Multiple Disability	27	320	16	4
	Orthopedic Impairment	162	330	21	14
	Other Health Impairment	1,449	332	23	21
	Specific Learning Disability	13,870	329	21	15
	Speech or Language Impairment	768	330	21	16
	Traumatic Brain Injury	79	327	19	11
	Visual Impairment	73	335	30	26
Large Print Version	Visual Impairment	25	342	32	32
Math Manipulatives	Specific Learning Disability	33	335	24	24

Table 2.C.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Autism	79	331	24	20
	Emotional Disturbance	99	330	26	21
	Hard of Hearing	31	339	23	29
	Mental Retardation	120	313	16	2
	Orthopedic Impairment	38	330	24	16
	Other Health Impairment	153	331	21	18
	Specific Learning Disability	1,926	331	22	19
	Speech or Language Impairment	153	332	21	17
	Traumatic Brain Injury	15	321	10	0
Oral Responses Dictated to a Scribe	Orthopedic Impairment	12	346	20	33
Sign Language	Deaf	69	327	21	6
	Hard of Hearing	13	326	15	8
	Specific Learning Disability	16	327	25	13
Supervised Breaks	Autism	76	329	25	13
	Emotional Disturbance	291	325	24	16
	Hard of Hearing	27	332	18	11
	Mental Retardation	84	309	13	1
	Orthopedic Impairment	17	328	21	24
	Other Health Impairment	222	331	24	19
	Specific Learning Disability	1,996	328	21	14
	Speech or Language Impairment	106	325	22	14
	Traumatic Brain Injury	13	324	18	15
	Visual Impairment	12	324	26	17
Test Over More Than One Day	Autism	12	342	21	33
	Emotional Disturbance	49	328	33	22
	Mental Retardation	12	314	19	0
	Other Health Impairment	39	330	27	18
	Specific Learning Disability	371	331	22	19
	Speech or Language Impairment	15	323	18	7
Transfer of Student T/B Responses to A/D	Specific Learning Disability	30	331	20	23
	Visual Impairment	21	347	34	38
Unlisted Accommodation	Autism	19	335	17	16
	Emotional Disturbance	22	328	22	14
	Other Health Impairment	29	330	23	17
	Specific Learning Disability	347	329	21	14
Unlisted Modification	Specific Learning Disability	72	335	25	28

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	17	337	23	35
Assistive Device No Interference	English Only	39	337	26	26
Beneficial Time	English Only	284	327	29	19
	English Learner	143	319	26	9
	Unknown	11	327	36	27
Dictionary	English Only	831	331	24	22
	Initially Fluent English Proficient	41	330	24	17
	Reclassified Fluent English Proficient	59	336	25	32
	English Learner	778	324	21	11
	Unknown	14	345	32	43
Essay Reponses(EO)	English Only	48	338	31	38
	English Learner	24	329	27	13
Essay Reponses(ER)	English Only	43	336	28	35
	English Learner	27	320	18	4
IEP or Section 504 Plan	English Only	11,747	330	27	22
	Initially Fluent English Proficient	538	330	26	21
	Reclassified Fluent English Proficient	778	336	26	30
	English Learner	7,647	324	23	12
	Unknown	192	331	29	24
Large Print Version	English Only	28	338	37	29
	English Learner	15	323	20	7
Oral Presentation	English Only	2,626	332	24	23
	Initially Fluent English Proficient	115	335	22	23
	Reclassified Fluent English Proficient	177	335	24	27
	English Learner	2,034	326	22	14
	Unknown	43	334	32	30
Oral Responses Dictated to a Scribe	English Only	44	340	35	36
	English Learner	23	324	25	9
Sign Language	English Only	48	323	24	10
	English Learner	39	308	22	8
Spell Checker Or Grammar Checker	English Only	272	335	23	26
	Initially Fluent English Proficient	16	332	17	13
	Reclassified Fluent English Proficient	11	345	25	55
	English Learner	142	327	22	17
	Unknown	11	344	30	45
Spell Checker Or Grammar Checker Off	English Only	129	340	28	36
	English Learner	43	320	20	7

Table 2.C.9 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Supervised Breaks	English Only	2,074	329	27	22
	Initially Fluent English Proficient	97	326	24	13
	Reclassified Fluent English Proficient	158	331	28	24
	English Learner	1,397	322	23	12
	Unknown	44	332	29	27
Test Over More Than One Day	English Only	493	332	28	26
	Initially Fluent English Proficient	30	331	27	20
	Reclassified Fluent English Proficient	27	332	30	26
	English Learner	418	327	24	18
	Unknown	12	326	27	17
Tested At Home Or Hospital	English Only	21	334	44	38
Transfer of Student T/B Responses to A/D	English Only	77	340	35	35
	English Learner	23	326	27	26
Unlisted Accommodation	English Only	251	333	26	22
	Initially Fluent English Proficient	16	332	32	19
	Reclassified Fluent English Proficient	23	340	23	35
	English Learner	141	317	22	9
Unlisted Modification	English Only	16	335	33	25
Writing ONLY	English Only	82	327	28	16
	English Learner	91	317	22	5

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	425	333	23	19
	Initially Fluent English Proficient	20	339	20	30
	Reclassified Fluent English Proficient	23	336	20	22
	English Learner	187	328	19	10
Assistive Device	English Only	18	342	19	39
Beneficial Time	English Only	252	323	23	14
	English Learner	118	325	19	11
Braille Version	English Only	12	327	31	17
Calculator	English Only	5,529	333	22	20
	Initially Fluent English Proficient	240	335	20	21
	Reclassified Fluent English Proficient	366	338	21	24
	English Learner	3,032	329	20	14
	Unknown	67	329	22	18
Dictionary for Math	English Only	166	334	23	18
	English Learner	97	331	20	19
IEP or Section 504 Plan	English Only	12,517	330	22	17
	Initially Fluent English Proficient	562	331	23	17
	Reclassified Fluent English Proficient	796	337	21	23
	English Learner	6,838	327	20	12
	Unknown	189	329	25	16
Large Print Version	English Only	39	337	29	26
Math Manipulatives	English Only	42	335	22	17
Oral Presentation	English Only	1,619	332	23	20
	Initially Fluent English Proficient	63	333	24	16
	Reclassified Fluent English Proficient	94	335	23	22
	English Learner	1,113	328	20	14
	Unknown	20	335	21	25
Oral Responses Dictated to a Scribe	English Only	18	333	26	28
	English Learner	11	333	23	27
Sign Language	English Only	65	326	20	8
	English Learner	35	324	14	3
Supervised Breaks	English Only	1,838	328	23	16
	Initially Fluent English Proficient	79	326	20	9
	Reclassified Fluent English Proficient	112	331	18	17
	English Learner	1,023	326	20	12
	Unknown	28	323	23	14
Test Over More Than One Day	English Only	308	332	26	23

Table 2.C.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Initially Fluent English Proficient	21	332	20	14
	Reclassified Fluent English Proficient	11	334	13	0
	English Learner	194	328	19	14
Tested At Home Or Hospital	English Only	23	349	42	57
Transfer of Student T/B Responses to A/D	English Only	72	336	28	26
	English Learner	16	329	17	19
Unlisted Accommodation	English Only	333	330	21	14
	Initially Fluent English Proficient	11	320	13	0
	Reclassified Fluent English Proficient	18	331	19	17
	English Learner	143	328	21	14
Unlisted Modification	English Only	67	330	24	18
	English Learner	15	330	18	20

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

[2] SD — Standard Deviation

Appendix 2.D: Results of Testing Variations and Disability Analyses—December 2010

Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	36	326	24	14
Modifications	126	334	26	28
All	162	332	25	25
Disability				
Autism	-	-	-	-
Emotional Disturbance	-	-	-	-
Hard of Hearing	-	-	-	-
Mental Retardation	-	-	-	-
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	26	341	23	35
Specific Learning Disability	326	331	24	22
Speech or Language Impairment	15	344	18	40

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

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

[2] SD — Standard Deviation

Table 2.D.2: Summary Statistics by Testing Variations and Disability—Mathematics, December 2010

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	38	328	25	13
Modifications	136	337	26	24
All	174	335	26	22
Disability				
Autism	-	-	-	-
Emotional Disturbance	-	-	-	-
Hard of Hearing	-	-	-	-
Mental Retardation	-	-	-	-
Multiple Disability	-	-	-	-
Orthopedic Impairment	-	-	-	-
Other Health Impairment	35	333	20	20
Specific Learning Disability	291	333	23	19
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 less than 11 students are not reported.

[2] SD — Standard Deviation

Table 2.D.3: Demographic Summary for All Examinees by Testing Variations—ELA, December 2010

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	287	63	22	224	78	331	57	50	52	40	52	2.0
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)	15	4	27	11	73	327	55	43	51	33	44	2.4
Supervised Breaks (SB)	45	9	20	36	80	332	60	53	52	39	51	2.1
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-	-	-
Tested At Home Or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	43	10	23	33	77	328	62	50	48	41	46	2.1
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	119	34	29	85	71	335	62	55	54	44	52	2.0
Spell Checker Or Grammar Checker (SC)	42	11	26	31	74	336	61	56	55	46	56	1.9
Essay Reponses (ER)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-	-	-
Writing Only (WO)	12	1	8	11	92	320	54	49	43	29	40	2.0

[1] Results for groups with less 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 2010

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	277	50	18	227	82	332	48	50	43	36	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)	-	-	-	-	-	-	-	-	-	-	-
Supervised Breaks (SB)	36	9	25	27	75	334	49	52	44	37	32
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-	-
Tested At Home Or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	56	18	32	38	68	340	50	59	49	41	35
Calculator (CA)	135	33	24	102	76	337	49	57	45	38	34
Arithmetic Table (AT)	19	0	0	19	100	328	43	52	37	31	32
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-	-

[1] Results for groups with less 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 2010

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	314	332	348	372	396	331	26	287
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	314	334	352	356	356	327	27	15
Supervised Breaks (SB)	277	292	318	334	346	367	377	332	23	45
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-
Tested At Home Or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	279	290	310	320	348	385	396	328	29	43
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	277	283	318	334	350	372	393	335	25	119
Spell Checker Or Grammar Checker (SC)	275	279	328	339	350	367	387	336	26	42
Essay Reponses (ER)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-
Writing Only (WO)	290	290	299	323	332	352	352	320	21	12

[1] SD — Standard Deviation

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

Table 2.D.6: Percentiles of Scale Scores for Students with Testing Variations— Mathematics, December 2010

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 Plan	284	301	316	330	344	385	412	332	25	277
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)	277	284	317	330	350	395	398	334	28	36
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-
Tested At Home Or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	277	293	321	334	357	398	412	340	30	56
Calculator (CA)	290	303	320	333	349	395	412	337	26	135
Arithmetic Table (AT)	301	301	320	332	335	349	349	328	12	19
Math Manipulative (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	-	-	-	-	-	-	-	-	-	-

[1] SD — Standard Deviation

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

Table 2.D.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, December 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Dictionary	Specific Learning Disability	41	328	29	22
IEP or Section 504 Plan	Other Health Impairment	21	341	24	29
	Specific Learning Disability	220	330	25	21
Oral Presentation	Specific Learning Disability	99	336	25	29
Spell Checker Or Grammar Checker	Specific Learning Disability	34	339	24	29
Supervised Breaks	Specific Learning Disability	33	330	23	15

[1] Results for groups with less 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 2010

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	12	328	10	0
Calculator	Other Health Impairment	16	327	17	13
	Specific Learning Disability	99	339	28	24
IEP or Section 504 Plan	Other Health Impairment	27	330	19	11
	Specific Learning Disability	198	332	26	18
Oral Presentation	Specific Learning Disability	38	346	32	39
Supervised Breaks	Specific Learning Disability	27	339	28	33

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Dictionary	English Only	12	336	32	33
	English Learner	29	326	28	21
IEP or Section 504 Plan	English Only	155	334	27	25
	English Learner	119	327	25	18
Oral Presentation	English Only	54	335	26	30
	English Learner	59	333	25	27
Spell Checker Or Grammar Checker	English Only	21	332	27	24
	English Learner	17	336	23	24
Supervised Breaks	English Only	29	336	20	24
	English Learner	14	327	25	14
Test Over More Than One Day	English Only	12	333	25	33

[1] Results for groups with less 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 2010

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	17	329	12	0
Calculator	English Only	80	338	25	24
	English Learner	46	336	27	28
IEP or Section 504 Plan	English Only	178	334	25	19
	English Learner	86	329	25	17
Oral Presentation	English Only	37	342	29	32
	English Learner	19	338	32	32
Supervised Breaks	English Only	32	334	29	25

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

[2] SD — Standard Deviation

Appendix 2.E: Results of Testing Variations and Disability Analyses—February 2011

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

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3,081	328	31	22
Modifications	3,382	326	25	18
All	6,463	327	28	20
Disability				
Unknown	-	-	-	-
Autism	673	347	41	42
Deaf	204	318	31	12
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,593	333	38	31
Hard of Hearing	212	336	36	31
Mental Retardation	516	303	21	3
Multiple Disability	42	316	31	14
Orthopedic Impairment	179	332	34	26
Other Health Impairment	1,721	341	33	40
Specific Learning Disability	15,221	328	27	21
Speech or Language Impairment	1,169	338	34	30
Traumatic Brain Injury	90	321	25	10
Visual Impairment	83	356	46	54

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

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

[2] SD — Standard Deviation

Table 2.E.2: Summary Statistics by Testing Variations and Disability—Mathematics, February 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	2,861	333	27	23
Modifications	5,728	331	21	18
All	8,589	332	23	19
Disability				
Unknown	-	-	-	-
Autism	689	351	39	42
Deaf	178	335	25	26
Deaf-Blindness	-	-	-	-
Emotional Disturbance	1,698	333	31	25
Hard of Hearing	200	347	35	35
Mental Retardation	515	315	17	4
Multiple Disability	49	325	20	16
Orthopedic Impairment	196	337	33	27
Other Health Impairment	1,870	341	30	33
Specific Learning Disability	15,255	334	25	23
Speech or Language Impairment	1,150	344	34	33
Traumatic Brain Injury	88	327	19	13
Visual Impairment	86	358	42	49

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

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

[2] SD — Standard Deviation

Table 2.E.3: Demographic Summary for All Examinees by Testing Variations—ELA, February 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	16,338	3,622	22	12,716	78	329	51	48	51	43	47	1.8
Transfer of T/B Responses to A/D (TS)	140	60	43	80	57	344	62	58	59	53	56	2.0
Oral Responses Dictated to a Scribe (OR)	74	21	28	53	72	333	57	50	56	49	43	1.9
Spell Checker/Grammar Checker Off (SO)	130	52	40	78	60	349	63	60	60	52	59	2.2
Essay Reponses (EO)	87	26	30	61	70	335	54	52	54	46	48	2.0
Assistive Device No Interference (AN)	64	17	27	47	73	335	56	52	55	47	48	2.0
Braille Version (BV)	11	8	73	3	27	365	73	68	67	62	69	2.3
Large Print Version (LV)	73	27	37	46	63	340	57	55	57	51	54	1.9
Test Over More Than One Day (TD)	614	115	19	499	81	325	50	47	48	40	43	1.8
Supervised Breaks (SB)	2,781	545	20	2,236	80	325	50	46	49	42	45	1.8
Beneficial Time (BT)	534	98	18	436	82	325	49	47	50	40	45	1.8
Tested At Home Or Hospital (HH)	30	11	37	19	63	337	54	54	58	51	52	1.8
Dictionary (DI)	1,138	185	16	953	84	326	55	47	48	42	45	1.9
Sign Language (SL)	58	7	12	51	88	315	47	43	43	30	42	1.5
Oral Presentation (OP)	2,745	476	17	2,269	83	327	53	48	49	43	45	1.9
Spell Checker Or Grammar Checker (SC)	212	39	18	173	82	327	52	49	49	39	45	1.9
Essay Reponses (ER)	77	18	23	59	77	330	57	48	49	44	46	2.1
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	18	4	22	14	78	322	45	45	47	36	50	1.7
Unlisted Accommodation (UA)	481	91	19	390	81	328	51	48	51	43	47	1.8
Writing Only (WO)	185	33	18	152	82	325	50	46	49	39	44	1.8

[1] Results for groups with less 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 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	16,867	3,803	23	13,064	77	334	49	42	41	40	35
Transfer of T/B Responses to A/D (TS)	118	52	44	66	56	351	60	51	54	51	43
Oral Responses Dictated to a Scribe (OR)	35	17	49	18	51	350	56	56	54	47	46
Braille Version (BV)	13	4	31	9	69	340	44	56	45	47	34
Large Print Version (LV))	61	22	36	39	64	348	56	51	52	44	45
Test Over More Than One Day (TD)	413	109	26	304	74	334	49	43	42	40	35
Supervised Breaks (SB)	2,615	545	21	2,070	79	331	47	40	40	38	35
Beneficial Time (BT)	485	97	20	388	80	330	48	40	38	37	33
Tested At Home Or Hospital (HH)	19	5	26	14	74	334	54	39	41	41	35
Dictionary for Math (DM)	234	41	18	193	82	333	51	41	39	39	39
Sign Language (SL)	48	7	15	41	85	328	41	38	36	35	38
Oral Presentation (OP)	1,850	416	22	1,434	78	334	49	43	41	40	36
Calculator (CA)	5,658	984	17	4,674	83	331	47	42	39	38	34
Arithmetic Table (AT)	535	109	20	426	80	334	51	43	41	40	36
Math Manipulative (MM)	80	28	35	52	65	340	51	48	46	45	38
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	32	9	28	23	72	338	48	46	44	41	41
Unlisted Accommodation (UA)	507	100	20	407	80	332	47	41	40	39	35

[1] Results for groups with less 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 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	325	346	382	416	329	30	16,338
Transfer of T/B Responses to A/D (TS)	275	290	318	344	368	404	420	344	35	140
Oral Responses Dictated to a Scribe (OR)	275	281	307	327	355	400	450	333	36	74
Spell Checker/Grammar Checker Off (SO)	283	299	323	343	372	420	450	349	37	130
Essay Reponses (EO)	275	287	303	329	359	400	416	335	35	87
Assistive Device No Interference (AN)	275	291	310	336	352	391	450	335	32	64
Braille Version (BV)	294	294	332	365	379	450	450	365	51	11
Large Print Version (LV)	277	281	314	337	361	411	442	340	37	73
Test Over More Than One Day (TD)	275	277	303	323	344	370	400	325	28	614
Supervised Breaks (SB)	275	283	303	323	344	377	400	325	29	2,781
Beneficial Time (BT)	275	283	305	323	344	372	397	325	27	534
Tested At Home Or Hospital (HH)	283	293	310	343	357	411	416	337	34	30
Dictionary (DI)	277	289	308	325	343	370	388	326	24	1,138
Sign Language (SL)	277	279	297	310	329	368	400	315	27	58
Oral Presentation (OP)	277	289	308	325	343	370	391	327	25	2,745
Spell Checker Or Grammar Checker (SC)	277	289	308	325	344	368	380	327	24	212
Essay Reponses (ER)	279	283	308	329	346	380	391	330	28	77
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	283	283	299	324	343	372	372	322	27	18
Unlisted Accommodation (UA)	275	285	307	325	344	377	400	328	28	481
Writing Only (WO)	277	287	303	323	343	377	400	325	28	185

[1] SD — Standard Deviation

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

Table 2.E.6: Percentiles of Scale Scores for Students with Testing Variations — Mathematics, February 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	300	316	330	346	385	418	334	26	16,867
Transfer of T/B Responses to A/D (TS)	298	308	322	344	368	418	450	351	36	118
Oral Responses Dictated to a Scribe (OR)	292	292	310	348	370	422	439	350	38	35
Braille Version (BV)	298	298	310	332	351	405	405	340	36	13
Large Print Version (LV)	282	303	318	336	364	433	450	348	41	61
Test Over More Than One Day (TD)	275	295	316	334	350	378	410	334	26	413
Supervised Breaks (SB)	279	300	314	328	344	376	407	331	24	2,615
Beneficial Time (BT)	275	298	314	328	344	368	387	330	23	485
Tested At Home Or Hospital (HH)	289	289	312	328	352	422	422	334	30	19
Dictionary for Math (DM)	298	305	321	332	343	364	392	333	19	234
Sign Language (SL)	300	305	316	322	337	362	372	328	17	48
Oral Presentation (OP)	292	303	316	330	346	383	407	334	24	1,850
Calculator (CA)	289	303	316	330	343	368	397	331	21	5,658
Arithmetic Table (AT)	289	305	321	334	346	370	403	334	22	535
Math Manipulatives (MM)	300	308	322	337	357	371	439	340	25	80
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	295	303	317	337	350	390	418	338	27	32
Unlisted Accommodation (UA)	295	300	316	330	344	374	410	332	24	507

[1] SD — Standard Deviation

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

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

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	Specific Learning Disability	40	327	27	18
Beneficial Time	Emotional Disturbance	124	321	30	19
	Other Health Impairment	30	328	26	20
	Specific Learning Disability	286	325	24	15
	Speech or Language Impairment	13	336	34	38
Dictionary	Autism	23	321	23	4
	Emotional Disturbance	38	336	27	32
	Mental Retardation	36	312	23	8
	Other Health Impairment	43	331	22	28
	Specific Learning Disability	843	326	24	16
	Speech or Language Impairment	37	318	23	8
Essay Reponses (EO)	Specific Learning Disability	40	331	31	23
Essay Reponses (ER)	Specific Learning Disability	38	330	26	24
IEP or Section 504 Plan	Autism	470	342	38	36
	Deaf	106	316	30	9
	Emotional Disturbance	1,184	332	37	30
	Hard of Hearing	138	330	35	22
	Mental Retardation	392	303	21	4
	Multiple Disability	36	314	32	14
	Orthopedic Impairment	127	330	31	20
	Other Health Impairment	1,117	337	32	35
	Specific Learning Disability	10,700	326	27	18
	Speech or Language Impairment	671	328	26	19
	Traumatic Brain Injury	75	320	24	9
	Visual Impairment	64	350	44	48
Large Print Version	Specific Learning Disability	11	330	37	27
	Visual Impairment	27	343	40	41
Oral Presentation	Autism	64	328	28	19
	Emotional Disturbance	105	328	30	24
	Hard of Hearing	21	329	30	24
	Mental Retardation	93	313	22	6
	Orthopedic Impairment	29	319	20	3
	Other Health Impairment	150	330	25	25
	Specific Learning Disability	2,008	327	24	17
	Speech or Language Impairment	116	328	23	14
	Traumatic Brain Injury	13	320	24	15
Oral Responses Dictated to a Scribe	Specific Learning Disability	28	328	31	25
Sign Language	Deaf	38	312	22	8
	Hard of Hearing	13	314	19	0
Spell Checker Or Grammar Checker	Deaf	13	302	13	0
	Other Health Impairment	12	324	25	25
	Specific Learning Disability	122	331	23	21
	Speech or Language Impairment	18	324	23	11
	Other Health Impairment	18	362	39	67

Table 2.E.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Spell Checker Or Grammar Checker Off	Emotional Disturbance	13	342	36	31
	Specific Learning Disability	53	338	29	28
Supervised Breaks	Autism	95	338	38	33
	Deaf	16	321	28	13
	Emotional Disturbance	318	324	33	21
	Hard of Hearing	21	333	43	33
	Mental Retardation	70	301	23	6
	Orthopedic Impairment	18	333	34	22
	Other Health Impairment	212	335	29	33
	Specific Learning Disability	1,717	325	27	17
	Speech or Language Impairment	101	321	23	14
	Traumatic Brain Injury	12	325	20	8
	Visual Impairment	16	333	39	44
Test Over More Than One Day	Autism	18	342	34	28
	Deaf	13	302	13	0
	Emotional Disturbance	57	319	31	16
	Mental Retardation	14	294	17	0
	Other Health Impairment	52	333	31	27
	Specific Learning Disability	366	324	26	18
	Speech or Language Impairment	28	325	26	18
Transfer of Student T/B Responses to A/D	Orthopedic Impairment	12	345	42	25
	Other Health Impairment	16	361	33	75
	Specific Learning Disability	42	343	32	40
	Visual Impairment	23	354	38	57
Unlisted Accommodation	Autism	14	338	33	29
	Emotional Disturbance	37	334	34	27
	Mental Retardation	11	302	18	0
	Other Health Impairment	34	336	35	29
	Specific Learning Disability	313	327	25	17
	Speech or Language Impairment	13	323	29	23
Unlisted Modification	Specific Learning Disability	11	322	31	27
Writing ONLY	Deaf	23	310	20	0
	Hard of Hearing	11	313	19	0
	Other Health Impairment	12	352	35	58
	Specific Learning Disability	98	325	23	14

[1] Results for groups with less 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 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Emotional Disturbance	20	331	31	25
	Other Health Impairment	49	335	22	22
	Specific Learning Disability	375	333	20	18
	Speech or Language Impairment	19	331	21	16
Beneficial Time	Autism	11	338	24	45
	Emotional Disturbance	110	325	25	18
	Other Health Impairment	36	333	20	17
	Specific Learning Disability	251	331	22	22
	Speech or Language Impairment	16	342	25	25
Calculator	Autism	154	333	24	21
	Deaf	23	329	21	22
	Emotional Disturbance	248	331	26	20
	Hard of Hearing	45	331	13	13
	Mental Retardation	155	319	20	8
	Multiple Disability	21	330	19	19
	Orthopedic Impairment	58	334	23	22
	Other Health Impairment	399	334	23	19
	Specific Learning Disability	3,894	331	21	17
	Speech or Language Impairment	272	334	20	19
	Traumatic Brain Injury	30	329	16	13
	Visual Impairment	18	346	41	33
Dictionary for Math	Specific Learning Disability	169	333	17	17
	Speech or Language Impairment	12	320	18	0
IEP or Section 504 Plan	Autism	487	346	37	35
	Deaf	89	331	23	22
	Emotional Disturbance	1,249	333	30	24
	Hard of Hearing	138	341	30	26
	Mental Retardation	394	315	17	4
	Multiple Disability	41	326	20	17
	Orthopedic Impairment	145	336	31	26
	Other Health Impairment	1,238	338	28	28
	Specific Learning Disability	10,899	332	23	20
	Speech or Language Impairment	686	335	25	22
	Traumatic Brain Injury	73	327	19	14
	Visual Impairment	69	352	40	43
Large Print Version	Visual Impairment	27	351	40	41
Math Manipulatives	Specific Learning Disability	44	338	26	34

Table 2.E.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Autism	46	336	25	24
	Emotional Disturbance	81	339	28	33
	Hard of Hearing	21	341	29	24
	Mental Retardation	68	323	23	13
	Orthopedic Impairment	23	330	28	26
	Other Health Impairment	120	334	22	20
	Specific Learning Disability	1,261	334	24	22
	Speech or Language Impairment	86	337	23	26
Sign Language	Deaf	23	328	19	17
Supervised Breaks	Autism	81	345	36	41
	Emotional Disturbance	287	325	24	15
	Hard of Hearing	21	345	31	33
	Mental Retardation	60	313	16	3
	Orthopedic Impairment	20	336	35	25
	Other Health Impairment	200	335	24	23
	Specific Learning Disability	1,619	331	23	21
	Speech or Language Impairment	102	331	23	23
	Traumatic Brain Injury	12	335	15	17
	Visual Impairment	18	341	36	28
Test Over More Than One Day	Autism	17	335	24	41
	Emotional Disturbance	57	324	29	19
	Other Health Impairment	42	342	30	36
	Specific Learning Disability	222	334	23	25
	Speech or Language Impairment	24	341	24	38
Transfer of Student T/B Responses to A/D	Orthopedic Impairment	11	347	45	36
	Other Health Impairment	12	378	40	75
	Specific Learning Disability	43	347	29	42
	Visual Impairment	23	353	43	43
Unlisted Accommodation	Autism	15	340	35	40
	Emotional Disturbance	38	328	21	11
	Other Health Impairment	47	341	30	32
	Specific Learning Disability	313	330	20	17
	Speech or Language Impairment	16	330	21	19
Unlisted Modification	Specific Learning Disability	22	332	22	23

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	45	342	33	33
	English Learner	11	318	23	9
Beneficial Time	English Only	352	326	29	21
	Initially Fluent English Proficient	14	334	34	43
	English Learner	149	322	20	9
	Unknown	13	320	33	15
Dictionary	English Only	541	327	25	20
	Initially Fluent English Proficient	21	337	26	43
	Reclassified Fluent English Proficient	36	336	25	25
	English Learner	517	323	22	11
	Unknown	23	323	29	17
Essay Reponses (EO)	English Only	61	335	34	30
	English Learner	14	315	29	14
Essay Reponses (ER)	English Only	42	337	27	31
	English Learner	28	316	26	14
IEP or Section 504 Plan	English Only	10,095	332	32	27
	Initially Fluent English Proficient	424	333	32	28
	Reclassified Fluent English Proficient	557	341	30	38
	English Learner	5,084	320	23	10
	Unknown	178	327	31	23
Large Print Version	English Only	51	343	36	39
	English Learner	17	326	28	24
Oral Presentation	English Only	1,446	330	26	21
	Initially Fluent English Proficient	56	331	27	25
	Reclassified Fluent English Proficient	82	334	22	23
	English Learner	1,142	323	22	12
	Unknown	19	319	21	11
Oral Responses Dictated to a Scribe	English Only	49	338	36	31
	English Learner	17	312	27	18
Sign Language	English Only	32	317	24	6
	English Learner	24	311	29	17
Spell Checker Or Grammar Checker	English Only	119	330	26	25
	English Learner	75	319	19	4
Spell Checker Or Grammar Checker Off	English Only	94	354	41	49
	English Learner	26	332	18	8
Supervised Breaks	English Only	1,805	327	31	23
	Initially Fluent English Proficient	67	331	29	25
	Reclassified Fluent English Proficient	82	335	32	37
	English Learner	814	320	23	10
	Unknown	13	318	30	23

Table 2.E.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Test Over More Than One Day	English Only	387	325	30	20
	Initially Fluent English Proficient	14	316	27	14
	Reclassified Fluent English Proficient	22	329	28	36
	English Learner	190	324	24	14
Tested At Home Or Hospital	English Only	21	343	35	48
Transfer of Student T/B Responses to A/D	English Only	103	346	35	48
	English Learner	28	326	28	14
Unlisted Accommodation	English Only	313	331	28	22
	Initially Fluent English Proficient	13	341	32	46
	Reclassified Fluent English Proficient	14	333	18	7
	English Learner	140	319	25	11
Unlisted Modification	English Only	14	319	26	14
Writing ONLY	English Only	106	327	29	22
	English Learner	69	320	25	10

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	342	334	22	23
	Initially Fluent English Proficient	11	332	18	27
	Reclassified Fluent English Proficient	21	345	24	29
	English Learner	149	334	20	15
	Unknown	12	325	15	8
Beneficial Time	English Only	320	331	23	22
	Initially Fluent English Proficient	11	327	23	18
	Reclassified Fluent English Proficient	12	338	21	33
	English Learner	134	328	20	16
Calculator	English Only	3,355	332	22	19
	Initially Fluent English Proficient	151	336	23	25
	Reclassified Fluent English Proficient	191	337	23	25
	English Learner	1,876	329	19	14
	Unknown	85	333	23	25
Dictionary for Math	English Only	131	336	21	24
	English Learner	86	330	15	9
IEP or Section 504 Plan	English Only	10,800	335	28	25
	Initially Fluent English Proficient	459	338	29	28
	Reclassified Fluent English Proficient	630	345	29	37
	English Learner	4,790	329	20	15
	Unknown	188	333	25	23
Large Print Version	English Only	45	347	42	36
	English Learner	12	340	29	25
Math Manipulatives	English Only	68	340	24	35
Oral Presentation	English Only	1,072	335	26	25
	Initially Fluent English Proficient	45	343	28	36
	Reclassified Fluent English Proficient	44	341	28	27
	English Learner	668	331	21	18
	Unknown	21	334	28	24
Oral Responses Dictated to a Scribe	English Only	27	350	38	48
Sign Language	English Only	28	328	18	18
	English Learner	17	326	18	12
Supervised Breaks	English Only	1,743	332	25	22
	Initially Fluent English Proficient	64	331	27	25
	Reclassified Fluent English Proficient	90	339	26	32
	English Learner	704	329	21	17
	Unknown	14	332	36	36
Test Over More Than One Day	English Only	284	333	27	25
	English Learner	112	337	23	30
Tested At Home Or Hospital	English Only	12	326	20	8

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	86	350	35	45
	English Learner	19	338	31	26
Unlisted Accommodation	English Only	371	333	25	22
	English Learner	118	327	22	14
Unlisted Modification	English Only	27	339	28	33

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

[2] SD — Standard Deviation

Appendix 2.F: Results of Testing Variations and Disability Analyses—March 2011

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

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	6,506	336	35	32
Modifications	5,002	327	26	18
All	11,508	332	32	26
Disability				
Unknown	-	-	-	-
Autism	1,758	354	43	50
Deaf	389	312	31	9
Deaf-Blindness	12	343	43	33
Emotional Disturbance	2,925	338	42	37
Hard of Hearing	547	346	39	43
Mental Retardation	744	304	21	3
Multiple Disability	65	320	32	17
Orthopedic Impairment	389	348	41	45
Other Health Impairment	4,114	347	36	46
Specific Learning Disability	30,956	332	30	27
Speech or Language Impairment	2,374	349	39	44
Traumatic Brain Injury	133	325	29	20
Visual Impairment	182	362	41	60

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

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

[2] SD — Standard Deviation

Table 2.F.2: Summary Statistics by Testing Variations and Disability—Mathematics, March 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	5,636	338	30	30
Modifications	8,943	335	23	24
All	14,579	336	26	26
Disability				
Unknown	-	-	-	-
Autism	1,761	359	42	50
Deaf	350	332	29	18
Deaf-Blindness	11	346	37	27
Emotional Disturbance	3,053	338	33	31
Hard of Hearing	522	353	35	48
Mental Retardation	749	315	19	4
Multiple Disability	60	326	29	17
Orthopedic Impairment	417	348	35	42
Other Health Impairment	4,221	346	31	41
Specific Learning Disability	30,882	338	26	29
Speech or Language Impairment	2,369	356	38	50
Traumatic Brain Injury	133	335	26	28
Visual Impairment	192	358	39	54

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

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

[2] SD — Standard Deviation

Table 2.F.3: Demographic Summary for All Examinees by Testing Variations—ELA, March 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	31,795	9,111	29	22,684	71	334	55	51	57	47	48	1.8
Transfer of T/B Responses to A/D (TS)	193	94	49	99	51	351	64	61	65	56	57	2.0
Oral Responses Dictated to a Scribe (OR)	114	40	35	74	65	336	60	52	56	49	49	1.8
Spell Checker/Grammar Checker Off (SO)	261	154	59	107	41	362	69	64	70	61	61	2.2
Essay Reponses (EO)	119	55	46	64	54	349	63	59	64	56	55	2.0
Assistive Device No Interference (AN)	124	54	44	70	56	345	62	56	62	51	55	2.1
Braille Version (BV)	33	20	61	13	39	373	72	68	72	61	72	2.3
Large Print Version (LV)	134	69	51	65	49	353	65	62	65	57	60	2.0
Test Over More Than One Day (TD)	1,680	552	33	1,128	67	336	57	52	58	49	49	1.9
Supervised Breaks (SB)	5,733	1,439	25	4,294	75	330	52	48	55	44	46	1.8
Beneficial Time (BT)	958	265	28	693	72	329	52	48	55	44	45	1.7
Tested At Home Or Hospital (HH)	118	72	61	46	39	363	69	65	72	61	63	2.1
Dictionary (DI)	1,610	242	15	1,368	85	325	52	46	52	41	42	1.8
Sign Language (SL)	104	17	16	87	84	325	51	45	48	42	43	1.7
Oral Presentation (OP)	3,977	686	17	3,291	83	327	53	47	53	43	44	1.8
Spell Checker Or Grammar Checker (SC)	422	72	17	350	83	328	53	47	53	43	44	1.9
Essay Reponses (ER)	123	30	24	93	76	329	54	48	52	45	44	2.0
Assistive Device (AD)	23	6	26	17	74	343	68	53	61	52	52	2.1
Unlisted Modification (UM)	34	15	44	19	56	342	58	56	61	53	54	1.9
Unlisted Accommodation (UA)	723	194	27	529	73	333	53	50	57	46	47	1.9
Writing Only(WO)	392	94	24	298	76	331	52	49	56	46	46	1.8

[1] Results for groups with less 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 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	32,062	9,486	30	22,576	70	339	49	52	45	40	34
Transfer of T/B Responses to A/D (TS)	167	73	44	94	56	352	58	59	53	47	43
Oral Responses Dictated to a Scribe (OR)	108	48	44	60	56	347	55	57	51	45	40
Braille Version (BV))	18	10	56	8	44	360	61	65	62	48	44
Large Print Version (LV))	119	62	52	57	48	356	60	61	56	49	45
Test Over More Than One Day (TD)	1,227	438	36	789	64	341	52	53	47	41	36
Supervised Breaks (SB))	5,079	1,348	27	3,731	73	335	48	49	43	38	33
Beneficial Time (BT)	938	255	27	683	73	333	46	48	42	38	32
Tested At Home Or Hospital (HH)	70	34	49	36	51	353	61	59	54	49	42
Dictionary for Math (DM)	207	43	21	164	79	333	47	53	40	36	29
Sign Language (SL)	149	23	15	126	85	331	41	48	40	34	35
Oral Presentation (OP)	2,766	646	23	2,120	77	335	47	52	42	38	32
Calculator (CA)	8,834	2,069	23	6,765	77	335	46	54	41	38	31
Arithmetic Table (AT)	537	109	20	428	80	334	47	53	39	36	31
Math Manipulatives (MM)	138	63	46	75	54	350	56	62	54	47	40
Assistive Device (AD)	13	7	54	6	46	350	59	65	47	45	41
Unlisted Modification (UM)	47	12	26	35	74	336	48	53	40	37	37
Unlisted Accommodation (UA)	661	170	26	491	74	336	47	50	44	39	34

[1] Results for groups with less 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 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	287	309	331	354	392	427	334	33	31,795
Transfer of T/B Responses to A/D (TS)	282	295	319	349	378	422	450	351	40	193
Oral Responses Dictated to a Scribe (OR)	275	282	303	324	365	406	450	336	41	114
Spell Checker/Grammar Checker Off (SO)	284	295	329	358	392	445	450	362	43	261
Essay Reponses (EO)	280	295	313	343	383	422	450	349	42	119
Assistive Device No Interference (AN)	284	297	315	341	369	414	445	345	36	124
Braille Version (BV)	275	282	331	389	414	450	450	373	51	33
Large Print Version (LV)	275	289	319	356	383	433	445	353	43	134
Test Over More Than One Day (TD)	275	284	311	335	361	392	422	336	33	1,680
Supervised Breaks (SB)	275	284	307	327	352	386	414	330	31	5,733
Beneficial Time (BT)	275	275	303	324	356	389	433	329	35	958
Tested At Home Or Hospital (HH)	275	282	331	367	399	445	450	363	46	118
Dictionary (DI)	275	287	307	323	341	368	383	325	24	1,610
Sign Language (SL)	280	287	301	313	337	410	445	325	37	104
Oral Presentation (OP)	275	289	307	325	343	370	389	327	25	3,977
Spell Checker Or Grammar Checker (SC)	278	289	309	325	343	373	418	328	27	422
Essay Reponses (ER)	284	291	309	327	347	378	399	329	28	123
Assistive Device (AD)	309	309	323	335	361	402	422	343	28	23
Unlisted Modification (UM)	291	293	315	338	370	402	433	342	35	34
Unlisted Accommodation (UA)	275	289	309	331	352	389	427	333	32	723
Writing Only (WO)	275	287	309	327	349	378	422	331	30	392

[1] SD — Standard Deviation

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

Table 2.F.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, March 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	318	334	354	393	429	339	29	32,062
Transfer of T/B Responses to A/D (TS)	291	305	322	345	373	429	450	352	37	167
Oral Responses Dictated to a Scribe (OR)	302	307	320	345	367	408	450	347	32	108
Braille Version (BV))	307	307	316	364	384	450	450	360	46	18
Large Print Version (LV))	294	302	326	352	381	429	450	356	38	119
Test Over More Than One Day (TD)	277	300	318	337	359	396	429	341	31	1,227
Supervised Breaks (SB))	281	300	316	330	350	386	424	335	27	5,079
Beneficial Time (BT)	275	291	314	328	350	391	434	333	31	938
Tested At Home Or Hospital (HH)	275	294	320	348	381	429	450	353	41	70
Dictionary for Math (DM)	297	307	318	330	347	365	391	333	20	207
Sign Language (SL)	288	302	316	326	343	371	408	331	23	149
Oral Presentation (OP)	291	302	318	332	348	379	399	335	24	2,766
Calculator (CA)	291	305	320	332	348	375	402	335	22	8,834
Arithmetic Table (AT)	291	302	318	330	345	373	402	334	22	537
Math Manipulative (MM)	297	309	332	347	367	402	419	350	28	138
Assistive Device (AD)	302	302	326	350	357	424	424	350	33	13
Unlisted Modification (UM)	285	302	316	330	352	375	399	336	25	47
Unlisted Accommodation (UA)	294	302	316	332	350	386	419	336	26	661

[1] SD — Standard Deviation

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

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

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	17	341	26	29
Assistive Device No Interference	Specific Learning Disability	72	336	29	35
Beneficial Time	Autism	28	325	36	18
	Emotional Disturbance	222	325	40	28
	Hard of Hearing	13	339	47	46
	Mental Retardation	17	296	23	12
	Other Health Impairment	60	334	40	35
	Specific Learning Disability	507	328	30	24
	Speech or Language Impairment	26	328	34	19
Braille Version	Visual Impairment	15	362	54	53
Dictionary	Autism	41	326	29	17
	Deaf	25	309	19	0
	Emotional Disturbance	38	327	32	26
	Hard of Hearing	16	329	19	13
	Mental Retardation	37	303	16	0
	Other Health Impairment	82	333	27	27
	Specific Learning Disability	1,233	325	24	15
	Speech or Language Impairment	43	323	20	12
Essay Reponses (EO)	Orthopedic Impairment	16	355	46	50
	Other Health Impairment	13	353	39	46
	Specific Learning Disability	32	334	30	31
Essay Reponses (ER)	Specific Learning Disability	69	330	26	25
IEP or Section 504 Plan	Autism	1,110	349	42	45
	Deaf	297	310	26	7
	Emotional Disturbance	2,072	334	40	34
	Hard of Hearing	338	340	37	37
	Mental Retardation	557	304	20	3
	Multiple Disability	47	320	32	13
	Orthopedic Impairment	272	344	39	40
	Other Health Impairment	2,724	345	36	43
	Specific Learning Disability	20,652	330	29	24
	Speech or Language Impairment	1,137	334	32	27
	Traumatic Brain Injury	99	322	28	19
	Visual Impairment	144	357	40	56
Large Print Version	Specific Learning Disability	14	317	25	7
	Visual Impairment	64	361	40	58
Oral Presentation	Autism	116	321	26	8
	Emotional Disturbance	108	329	30	27
	Hard of Hearing	30	324	22	13
	Mental Retardation	135	309	21	6
	Orthopedic Impairment	38	325	27	24
	Other Health Impairment	225	332	28	25

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Specific Learning Disability	2,908	327	24	16
	Speech or Language Impairment	157	326	22	18
	Traumatic Brain Injury	17	320	24	12
	Visual Impairment	20	346	34	45
Oral Responses Dictated to a Scribe	Orthopedic Impairment	24	354	46	46
	Other Health Impairment	11	350	32	64
	Specific Learning Disability	30	327	35	27
Sign Language	Deaf	57	310	19	2
	Hard of Hearing	25	317	21	8
Spell Checker Or Grammar Checker	Autism	13	341	37	31
	Deaf	17	299	12	0
	Emotional Disturbance	12	331	21	17
	Hard of Hearing	11	316	28	9
	Mental Retardation	22	302	17	0
	Other Health Impairment	14	345	48	36
	Specific Learning Disability	284	329	23	15
	Speech or Language Impairment	17	334	37	24
Spell Checker Or Grammar Checker Off	Autism	24	370	43	71
	Emotional Disturbance	17	350	40	53
	Orthopedic Impairment	21	358	46	57
	Other Health Impairment	37	373	47	73
	Specific Learning Disability	95	349	36	44
Supervised Breaks	Autism	198	334	40	32
	Deaf	31	309	19	0
	Emotional Disturbance	509	328	39	28
	Hard of Hearing	59	337	37	37
	Mental Retardation	139	302	20	4
	Orthopedic Impairment	46	344	35	39
	Other Health Impairment	397	341	36	39
	Specific Learning Disability	3,778	329	28	23
	Speech or Language Impairment	191	331	28	23
	Traumatic Brain Injury	20	329	27	25
	Visual Impairment	33	345	39	48
Test Over More Than One Day	Autism	51	352	41	49
	Deaf	41	299	14	0
	Emotional Disturbance	119	324	36	22
	Hard of Hearing	22	331	38	41
	Mental Retardation	25	306	22	8
	Orthopedic Impairment	16	349	33	56
	Other Health Impairment	149	350	38	50
	Specific Learning Disability	1,098	336	30	31

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Speech or Language Impairment	66	335	28	30
	Visual Impairment	12	345	35	33
Tested At Home Or Hospital	Emotional Disturbance	12	352	57	50
	Other Health Impairment	12	346	47	42
	Specific Learning Disability	22	335	36	32
Transfer of Student T/B Responses to A/D	Autism	28	347	42	39
	Orthopedic Impairment	22	352	38	55
	Specific Learning Disability	50	339	28	34
	Visual Impairment	41	372	41	68
Unlisted Accommodation	Autism	27	347	44	37
	Emotional Disturbance	33	341	40	42
	Other Health Impairment	52	341	32	35
	Specific Learning Disability	501	330	29	23
	Speech or Language Impairment	25	328	23	8
Unlisted Modification	Specific Learning Disability	19	335	30	42
Writing ONLY	Autism	13	337	36	38
	Deaf	22	301	15	0
	Emotional Disturbance	12	332	35	33
	Mental Retardation	11	306	15	0
	Other Health Impairment	21	340	36	29
	Specific Learning Disability	263	330	24	22

[1] Results for groups with less 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 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Autism	12	327	14	8
	Emotional Disturbance	19	329	22	16
	Mental Retardation	14	314	15	7
	Other Health Impairment	48	339	27	31
	Specific Learning Disability	377	333	20	18
	Speech or Language Impairment	20	342	26	45
Beneficial Time	Autism	23	342	34	35
	Emotional Disturbance	220	326	30	17
	Hard of Hearing	14	338	39	29
	Mental Retardation	16	311	16	0
	Other Health Impairment	56	331	29	29
	Specific Learning Disability	502	334	27	28
	Speech or Language Impairment	23	332	33	22
Calculator	Autism	276	341	27	32
	Deaf	45	325	16	7
	Emotional Disturbance	419	335	26	28
	Hard of Hearing	61	335	20	26
	Mental Retardation	212	318	16	5
	Multiple Disability	18	325	18	11
	Orthopedic Impairment	104	340	23	36
	Other Health Impairment	727	339	24	29
	Specific Learning Disability	6,115	335	22	22
	Speech or Language Impairment	341	337	23	28
	Traumatic Brain Injury	37	333	20	19
	Visual Impairment	32	352	38	47
Dictionary for Math	Specific Learning Disability	150	333	20	17
	Speech or Language Impairment	14	339	24	43
IEP or Section 504 Plan	Autism	1,109	353	39	45
	Deaf	264	331	24	19
	Emotional Disturbance	2,167	336	31	28
	Hard of Hearing	319	348	33	45
	Mental Retardation	559	315	17	4
	Multiple Disability	44	323	26	11
	Orthopedic Impairment	295	344	32	37
	Other Health Impairment	2,807	345	31	39
	Specific Learning Disability	20,687	336	25	26
	Speech or Language Impairment	1,122	341	28	34
	Traumatic Brain Injury	102	333	23	24
	Visual Impairment	150	354	38	52
Large Print Version	Specific Learning Disability	13	326	22	15
	Visual Impairment	66	366	39	65
Math Manipulatives	Emotional Disturbance	13	340	17	15
	Specific Learning Disability	70	347	27	46

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Autism	89	337	27	24
	Emotional Disturbance	88	333	23	22
	Hard of Hearing	18	336	25	33
	Mental Retardation	100	319	17	6
	Orthopedic Impairment	26	333	25	35
	Other Health Impairment	173	336	28	27
	Specific Learning Disability	1,959	335	22	23
	Speech or Language Impairment	116	337	23	31
	Traumatic Brain Injury	18	329	22	22
	Visual Impairment	17	345	42	35
Oral Responses Dictated to a Scribe	Orthopedic Impairment	21	351	26	52
	Specific Learning Disability	34	337	20	32
Sign Language	Deaf	103	330	21	16
	Hard of Hearing	29	323	16	7
Supervised Breaks	Autism	179	342	35	32
	Deaf	24	328	23	13
	Emotional Disturbance	482	329	28	19
	Hard of Hearing	53	337	31	34
	Mental Retardation	121	314	17	2
	Orthopedic Impairment	44	340	27	36
	Other Health Impairment	370	342	30	36
	Specific Learning Disability	3,286	335	26	26
	Speech or Language Impairment	160	336	27	28
	Traumatic Brain Injury	18	330	19	17
	Visual Impairment	34	340	33	38
Test Over More Than One Day	Autism	38	354	36	47
	Emotional Disturbance	98	321	25	12
	Mental Retardation	18	319	21	6
	Orthopedic Impairment	11	346	27	36
	Other Health Impairment	111	351	35	53
	Specific Learning Disability	836	341	29	36
	Speech or Language Impairment	38	340	23	39
	Visual Impairment	11	340	37	27
Tested At Home Or Hospital	Emotional Disturbance	14	339	37	36
	Specific Learning Disability	14	342	36	43
Transfer of Student T/B Responses to A/D	Autism	21	341	40	24
	Orthopedic Impairment	16	368	45	63
	Other Health Impairment	12	335	27	25
	Specific Learning Disability	45	338	25	24
	Visual Impairment	35	361	40	60
Unlisted Accommodation	Autism	28	346	37	36
	Emotional Disturbance	36	338	28	28

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Other Health Impairment	39	342	25	36
	Specific Learning Disability	459	334	24	22
	Speech or Language Impairment	18	337	25	28
Unlisted Modification	Specific Learning Disability	35	335	23	26

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	15	339	25	27
Assistive Device No Interference	English Only	64	349	35	53
	Initially Fluent English Proficient	11	367	46	55
	English Learner	40	327	26	18
Beneficial Time	English Only	580	332	38	33
	Initially Fluent English Proficient	14	330	28	21
	Reclassified Fluent English Proficient	31	353	32	61
	English Learner	322	321	27	15
	Unknown	11	328	38	36
Braille Version	English Only	15	384	41	80
Dictionary	English Only	722	327	25	17
	Initially Fluent English Proficient	37	326	27	22
	Reclassified Fluent English Proficient	54	339	28	33
	English Learner	786	322	23	11
	Unknown	11	322	36	18
Essay Reponses(EO)	English Only	87	347	42	43
	English Learner	17	332	35	35
Essay Reponses(ER)	English Only	81	330	30	26
	English Learner	34	327	25	21
IEP or Section 504 Plan	English Only	17,984	338	35	34
	Initially Fluent English Proficient	957	344	36	41
	Reclassified Fluent English Proficient	1,467	352	33	52
	English Learner	10,899	323	25	14
	Unknown	488	340	40	39
Large Print Version	English Only	72	357	43	53
	Reclassified Fluent English Proficient	16	393	24	100
	English Learner	35	326	29	20
Oral Presentation	English Only	1,974	329	26	20
	Initially Fluent English Proficient	99	333	27	23
	Reclassified Fluent English Proficient	144	335	26	26
	English Learner	1,718	323	24	13
	Unknown	42	324	23	14
Oral Responses Dictated to a Scribe	English Only	72	340	40	39
	English Learner	28	311	27	11
Sign Language	English Only	54	328	37	20
	English Learner	40	312	19	5
Spell Checker Or Grammar Checker	English Only	236	330	30	19
	Initially Fluent English Proficient	18	342	28	33
	English Learner	158	323	23	11
Spell Checker Or Grammar Checker Off	English Only	202	365	44	62

Table 2.F.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Reclassified Fluent English Proficient	16	374	42	69
	English Learner	32	339	26	34
Supervised Breaks	English Only	3,078	333	34	30
	Initially Fluent English Proficient	188	335	33	29
	Reclassified Fluent English Proficient	280	346	31	45
	English Learner	2,102	323	25	15
	Unknown	85	324	34	20
Test Over More Than One Day	English Only	736	339	37	37
	Initially Fluent English Proficient	58	351	32	53
	Reclassified Fluent English Proficient	73	366	33	70
	English Learner	748	330	26	23
	Unknown	65	336	36	34
Tested At Home Or Hospital	English Only	75	364	43	63
	Reclassified Fluent English Proficient	12	404	32	100
	English Learner	27	333	36	33
Transfer of Student T/B Responses to A/D	English Only	152	352	41	49
	English Learner	25	335	25	32
Unlisted Accommodation	English Only	413	337	33	34
	Initially Fluent English Proficient	12	348	27	42
	Reclassified Fluent English Proficient	44	351	41	48
	English Learner	252	322	23	10
Unlisted Modification	English Only	20	346	35	55
	English Learner	12	332	25	25
Writing ONLY	English Only	195	332	31	24
	Initially Fluent English Proficient	18	324	32	28
	Reclassified Fluent English Proficient	41	340	32	39
	English Learner	133	325	22	17

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	323	334	22	23
	Initially Fluent English Proficient	13	332	12	8
	Reclassified Fluent English Proficient	13	333	21	15
	English Learner	186	332	22	18
Beneficial Time	English Only	590	334	31	28
	Initially Fluent English Proficient	16	339	44	38
	Reclassified Fluent English Proficient	30	356	33	60
	English Learner	289	331	27	22
	Unknown	13	312	30	8
Calculator	English Only	5,166	337	23	26
	Initially Fluent English Proficient	243	338	25	28
	Reclassified Fluent English Proficient	350	344	24	35
	English Learner	2,982	331	20	17
	Unknown	93	332	25	22
Dictionary for Math	English Only	109	335	21	27
	English Learner	84	331	17	12
IEP or Section 504 Plan	English Only	18,718	341	30	33
	Initially Fluent English Proficient	975	345	34	37
	Reclassified Fluent English Proficient	1,557	353	31	50
	English Learner	10,329	332	23	20
	Unknown	483	342	35	35
Large Print Version	English Only	68	356	36	54
	Reclassified Fluent English Proficient	16	378	42	69
	English Learner	27	338	36	33
Math Manipulatives	English Only	96	351	28	46
	English Learner	22	342	20	32
Oral Presentation	English Only	1,455	335	24	25
	Initially Fluent English Proficient	68	333	21	21
	Reclassified Fluent English Proficient	124	346	26	43
	English Learner	1,081	333	22	20
	Unknown	38	334	20	13
Oral Responses Dictated to a Scribe	English Only	71	348	30	48
	English Learner	25	329	21	16
Sign Language	English Only	84	335	27	19
	English Learner	53	326	18	11
Supervised Breaks	English Only	2,855	336	28	28
	Initially Fluent English Proficient	168	334	28	24
	Reclassified Fluent English Proficient	260	349	30	46

Table 2.F.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	1,729	332	25	21
	Unknown	67	332	30	19
Test Over More Than One Day	English Only	544	342	32	38
	Initially Fluent English Proficient	48	347	35	42
	Reclassified Fluent English Proficient	57	361	26	67
	English Learner	533	337	28	30
	Unknown	45	339	29	27
Tested At Home Or Hospital	English Only	59	352	41	47
Transfer of Student T/B Responses to A/D	English Only	122	351	36	44
	English Learner	25	341	31	32
Unlisted Accommodation	English Only	385	337	26	25
	Reclassified Fluent English Proficient	44	350	34	50
	English Learner	218	332	24	20
Unlisted Modification	English Only	27	335	27	19
	English Learner	13	334	23	23

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

[2] SD — Standard Deviation

Appendix 2.G: Results of Testing Variations and Disability Analyses—May 2011

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

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1,207	321	27	15
Modifications	2,045	324	25	16
All	3,252	323	26	16
Disability				
Unknown	-	-	-	-
Autism	294	325	29	19
Deaf	162	310	25	5
Deaf-Blindness	-	-	-	-
Emotional Disturbance	841	321	35	20
Hard of Hearing	107	322	24	8
Mental Retardation	250	304	22	4
Multiple Disability	18	313	29	11
Orthopedic Impairment	96	325	29	22
Other Health Impairment	729	326	30	22
Specific Learning Disability	7,048	321	25	14
Speech or Language Impairment	434	325	26	16
Traumatic Brain Injury	46	321	29	17
Visual Impairment	25	337	35	48

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

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

[2] SD — Standard Deviation

Table 2.G.2: Summary Statistics by Testing Variations and Disability—Mathematics, May 2011

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	996	328	23	15
Modifications	3,539	331	20	17
All	4,535	330	20	16
Disability				
Autism	297	332	24	19
Deaf	117	329	26	10
Deaf-Blindness	-	-	-	-
Emotional Disturbance	929	328	27	17
Hard of Hearing	87	331	17	18
Mental Retardation	238	318	16	3
Multiple Disability	21	327	13	5
Orthopedic Impairment	106	332	22	23
Other Health Impairment	839	332	22	21
Specific Learning Disability	7,077	329	20	14
Speech or Language Impairment	406	333	21	19
Traumatic Brain Injury	49	330	25	24
Visual Impairment	28	331	27	32

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

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

[2] SD — Standard Deviation

Table 2.G.3: Demographic Summary for All Examinees by Testing Variations—ELA, May 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	7,383	1,087	15	6,296	85	321	45	47	45	36	42	1.8
Transfer of T/B Responses to A/D (TS)	43	13	30	30	70	337	61	54	54	46	56	2.0
Oral Responses Dictated to a Scribe (OR)	41	9	22	32	78	322	47	47	44	38	42	1.8
Spell Checker/Grammar Checker Off (SO)	74	26	35	48	65	336	55	55	57	45	48	2.0
Essay Reponses (EO)	40	12	30	28	70	327	52	49	50	41	42	1.8
Assistive Device No Interference (AN)	41	5	12	36	88	324	50	51	47	37	40	2.0
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	23	5	22	18	78	326	53	51	46	38	47	1.9
Test Over More Than One Day (TD)	331	59	18	272	82	323	47	48	46	38	43	1.8
Supervised Breaks (SB)	1,313	173	13	1,140	87	320	45	46	44	35	42	1.8
Beneficial Time (BT)	324	46	14	278	86	318	44	45	44	33	41	1.7
Tested At Home Or Hospital (HH)	42	11	26	31	74	331	55	52	50	46	49	1.9
Dictionary (DI)	613	85	14	528	86	323	52	48	45	37	43	1.9
Sign Language (SL)	64	0	0	64	100	308	40	41	37	28	36	1.5
Oral Presentation (OP)	1,637	277	17	1,360	83	325	48	50	47	39	45	1.9
Spell Checker Or Grammar Checker (SC)	176	30	17	146	83	325	46	48	45	39	47	2.0
Essay Reponses (ER)	40	14	35	26	65	338	58	53	49	50	52	2.3
Assistive Device (AD)	23	8	35	15	65	329	52	53	46	42	47	2.1
Unlisted Modification (UM)	16	1	6	15	94	326	48	44	50	44	47	1.9
Unlisted Accommodation (UA)	190	32	17	158	83	325	50	50	48	39	44	1.9
Writing Only (WO)	111	3	3	108	97	312	41	39	39	33	37	1.7

[1] Results for groups with less 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 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	7,748	1,178	15	6,570	85	329	41	43	39	36	31
Transfer of T/B Responses to A/D (TS)	43	9	21	34	79	332	45	43	40	39	33
Oral Responses Dictated to a Scribe (OR)	31	9	29	22	71	339	49	47	49	41	34
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	18	4	22	14	78	339	49	48	48	39	34
Test Over More Than One Day (TD)	199	37	19	162	81	331	43	44	40	37	31
Supervised Breaks (SB)	1,119	160	14	959	86	328	40	43	38	35	30
Beneficial Time (BT)	250	35	14	215	86	326	39	42	37	34	30
Tested At Home Or Hospital (HH)	32	14	44	18	56	344	50	51	52	47	34
Dictionary for Math (DM)	123	22	18	101	82	331	43	44	40	39	32
Sign Language (SL)	34	2	6	32	94	324	34	37	35	35	32
Oral Presentation (OP)	895	177	20	718	80	333	43	46	40	39	33
Calculator (CA)	3,493	576	16	2,917	84	331	42	46	39	38	31
Arithmetic Table (AT)	163	41	25	122	75	335	49	48	41	39	34
Math Manipulatives (MM)	32	5	16	27	84	335	47	53	42	37	31
Assistive Device (AD)	17	5	29	12	71	340	44	55	49	39	37
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	176	28	16	148	84	329	40	43	38	35	32

[1] Results for groups with less 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 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	280	300	318	338	366	395	321	27	7,383
Transfer of T/B Responses to A/D (TS)	280	297	318	342	350	373	405	337	27	43
Oral Responses Dictated to a Scribe (OR)	275	275	299	316	336	383	401	322	32	41
Spell Checker/Grammar Checker Off (SO)	275	291	312	332	356	386	433	336	32	74
Essay Reponses (EO)	275	275	297	320	357	392	427	327	39	40
Assistive Device No Interference (AN)	291	297	304	323	342	362	381	324	23	41
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	275	278	310	323	346	369	405	326	32	23
Test Over More Than One Day (TD)	275	285	300	321	340	371	392	323	28	331
Supervised Breaks (SB)	275	280	300	318	338	366	386	320	26	1,313
Beneficial Time (BT)	275	276	297	316	334	369	386	318	27	324
Tested At Home Or Hospital (HH)	275	283	312	327	350	386	398	331	32	42
Dictionary (DI)	275	287	304	323	340	364	381	323	24	613
Sign Language (SL)	275	285	297	306	318	334	348	308	16	64
Oral Presentation (OP)	275	287	306	323	342	369	386	325	25	1,637
Spell Checker Or Grammar Checker (SC)	283	287	303	324	342	369	381	325	25	176
Essay Reponses (ER)	275	286	305	334	360	403	450	338	39	40
Assistive Device (AD)	280	287	312	323	352	369	378	329	26	23
Unlisted Modification (UM)	275	275	311	325	340	376	376	326	22	16
Unlisted Accommodation (UA)	275	287	304	325	340	373	395	325	27	190
Writing Only (WO)	283	287	299	308	323	346	371	312	19	111

[1] SD — Standard Deviation

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

Table 2.G.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, May 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	301	315	327	341	366	392	329	21	7,748
Transfer of T/B Responses to A/D (TS)	275	291	315	329	347	363	426	332	27	43
Oral Responses Dictated to a Scribe (OR)	275	275	319	327	364	406	422	339	37	31
Braille Version (BV))	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV))	287	287	321	338	347	426	426	339	31	18
Test Over More Than One Day (TD)	275	296	315	329	345	376	409	331	25	199
Supervised Breaks (SB))	280	299	313	325	340	368	387	328	21	1,119
Beneficial Time (BT)	275	293	311	325	340	364	397	326	23	250
Tested At Home Or Hospital (HH)	275	275	316	340	374	400	450	344	41	32
Dictionary for Math (DM)	299	304	319	329	343	370	376	331	19	123
Sign Language (SL)	301	304	315	321	331	357	364	324	14	34
Oral Presentation (OP)	293	304	319	331	345	372	397	333	21	895
Calculator (CA)	290	304	317	329	343	366	385	331	20	3,493
Arithmetic Table (AT)	299	309	321	334	350	368	378	335	19	163
Math Manipulatives (MM)	294	301	322	335	344	376	378	335	21	32
Assistive Device (AD)	313	313	333	336	350	372	372	340	18	17
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-
Unlisted Accommodation (UA)	296	306	315	325	340	363	385	329	19	176

[1] SD — Standard Deviation

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

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

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	17	328	26	35
Assistive Device No Interference	Specific Learning Disability	25	327	22	12
Beneficial Time	Autism	19	317	31	11
	Emotional Disturbance	77	312	29	10
	Other Health Impairment	20	319	24	15
	Specific Learning Disability	165	321	26	15
Dictionary	Autism	13	334	27	15
	Deaf	11	310	17	0
	Emotional Disturbance	19	312	24	5
	Mental Retardation	22	308	24	9
	Other Health Impairment	28	322	25	14
	Specific Learning Disability	446	324	23	14
	Speech or Language Impairment	20	319	23	10
Essay Reponses(ER)	Specific Learning Disability	23	345	31	39
IEP or Section 504 Plan	Autism	212	324	28	19
	Deaf	95	315	27	7
	Emotional Disturbance	632	319	34	18
	Hard of Hearing	71	319	24	8
	Mental Retardation	190	304	20	3
	Multiple Disability	11	311	24	0
	Orthopedic Impairment	68	329	27	28
	Other Health Impairment	493	325	31	21
	Specific Learning Disability	4,879	320	25	13
	Speech or Language Impairment	283	325	24	14
	Traumatic Brain Injury	31	318	29	16
	Visual Impairment	22	333	35	45
Oral Presentation	Autism	51	326	23	16
	Emotional Disturbance	40	321	26	15
	Hard of Hearing	14	323	23	7
	Mental Retardation	58	310	25	9
	Orthopedic Impairment	22	337	24	45
	Other Health Impairment	99	326	27	23
	Specific Learning Disability	1,162	325	25	16
	Speech or Language Impairment	105	324	22	11
Sign Language	Deaf	46	307	14	0
	Hard of Hearing	16	315	17	0
Spell Checker Or Grammar Checker	Deaf	14	302	13	0
	Other Health Impairment	15	324	35	27
	Specific Learning Disability	107	328	22	16
	Speech or Language Impairment	12	326	28	25
Spell Checker Or Grammar Checker Off	Specific Learning Disability	34	328	30	26
Supervised Breaks	Autism	47	322	29	15
	Deaf	12	323	14	0
	Emotional Disturbance	150	312	28	10

Table 2.G.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Hard of Hearing	14	315	22	7
	Mental Retardation	40	302	20	5
	Other Health Impairment	88	328	31	27
	Specific Learning Disability	847	321	24	13
	Speech or Language Impairment	38	326	17	8
Test Over More Than One Day	Autism	14	317	22	7
	Deaf	15	301	13	0
	Emotional Disturbance	37	315	30	14
	Other Health Impairment	22	326	28	18
	Specific Learning Disability	188	324	26	18
	Speech or Language Impairment	11	340	28	36
Tested At Home Or Hospital	Specific Learning Disability	11	335	28	27
Transfer of Student T/B Responses to A/D	Specific Learning Disability	12	345	20	58
Unlisted Accommodation	Emotional Disturbance	16	331	36	31
	Other Health Impairment	15	325	25	13
	Specific Learning Disability	126	322	25	13
Unlisted Modification	Specific Learning Disability	13	332	18	8
Writing ONLY	Deaf	38	307	16	0
	Hard of Hearing	13	310	20	0
	Specific Learning Disability	42	316	21	5

[1] Results for groups with less 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 2011

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Other Health Impairment	16	343	20	38
	Specific Learning Disability	103	333	19	21
Assistive Device	Specific Learning Disability	11	340	16	27
Beneficial Time	Autism	15	319	23	7
	Emotional Disturbance	64	323	24	13
	Other Health Impairment	20	327	24	25
	Specific Learning Disability	117	328	23	13
Calculator	Autism	98	333	18	15
	Deaf	24	326	13	4
	Emotional Disturbance	167	330	18	16
	Hard of Hearing	30	333	19	23
	Mental Retardation	83	322	16	6
	Orthopedic Impairment	40	338	22	28
	Other Health Impairment	265	333	21	23
	Specific Learning Disability	2,406	331	20	16
	Speech or Language Impairment	158	330	17	12
	Traumatic Brain Injury	15	340	26	33
Dictionary for Math	Specific Learning Disability	86	333	19	21
IEP or Section 504 Plan	Autism	222	330	20	18
	Deaf	66	333	31	12
	Emotional Disturbance	691	326	26	15
	Hard of Hearing	62	331	18	19
	Mental Retardation	183	318	15	3
	Multiple Disability	18	328	12	6
	Orthopedic Impairment	73	334	22	25
	Other Health Impairment	587	332	22	21
	Specific Learning Disability	5,026	329	20	14
	Speech or Language Impairment	281	332	19	17
	Traumatic Brain Injury	37	330	24	24
	Visual Impairment	25	327	25	28
Math Manipulatives	Specific Learning Disability	21	336	21	19
Oral Presentation	Autism	26	335	17	19
	Emotional Disturbance	27	327	18	15
	Hard of Hearing	12	343	16	25
	Mental Retardation	29	325	18	10
	Orthopedic Impairment	17	343	23	35
	Other Health Impairment	69	334	22	28
	Specific Learning Disability	601	333	22	19
	Speech or Language Impairment	51	334	17	22
Sign Language	Deaf	25	323	13	4
Supervised Breaks	Autism	47	329	21	15
	Emotional Disturbance	128	321	19	5
	Hard of Hearing	12	330	15	8

Table 2.G.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Mental Retardation	31	317	18	6
	Other Health Impairment	80	329	21	19
	Specific Learning Disability	702	329	22	16
	Speech or Language Impairment	32	333	17	13
Test Over More Than One Day	Autism	11	332	23	18
	Emotional Disturbance	23	330	24	17
	Other Health Impairment	17	336	16	24
	Specific Learning Disability	116	331	26	19
Transfer of Student T/B Responses to A/D	Specific Learning Disability	14	341	19	36
Unlisted Accommodation	Emotional Disturbance	14	324	21	14
	Other Health Impairment	18	329	18	22
	Specific Learning Disability	110	328	18	13

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	12	330	27	33
	English Learner	11	328	26	36
Assistive Device No Interference	English Only	21	325	23	14
	English Learner	16	323	24	13
Beneficial Time	English Only	208	320	28	17
	English Learner	98	315	21	6
Dictionary	English Only	310	323	24	14
	Initially Fluent English Proficient	13	328	27	23
	Reclassified Fluent English Proficient	26	327	27	23
	English Learner	253	322	22	12
	Unknown	11	315	40	18
Essay Reponses(EO)	English Only	20	331	37	35
Essay Reponses(ER)	English Only	14	334	44	29
	English Learner	14	327	31	29
IEP or Section 504 Plan	English Only	4,209	322	29	17
	Initially Fluent English Proficient	169	325	25	22
	Reclassified Fluent English Proficient	220	330	27	26
	English Learner	2,618	317	23	10
	Unknown	167	323	35	19
Large Print Version	English Only	13	333	35	31
Oral Presentation	English Only	853	327	26	20
	Initially Fluent English Proficient	31	334	23	29
	Reclassified Fluent English Proficient	47	335	27	30
	English Learner	677	321	23	11
	Unknown	29	329	36	28
Oral Responses Dictated to a Scribe	English Only	21	324	34	24
	English Learner	15	317	23	13
Sign Language	English Only	42	308	16	0
	English Learner	20	309	14	0
Spell Checker Or Grammar Checker	English Only	95	329	27	21
	English Learner	72	320	22	11
Spell Checker Or Grammar Checker Off	English Only	53	334	35	34
	English Learner	18	338	25	39
Supervised Breaks	English Only	712	321	27	15
	Initially Fluent English Proficient	38	326	27	26
	Reclassified Fluent English Proficient	43	330	24	26
	English Learner	499	317	22	9
	Unknown	21	310	37	14
Test Over More Than One Day	English Only	187	323	28	19

Table 2.G.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	122	321	24	13
Tested At Home Or Hospital	English Only	28	332	36	32
	English Learner	14	331	25	14
Transfer of Student T/B Responses to A/D	English Only	33	339	27	33
Unlisted Accommodation	English Only	99	330	29	23
	English Learner	79	318	23	5
Writing ONLY	English Only	47	313	19	4
	English Learner	54	308	17	0

[1] Results for groups with less 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 2011

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	97	335	19	27
	English Learner	55	333	18	18
Assistive Device	English Only	16	341	17	31
Beneficial Time	English Only	170	328	24	16
	English Learner	70	324	19	10
Calculator	English Only	2,155	332	20	18
	Initially Fluent English Proficient	73	336	22	19
	Reclassified Fluent English Proficient	118	335	18	23
	English Learner	1,070	328	18	11
	Unknown	77	334	22	22
Dictionary for Math	English Only	71	333	19	21
	English Learner	47	330	17	13
IEP or Section 504 Plan	English Only	4,716	330	22	17
	Initially Fluent English Proficient	193	331	20	17
	Reclassified Fluent English Proficient	259	337	22	26
	English Learner	2,398	327	19	11
	Unknown	182	331	25	20
Large Print Version	English Only	13	342	32	23
Math Manipulatives	English Only	17	334	15	12
Oral Presentation	English Only	506	332	22	19
	Initially Fluent English Proficient	24	337	28	21
	Reclassified Fluent English Proficient	27	352	19	52
	English Learner	311	331	19	18
	Unknown	27	334	22	22
Oral Responses Dictated to a Scribe	English Only	19	336	35	26
Sign Language	English Only	19	326	15	11
	English Learner	11	322	14	0
Supervised Breaks	English Only	637	328	21	14
	Initially Fluent English Proficient	36	328	23	17
	Reclassified Fluent English Proficient	44	341	24	32
	English Learner	387	327	19	13
	Unknown	15	311	17	0
Test Over More Than One Day	English Only	122	333	23	20
	English Learner	60	327	27	15
Tested At Home Or Hospital	English Only	19	337	42	32
Transfer of Student T/B Responses to A/D	English Only	34	334	28	24
Unlisted Accommodation	English Only	95	328	19	18
	English Learner	72	328	16	10

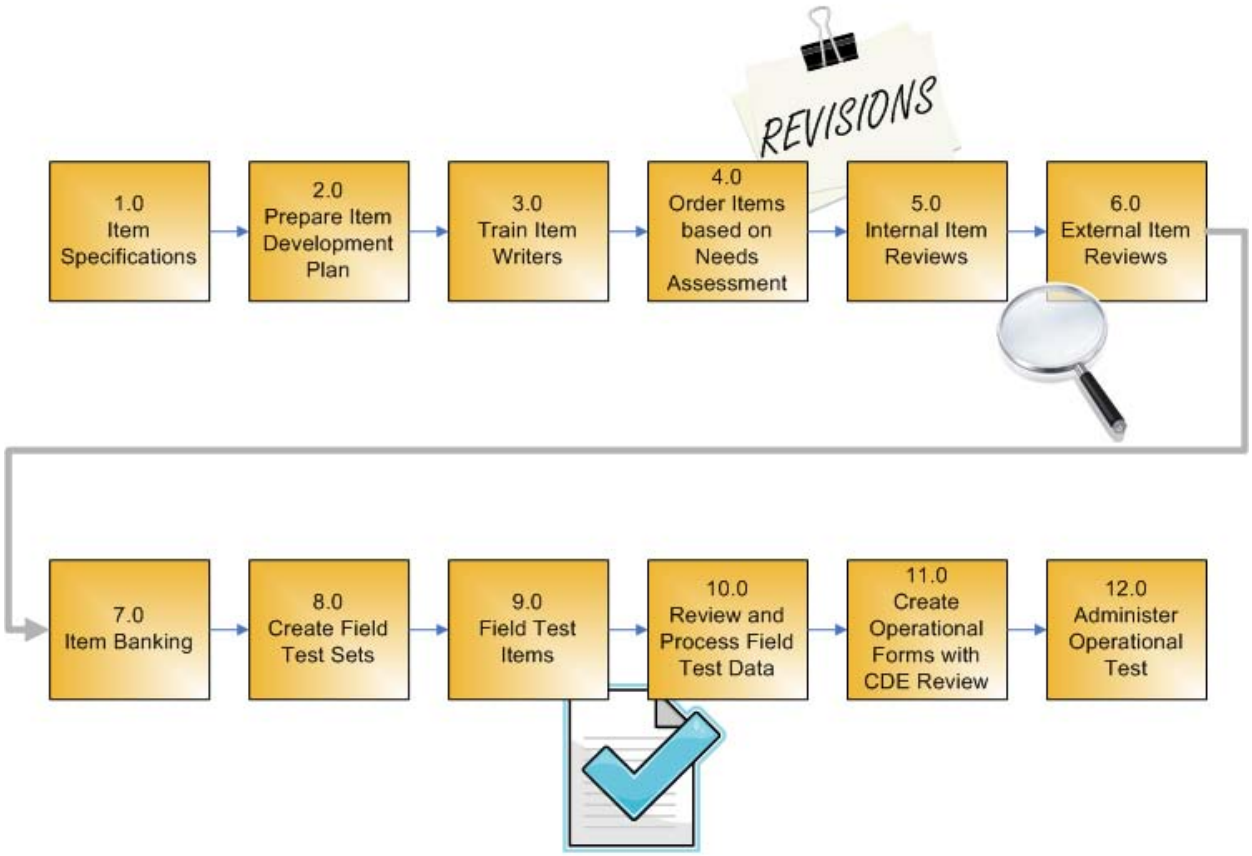
[1] Results for groups with less than 11 students are not reported.

[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, 2010c).
- 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 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).

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.

The field-test percentages and item counts for the 2011 field test 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 1.20 times or 120 percent of the number of unique items used to score students across the administrations. The number of mathematics items field tested was 0.03 times or 3 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	≈ 120	500
Mathematics	≈ 3	16

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 2010-11 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 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 used the Item Bank Inventory and the Item Development Plan to determine how many items were still needed for each standard and ordered items from item writers based on these numbers. The item writers were 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 provided 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.
- Checking the language of the items 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 2010-11 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 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 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 into 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 maintain 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. However, due to state budget cuts the 2010 stand-alone writing prompt field testing was not conducted.

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 over 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 2011 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 mathematics field-test form was administered in February and March due to the state budget cuts. One ELA field-test form was administered in February and 76 ELA field-test forms were administered in March. Additional ELA item development took place to meet the requirements for operationally-ready items for the item bank. Due to the timing of item development, the majority of field testing occurred in March.

Table 3.2 Summary of Items and Forms Presented in 2011 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**	**No. of Examinees Total[1]**
ELA	**February**	73	157,897	156,783	1	7	157,897	156,783
	March	73	407,128	405,036	76	493	5,129-10,698[2]	405,036
Mathematics	**February**	80	160,859	159,771	1	8	160,859	159,771
	March	80	410,152	406,798	1	8	410,152	406,798

[1] This number 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) the distractors 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 2010–2011 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 2010–11 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 are 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

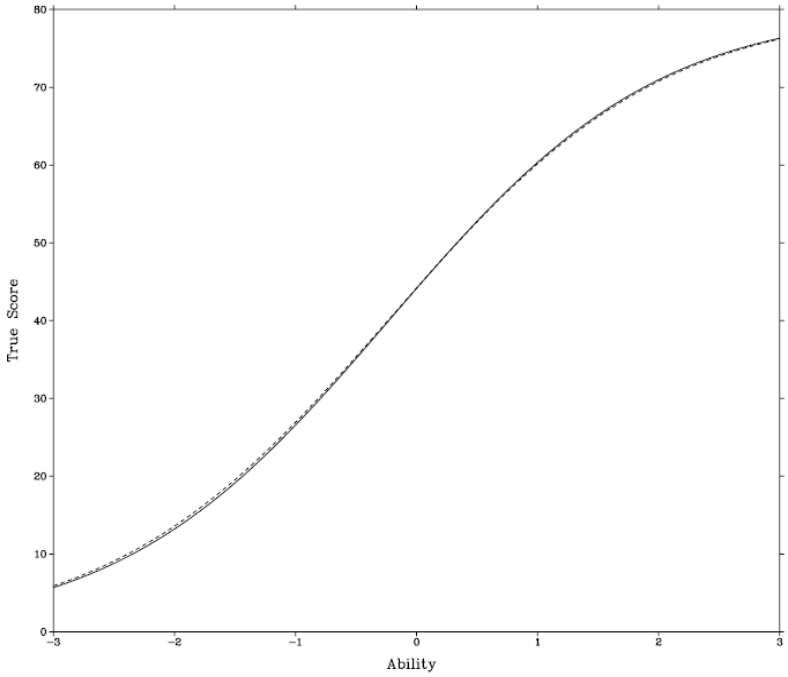


Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items

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, Adobe Acrobat PDF, 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, 2010a) 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, 2010a) 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 student's 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 a 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 CAHSEE tests, as described in this section.

Test Administrators

The CAHSEE tests are administered seven times in a school year. In that respect, 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, 2010a). This manual is described in the next 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. They are also 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, 2010b) exactly.

Test Proctor

A test proctor is an employee of the school district or a person, assigned by an NPS to implement the IEP of a student, who has received training designed to prepare him or her to assist the test examiner in the administration of the CAHSEE. Test proctors must sign the *CAHSEE Test Security Affidavits*.

Scribe

A scribe is an employee of the school district or a person, assigned by an NPS to implement the IEP of a student, who 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 Affidavits*.

CAHSEE Directions for Administration

The CAHSEE *DFA Manuals* are 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, 2010a) 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 IEP or Section 504 plan. If students use 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 actually used on each test – ELA and mathematics. The coordinator also indicates, by signature, that the student had access to the modifications and/or accommodations as 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 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 determine 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 "Score Invalidated." The information and procedures to assist in identifying irregularities and notifying the CDE are provided in the *CAHSEE DTSCM* (CDE, 2010a).

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, 2010a). 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 2010–11. 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 percent 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.

Percent of students choosing each response option:

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

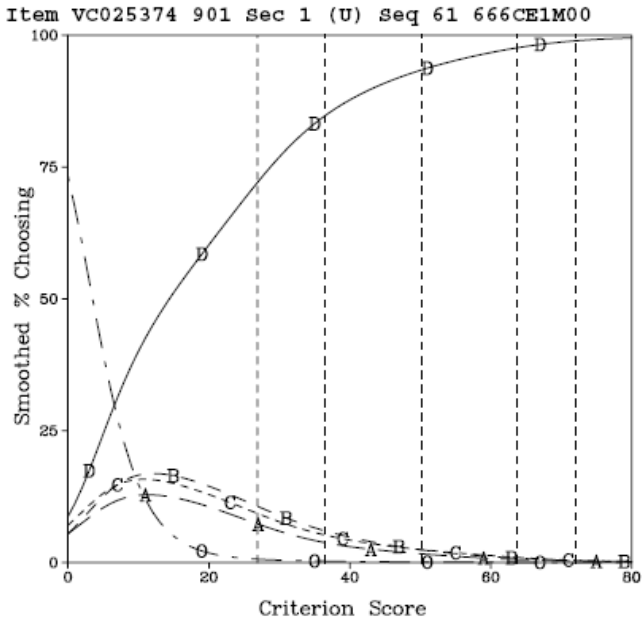
Percent of students omitting an item:

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 percent 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 percent 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

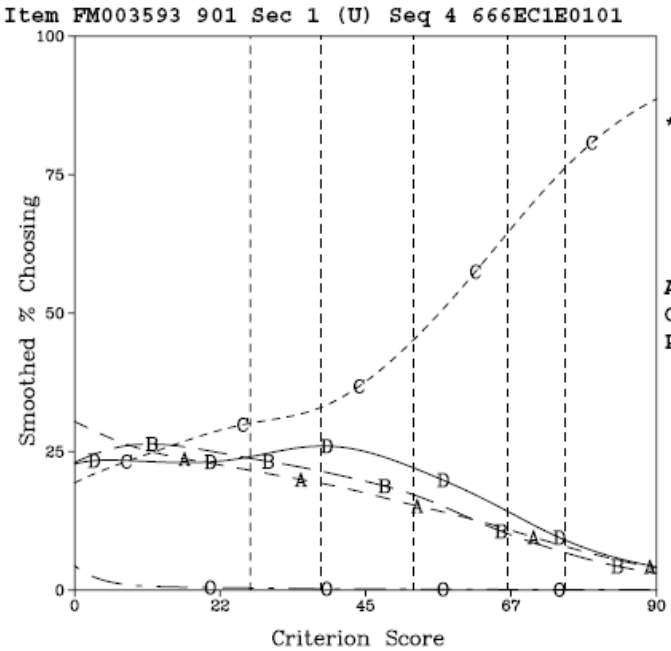


Multiple Choice, 4 choice

	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 2009 through May 2010 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, there is one clear and correct answer, and 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 2010 items were placed on the reference scale through a set of linking items that appeared in the 2009 operational forms and were re-administered in 2010. 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 2010–11 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:

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

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, ξ_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 2010–11 forms were linked back to scale established in February 2004.

Complete raw-to-scale score conversion tables for the 2010–11 CAHSEE are presented in Appendix 6.O. The raw scores and corresponding unrounded converted scale scores are listed in those tables. For all the 2010–11 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 2010 operational form was used as the Braille, Braille large-print, and Braille audio CD forms for the July, October, and November 2010 administrations for both ELA and mathematics. The December 2010 operational form was used for the December 2010 and February 2011 administrations, and the March operational form was used for the March and May 2011 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	7 (100)
	Mathematics	8 (88)
March 2011	ELA	493 (94)
	Mathematics	8 (63)

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 0.86 to 0.94. Reliability estimates for the mathematics section across the seven administrations ranged from 0.86 to 0.95. Reliabilities for the MC strands for ELA ranged from 0.36 to 0.84, while the strand reliabilities for mathematics ranged from 0.39 to 0.86. The reliabilities for the ELA CR (unweighted) item score ranged from 0.21 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 2010–11 administrations ranged from 3.81 to 4.46 and 3.59 to 4.13, 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 2010–11 administrations, the correlations between ELA strands (including essay) ranged from 0.35 to 0.79, while the correlations between mathematics strands ranged from 0.33 to 0.81. The reliabilities of the ELA strands ranged from 0.36 to 0.84, and the reliability of the mathematics strands ranged from 0.39 to 0.86. The SEM for the MC strands for ELA ranged from 0.97 to 2.04, while the SEM for mathematics ranged from 1.39 to 2.08. The SEM for the ELA CR items ranged from 0.43 to 0.52.

Subgroup Reliabilities and Standard Errors of Measurement

Reliabilities (R_{XX}) and SEM estimates are reported for subgroups from the February and March 2011 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 2011 administration. Table 6.K.9 provides the same information for the mathematics test of the February 2011 administration. Tables 6.K.10 and 6.K.11 provide R_{XX} and SEM for the subgroups for ELA and mathematics for the March 2011 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 2.77 (Pacific Islander, March) to 3.64 (African American, February) for the MC total and a range of 3.59 (Pacific Islander, Two or More Races, March) to 4.33 (White, February) for the composite. All R_{XX} are greater than or equal to 0.87.

The R_{XX} for the accommodation group are the same for March and slightly lower compared to the nonaccommodation 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 2.80 (Pacific Islander, March) to 3.98 (White, February). The R_{XX} for the accommodation groups are slightly lower than 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 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 November 2010, February and March 2011 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 0.21 to 0.34. The correlation between first and second ratings across the seven administrations ranged from 0.62 to 0.74.

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 March 2011 administration and the largest percentage was in the May 2011 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 the Appendix 6.M. Generalizability coefficients for the CR item ranged from 0.76 to 0.85 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 0.95 to 1.00. The consistency of classifying students into the Advanced level versus Proficient level or below for ELA ranged from 0.89 to 0.99, while the consistency for mathematics ranged from 0.93 to 1.00.

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 0.94 to 0.98. The consistency of classifying students into the Proficient level or above versus Below Proficient level for ELA ranged from 0.89 to 0.96, 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 0.89 to 0.95, while the decision accuracy for mathematics at the Pass/Not Pass levels ranged from 0.91 to 0.95. The decision consistency for ELA at the Pass/Not Pass levels ranged from 0.84 to 0.92, while the decision consistency for mathematics at the Pass/Not Pass level ranged from 0.87 to 0.93. The magnitude of the estimates (0.84 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, the population to be assessed, as well as 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 has 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 2010 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 to establish whether the material on the test is the same as that which they are expected to know. 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 portions of the CAHSEE to verify the quality of the tests. The Independent Evaluation report for 2009 contains the results of an alignment study for the February 2009 ELA test form (Taylor & Wise, 2009), while the 2008 Independent Evaluation report includes the results of an alignment study for the February 2008 ELA and mathematics forms (Taylor, Johnstone, Wise, Thacker, & Hardoin-Mandeville, 2008).

There were two approaches to evaluating the content alignment of the February 2009 ELA test. The panelists-test developer agreement reveals the extent to which panelists chose the same category content and standard for the items as the test developers. Results indicated that panelists demonstrated high agreement with test developers at the content category level, but showed moderate agreement at the standard level (most specific level). The Webb alignment method was used to evaluate the alignment of the 2009 ELA test to the California content standards. The Webb method requires a set of raters to evaluate each test item on two different dimensions: (1) the standard(s) targeted by items and (2) the DOK required of students to respond to items. These ratings form the basis of the four separate Webb alignment analyses: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation. Webb Alignment statistics indicate acceptable assessment for most strands.

For the universal test design study the panelists were asked to rate the quality of the items in terms of lack of flaws and accessibility to all students. The analyses indicated that a large majority of items were well constructed, unbiased, and accessible to a wide range of students.

Overall, the findings of the 2009 study parallel the results of the 2008 alignment study. Both 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, as well as 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 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 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 2010–11 administrations few operational items were flagged for significant DIF per administration. Of the 30 items that were flagged, 27 were ELA items and 3 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 0.58 to 0.70) 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., 0.80 and 0.79 for the February and March administrations, respectively).

In addition, inter-correlations 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, inter-correlations 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 inter-correlations between the content area scores and the strand scores were 0.85 and 0.89 for ELA and mathematics, respectively. The mean inter-correlations between the ELA score and mathematics strands was 0.72, and the mean inter-correlations 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 0.85 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 67 to 79 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 0.45 for ELA and from 0.28 to 0.47 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 2010 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 is 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 0.71 (SD = 0.05) and 0.70 (SD = 0.05) for the February and March administrations, respectively. The mathematics strands exhibit mean correlations of 0.74 (SD = 0.05) and 0.73 (SD = 0.05) 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 = 0.04) 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 for a) word analysis, fluency and systematic vocabulary development; and b) reading comprehension; and for writing, the scores were 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.

- 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.

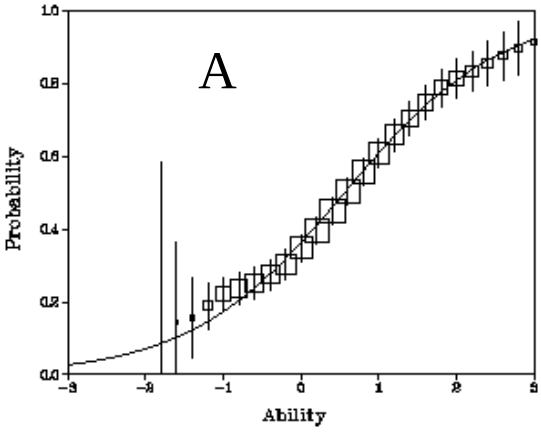
Flag D (Item 94, FM002620)

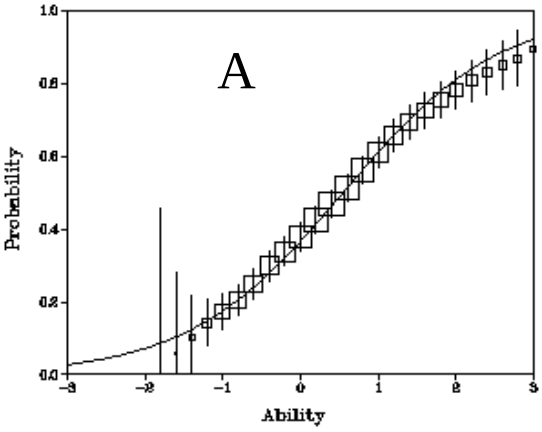
- Fit at the passing score may be slightly out of error range.
- Theoretical curve outside of error range at some regions across ability range.
- Empirical curve may have a zero slope at and around the passing score.
- Large Chi-square value relative to the other items in the calibration with similar 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)

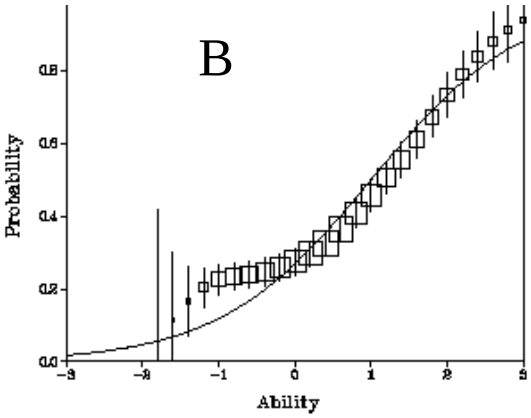
- Fit at the passing score outside error range.
- Theoretical curve outside of error range at most regions across ability range.
- Probability of answering item correctly may be greater at lower ability than higher ability (U shaped empirical curve).
- Very large Chi-square value (sometimes larger than 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

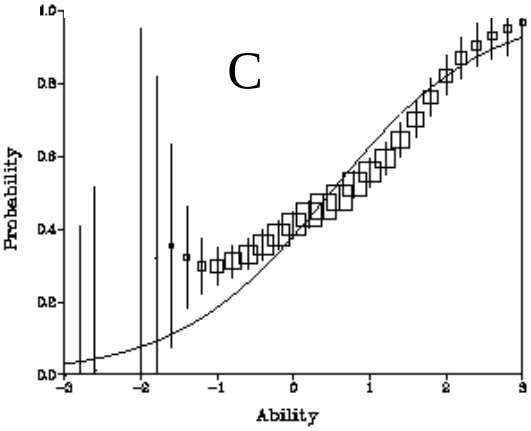




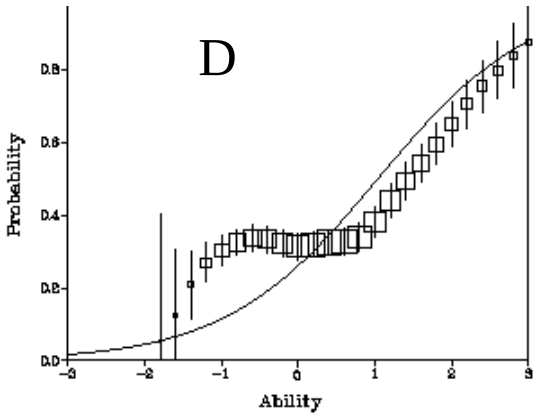
No.93 FM002619901 4 Choice P+=0.494
a=0.588 F, b=0.561, c=0.000 F, CHI=11.85



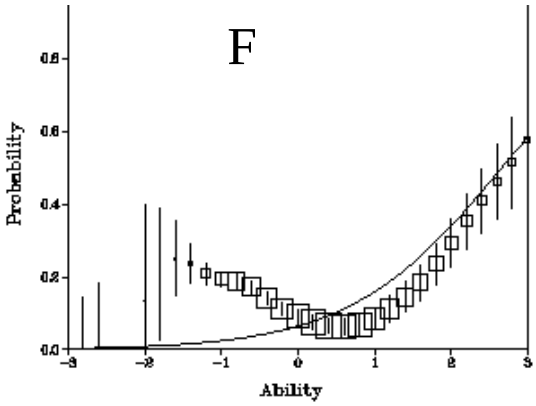
No.96 FM002641901 4 Choice P+=0.402
a=0.588 F, b=1.003, c=0.000 F, CHI=116.63



No.97 FM002767901 4 Choice P+=0.506
a=0.588 F, b=0.486, c=0.000 F, CHI=151.58



No.94 FM002620901 4 Choice P+=0.395
a=0.588 F, b=1.040, c=0.000 F, CHI=480.45



No.98 FM002768901 4 Choice P+=0.136
a=0.588 F, b= 2.664, c=0.000 F, CHI=720.60

Appendix 6.B: Summary Statistics for Operational Items—July 2010

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	-0.06	0.70	-1.81	1.31	0.31	0.08	0.14	0.46
RC	18	-0.16	0.75	-1.70	0.97	0.33	0.08	0.16	0.44
RL	20	-0.19	0.72	-1.81	0.84	0.32	0.08	0.20	0.44
RW	7	-0.54	0.41	-1.27	0.08	0.28	0.08	0.15	0.38
WA	1	0.66	-	0.66	0.66	0.60	-	0.60	0.60
WC	15	0.02	0.65	-0.94	1.31	0.30	0.09	0.14	0.46
WS	12	0.46	0.46	-0.41	1.12	0.29	0.06	0.19	0.41

[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	0.66	-1.88	1.35	0.28	0.07	0.11	0.47
A1	12	0.23	0.51	-0.71	1.35	0.26	0.06	0.17	0.37
AF	20	-0.36	0.60	-1.88	0.81	0.29	0.07	0.15	0.41
MG	18	-0.07	0.53	-0.90	0.84	0.30	0.10	0.11	0.47
MR	8	-0.50	0.75	-1.46	0.81	0.27	0.07	0.15	0.40
NS	17	-0.38	0.81	-1.75	0.87	0.27	0.06	0.17	0.38
PS	13	-0.57	0.58	-1.46	0.14	0.29	0.08	0.16	0.40

[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	32	44
B	Use	19	26
C	Use	22	30
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	38	48
B	Use	24	30
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.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	FM005151	L158A014	0	23	A	S	S	S	S	S	A	A	C+
ELA	VE046462	L6OSA1169	0	75	A	S	S	S	S	S	A	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	0	0	1	1
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5	2	3		
A	73	100	0	0	0	0	0	0	0	0	0	0	73	100	66	90	68	93		
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	2	3		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	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	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 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-	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1		
A	78	98	0	0	0	0	0	0	0	0	0	0	80	100	78	98	78	98		
B+	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	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	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	VC020381
CAHSEE ID	L30SA316
Polyserial correlation	0.66
IRT b-value	0.6632
Step category 1[1]	1.5906
Step category 2	1.4785
Step category 3	3.6582
Step category 4	-0.9642
Step category 5	-0.8428
Step category 6	-2.6369
Step category 7	-2.2833
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 2010

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	0.60	-1.54	1.17	0.35	0.06	0.21	0.52
RC	18	-0.04	0.56	-1.11	0.60	0.35	0.08	0.24	0.52
RL	20	-0.10	0.57	-1.54	0.68	0.36	0.05	0.25	0.45
RW	7	-0.46	0.42	-1.04	0.00	0.32	0.06	0.26	0.40
WA	1	0.82	-	0.82	0.82	0.64	-	0.64	0.64
WC	15	0.00	0.67	-1.21	1.17	0.35	0.06	0.21	0.46
WS	12	0.17	0.62	-1.16	0.89	0.36	0.05	0.29	0.45

[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	-0.25	0.59	-1.87	1.27	0.33	0.07	0.19	0.47
A1	12	0.16	0.52	-0.70	1.18	0.31	0.06	0.22	0.43
AF	20	-0.41	0.42	-1.03	0.54	0.36	0.08	0.19	0.47
MG	18	0.00	0.51	-0.89	1.27	0.35	0.07	0.21	0.47
MR	8	-0.19	0.43	-0.84	0.33	0.32	0.08	0.19	0.41
NS	17	-0.49	0.67	-1.87	0.41	0.31	0.06	0.19	0.41
PS	13	-0.45	0.58	-1.23	0.45	0.34	0.06	0.25	0.45

[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	36	49
B	Use	17	23
C	Use	20	27
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	31	39
B	Use	22	28
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.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	VC019812	L268C015-F	0	3	A	A	C-	S	B-	C-	A	A	A
ELA	VC021184	L390E004	0	18	A	A	A	S	C-	B-	B-	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	1	1	0	0	1	1	1	1	0	0	0	0	0	0	2	3
B-	4	5	0	0	3	4	0	0	4	5	3	4	3	4	0	0	3	4		
A	69	95	73	100	66	90	0	0	66	90	69	95	70	96	73	100	67	92		
B+	0	0	0	0	3	4	0	0	2	3	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	0	0	0	0	73	100	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.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]	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	2	3	3	4	1	1	0	0	0	0	0	0	1	1		
A	75	94	80	100	77	96	75	94	77	96	80	100	80	100	80	100	79	99		
B+	2	3	0	0	1	1	2	3	2	3	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	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.C.8: Listing of CR Item Statistics—ELA

Accession number	VE025109
CAHSEE ID	L6OSA1401
Polyserial correlation	0.64
IRT b-value	0.8154
Step category 1[1]	1.4825
Step category 2	1.5171
Step category 3	3.6680
Step category 4	-0.9064
Step category 5	-0.8334
Step category 6	-2.4858
Step category 7	-2.4421
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	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

[3] S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.D: Summary Statistics for Operational Items—November 2010

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	-0.05	0.65	-1.70	1.35	0.36	0.07	0.20	0.50
RC	18	-0.11	0.53	-0.72	1.19	0.37	0.07	0.20	0.46
RL	20	-0.25	0.64	-1.40	0.78	0.34	0.08	0.22	0.48
RW	7	-0.49	0.63	-1.70	0.18	0.33	0.07	0.26	0.44
WA	1	0.74	-	0.74	0.74	0.66	-	0.66	0.66
WC	15	0.24	0.75	-1.47	1.35	0.35	0.08	0.20	0.49
WS	12	0.18	0.50	-0.51	1.05	0.39	0.06	0.29	0.50

[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	-0.26	0.67	-2.09	1.22	0.35	0.07	0.19	0.50
A1	12	0.20	0.48	-0.75	1.22	0.33	0.08	0.19	0.43
AF	20	-0.47	0.52	-1.89	0.61	0.35	0.07	0.19	0.46
MG	18	-0.12	0.56	-0.89	1.22	0.33	0.08	0.22	0.50
MR	8	-0.58	0.70	-2.00	0.33	0.39	0.04	0.33	0.45
NS	17	-0.35	0.69	-1.79	0.65	0.36	0.08	0.22	0.47
PS	13	-0.46	0.91	-2.09	0.80	0.35	0.06	0.25	0.44

[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	22	30
B	Use	43	59
C	Use	8	11
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	35	44
B	Use	33	41
C	Use	12	15
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	VC392342	L579G008	0	1	B-	A	C-	B-	A	B-	A	A	C-
ELA	VC392340	L579G006	0	5	A	A	C-	B-	C-	C-	A	A	A
ELA	VC018472	L221E002	0	15	C-	A	C-	B-	C-	B-	B-	B-	A
ELA	VC018481	L221E010	0	17	B-	A	C-	A	B-	B-	B-	B-	C-
ELA	VC018487	L221E014	0	19	B-	A	B-	B-	A	A	A	B-	C-
ELA	VC392614	L5OSA561	0	78	A	A	B+	A	C+	B+	A	A	A
MATH	VE047928	M50826	0	91	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.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]	1	1	0	0	4	5	0	0	2	3	1	1	0	0	0	0	3	4	5	7
B-	4	5	0	0	3	4	6	8	1	1	3	4	2	3	3	4	0	0		
A	68	93	72	99	61	84	66	90	68	93	67	92	71	97	70	96	69	95		
B+	0	0	1	1	5	7	1	1	1	1	2	3	0	0	0	0	1	1		
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	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	6	8

[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	0	0	0	0	0	0	0	0	0	0	0	0
B-	0	0	0	0	2	3	3	4	3	4	3	4	0	0	0	0	0	0		
A	78	98	80	100	76	95	75	94	74	93	75	94	80	100	80	100	80	100		
B+	2	3	0	0	2	3	2	3	2	3	2	3	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	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	VE025113
CAHSEE ID	L6OSA1407
Polyserial correlation	0.66
IRT b-value	0.73971
Step category 1[1]	1.9380
Step category 2	1.4752
Step category 3	3.3247
Step category 4	-0.8440
Step category 5	-0.8847
Step category 6	-2.4420
Step category 7	-2.5673
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	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]	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.E: Summary Statistics for Operational Items—December 2010

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	-0.06	0.68	-2.10	1.34	0.31	0.06	0.13	0.51
RC	18	-0.15	0.48	-1.34	0.39	0.33	0.07	0.22	0.51
RL	20	-0.16	0.56	-1.56	0.59	0.32	0.06	0.13	0.42
RW	7	-0.39	1.11	-2.10	1.14	0.29	0.06	0.23	0.40
WA	1	0.60	-	0.60	0.60	0.57	-	0.57	0.57
WC	15	-0.11	0.78	-1.58	1.18	0.28	0.05	0.23	0.38
WS	12	0.43	0.54	-0.37	1.34	0.31	0.05	0.23	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	-0.25	0.64	-1.92	1.48	0.30	0.08	0.11	0.44
A1	12	0.29	0.52	-0.44	1.48	0.23	0.06	0.16	0.34
AF	20	-0.32	0.62	-1.39	1.17	0.32	0.08	0.11	0.44
MG	18	-0.10	0.59	-0.85	1.36	0.32	0.08	0.15	0.44
MR	8	-0.52	0.53	-1.39	0.22	0.32	0.05	0.25	0.41
NS	17	-0.37	0.33	-1.11	0.16	0.29	0.04	0.22	0.36
PS	13	-0.69	0.79	-1.92	0.67	0.32	0.06	0.23	0.42

[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	27	37
B	Use	31	42
C	Use	14	19
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	23	29
B	Use	31	39
C	Use	23	29
D	Review	3	4
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
ELA	VC254951	L457G007	0	1	A	S	S	S	S	S	S	S	C+
ELA	VE044990	L621J001	0	6	A	S	S	S	S	S	S	S	C+
ELA	VE046719	L6OSA705	0	79	A	S	S	S	S	S	S	S	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.E.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	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
B-	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	8		
A	71	97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62	85		
B+	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	2	3
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	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.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]	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	0	0	0	0	0	0	0	0	2	3	0	0	3	4		
A	73	91	0	0	0	0	0	0	0	0	0	0	78	98	0	0	77	96		
B+	5	6	0	0	0	0	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	80	100	80	100	80	100	80	100	80	100	0	0	80	100	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.E.8: Listing of CR Item Statistics—ELA

Accession number	VE025128
CAHSEE ID	L6OSA1428
Polyserial correlation	0.62
IRT b-value	0.60234
Step category 1[1]	2.20134
Step category 2	1.29447
Step category 3	3.50355
Step category 4	-1.19496
Step category 5	-1.16802
Step category 6	-2.72199
Step category 7	-1.91439
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	S
DIF category, White-African American	S
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.F: Summary Statistics for Operational Items—February 2011

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	-0.02	0.64	-1.37	1.20	0.45	0.08	0.22	0.57
RC	18	-0.15	0.50	-1.28	0.82	0.48	0.06	0.39	0.57
RL	20	-0.12	0.72	-1.37	1.20	0.44	0.09	0.22	0.56
RW	7	-0.02	0.50	-0.46	0.75	0.44	0.08	0.33	0.52
WA	1	0.89	-	0.89	0.89	0.74	-	0.74	0.74
WC	15	0.07	0.79	-1.28	1.13	0.42	0.08	0.30	0.52
WS	12	0.13	0.53	-0.79	1.04	0.49	0.05	0.39	0.57

[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	-0.18	0.64	-1.93	1.36	0.47	0.08	0.26	0.69
A1	12	0.31	0.52	-0.62	1.10	0.45	0.06	0.36	0.54
AF	20	-0.31	0.39	-1.01	0.43	0.49	0.08	0.35	0.69
MG	18	-0.06	0.73	-1.27	1.36	0.47	0.08	0.30	0.56
MR	8	-0.41	0.34	-0.91	0.25	0.50	0.07	0.37	0.56
NS	17	-0.25	0.65	-1.43	0.84	0.50	0.07	0.37	0.61
PS	13	-0.54	0.69	-1.93	0.40	0.41	0.09	0.26	0.53

[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	21	29
B	Use	34	47
C	Use	17	23
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.F.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	21	26
B	Use	33	41
C	Use	25	31
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.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	VC392370	L581G013	0	38	B-	A	C-	B-	C-	C-	C-	A	B-
ELA	VE339060	L7OSA1596	0	69	A	A	C-	B-	C-	C-	A	B-	A
ELA	VE046461	L6OSA1168	0	78	A	A	B-	C-	A	B-	B-	A	A
MATH	VC151330	M31263	0	74	B-	A	B-	C-	B-	B-	A	B-	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]	0	0	0	0	2	3	1	1	2	3	2	3	1	1	0	0	0	0	3	4
B-	3	4	0	0	5	7	3	4	2	3	2	3	1	1	1	1	2	3		
A	70	96	73	100	65	89	68	93	69	95	68	93	71	97	72	99	67	92		
B+	0	0	0	0	1	1	1	1	0	0	1	1	0	0	0	0	4	5		
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.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	1	1	0	0	0	0	0	0	0	0	0	0	1	1
B-	7	9	0	0	3	4	1	1	1	1	1	1	0	0	1	1	0	0		
A	72	90	80	100	77	96	78	98	78	98	79	99	79	99	79	99	80	100		
B+	1	1	0	0	0	0	0	0	1	1	0	0	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	VE025121
CAHSEE ID	L6OSA1417
Polyserial correlation	0.81
IRT b-value	0.89082
Step category 1[1]	1.9246
Step category 2	1.69534
Step category 3	3.20821
Step category 4	-0.73659
Step category 5	-0.92762
Step category 6	-2.41162
Step category 7	-2.75231
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.G: Summary Statistics for Operational Items—March 2011

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	-0.04	0.84	-2.50	1.66	0.44	0.08	0.21	0.61
RC	18	-0.02	0.67	-1.13	1.11	0.46	0.07	0.29	0.55
RL	20	-0.33	1.01	-2.50	1.66	0.41	0.11	0.21	0.58
RW	7	-0.15	0.85	-0.86	1.33	0.44	0.03	0.40	0.49
WA	1	1.11	-	1.11	1.11	0.73	-	0.73	0.73
WC	15	0.08	0.85	-1.13	1.29	0.46	0.07	0.34	0.59
WS	12	0.20	0.68	-1.30	1.14	0.46	0.09	0.32	0.61

[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	-0.18	0.73	-2.12	1.23	0.46	0.07	0.31	0.65
A1	12	0.33	0.47	-0.76	0.93	0.49	0.08	0.39	0.65
AF	20	-0.30	0.67	-1.46	1.07	0.47	0.07	0.33	0.55
MG	18	0.06	0.71	-1.70	1.14	0.47	0.06	0.35	0.57
MR	8	-0.23	0.69	-1.44	1.07	0.43	0.08	0.33	0.54
NS	17	-0.44	0.81	-2.12	1.23	0.43	0.06	0.34	0.56
PS	13	-0.47	0.68	-1.38	0.54	0.43	0.07	0.31	0.54

[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	25	34
B	Use	26	36
C	Use	21	29
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.G.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	34	43
B	Use	20	25
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.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	VE045155	L628H004	0	24	A	A	B-	B-	C-	B-	A	A	A
ELA	VE338800	L790J012	0	27	C-	A	B-	C-	B-	B-	C-	B-	A
ELA	VC017514	L10SA534	0	64	A	A	B+	A	C+	B+	A	A	A
ELA	VC017463	L10SA461	0	74	A	A	A	A	C-	A	A	A	A
ELA	VE046471	L6OSA1178	0	77	A	A	C-	A	B-	B-	A	A	A
MATH	VC151395	M32139	0	8	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]	1	1	0	0	1	1	1	1	2	3	0	0	1	1	0	0	0	0	4	5
B-	0	0	0	0	4	5	2	3	3	4	4	5	1	1	2	3	2	3		
A	72	99	73	100	65	89	70	96	65	89	66	90	71	97	70	96	70	96		
B+	0	0	0	0	3	4	0	0	2	3	3	4	0	0	1	1	1	1		
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	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	5	7

[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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
B-	2	3	0	0	5	6	1	1	2	3	3	4	0	0	2	3	2	3		
A	73	91	80	100	73	91	79	99	75	94	75	94	79	99	78	98	78	98		
B+	4	5	0	0	2	3	0	0	3	4	2	3	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 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	VE025132
CAHSEE ID	L6OSA1434
Polyserial correlation	0.83
IRT b-value	1.11269
Step category 1[1]	2.91546
Step category 2	1.5859
Step category 3	3.40633
Step category 4	-0.9865
Step category 5	-1.10151
Step category 6	-2.79953
Step category 7	-3.02014
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 2011

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.	Minim um	Maximum
Overall	73	-0.03	0.56	-1.42	1.20	0.40	0.07	0.21	0.51
RC	18	-0.23	0.55	-1.42	0.90	0.40	0.06	0.28	0.51
RL	20	-0.11	0.63	-1.24	1.20	0.40	0.06	0.21	0.48
RW	7	-0.08	0.40	-0.59	0.43	0.40	0.08	0.27	0.49
WA	1	0.73	-	0.73	0.73	0.68	-	0.68	0.68
WC	15	0.05	0.63	-1.39	1.17	0.38	0.07	0.23	0.47
WS	12	0.27	0.27	-0.23	0.70	0.42	0.07	0.33	0.50

[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	-0.18	0.58	-1.37	1.07	0.35	0.08	0.16	0.49
A1	12	0.26	0.42	-0.26	1.07	0.31	0.07	0.16	0.42
AF	20	-0.30	0.49	-1.34	0.62	0.39	0.08	0.22	0.49
MG	18	-0.04	0.62	-1.18	0.88	0.36	0.09	0.16	0.49
MR	8	0.11	0.44	-0.46	0.62	0.32	0.11	0.16	0.45
NS	17	-0.28	0.67	-1.37	0.65	0.32	0.08	0.16	0.43
PS	13	-0.42	0.41	-1.08	0.36	0.36	0.07	0.27	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.H.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	14	19
B	Use	22	30
C	Use	34	47
D	Review	3	4
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	17	21
B	Use	28	35
C	Use	33	41
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	VC392117	L563G005	0	3	A	A	C-	S	B-	B-	A	A	B-
ELA	VE044540	L603J009	0	11	A	A	C-	S	B-	C-	B-	B-	B-
ELA	VC392671	L5OSA620	0	65	B-	A	C-	S	C-	C-	B-	C-	A
ELA	VE046945	L6OSA931	0	66	A	A	C+	S	A	A	A	A	A
ELA	VE046773	L6OSA759	0	69	A	A	C+	S	A	B+	A	A	A
ELA	VC392641	L5OSA589	0	76	A	A	A	S	B-	A	A	A	C-

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	3	4	0	0	1	1	2	3	0	0	1	1	1	1	4	5
B-	2	3	0	0	2	3	0	0	4	5	1	1	2	3	1	1	2	3		
A	71	97	73	100	62	85	0	0	66	90	67	92	71	97	71	97	69	95		
B+	0	0	0	0	4	5	0	0	2	3	3	4	0	0	0	0	1	1		
C+[2]	0	0	0	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0	2	3
Small N[3]	0	0	0	0	0	0	73	100	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	6	8

[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	3	4	0	0	4	5	2	3	0	0	0	0	3	4		
A	78	98	79	99	76	95	0	0	74	93	78	98	80	100	80	100	77	96		
B+	0	0	1	1	1	1	0	0	2	3	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	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	VE025114
CAHSEE ID	L6OSA1408
Polyserial correlation	0.74
IRT b-value	0.73246
Step category 1[1]	2.03057
Step category 2	1.38129
Step category 3	3.37645
Step category 4	-0.89260
Step category 5	-0.72564
Step category 6	-2.57388
Step category 7	-2.59618
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	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.

[3] S indicates that DIF analysis was not performed due to insufficient sample size.

Appendix 6.I: Summary Statistics for Field-Test Items—February 2011

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	7	-0.01	1.09	-1.94	1.63	0.42	0.10	0.27	0.52
RL	5	-0.17	1.30	-1.94	1.63	0.42	0.10	0.27	0.51
RW	1	0.32	-	0.32	0.32	0.31	-	0.31	0.31
WC	1	0.46	-	0.46	0.46	0.52	-	0.52	0.52

[1] RL = Literary Response & Analysis, RW = Word Analysis, WC = Writing Conventions,

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	7	0.70	0.69	-0.22	1.53	0.41	0.09	0.26	0.51
A1	2	1.06	0.45	0.74	1.38	0.40	0.01	0.39	0.41
AF	2	-0.21	0.01	-0.22	-0.20	0.49	0.03	0.47	0.51
MG	3	1.07	0.39	0.82	1.53	0.35	0.10	0.26	0.46
MR	1	0.82	-	0.82	0.82	0.26	-	0.26	0.26

[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.

Table 6.I.3: IRT Model Data Fit Distribution of Field-Test Items—ELA

IRT review category	Status	Item count	Percent[1]
B	Use	4	57
C	Use	3	43
Total		7	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	1	13
B	Use	4	50
C	Use	2	25
N/A	Dropped from Calibration	1	13
Total		8	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
NONE													

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]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
A	7	100	7	100	6	86	7	100	6	86	6	86	7	100	7	100	6	86		
B+	0	0	0	0	1	14	0	0	1	14	1	14	0	0	0	0	1	14		
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]	7	100	7	100	7	100	7	100	7	100	7	100	7	100	7	100	7	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 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]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	1	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
A	7	88	8	100	8	100	8	100	8	100	8	100	8	100	8	100	8	100		
B+	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	8	100	8	100	8	100	8	100	8	100	8	100	8	100	8	100	8	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 percentages may differ slightly from the sum of the category percentages due to rounding.

Appendix 6.J: Summary Statistics for Field-Test Items—March 2011

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	475	0.66	0.90	-2.10	3.36	0.39	0.11	0.09	0.59
RC	158	0.75	0.75	-0.83	3.36	0.40	0.11	0.12	0.55
RL	120	0.70	0.89	-1.60	2.97	0.38	0.10	0.12	0.59
RW	48	0.62	0.99	-1.38	2.63	0.39	0.11	0.12	0.54
WC	83	0.18	1.03	-2.10	3.31	0.41	0.09	0.15	0.55
WS	66	0.99	0.79	-0.61	2.83	0.34	0.12	0.09	0.58

[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	5	1.25	0.85	0.41	2.60	0.43	0.09	0.31	0.55
A1	2	0.75	0.48	0.41	1.08	0.45	0.05	0.41	0.48
AF	1	1.45	-	1.45	1.45	0.31	-	0.31	0.31
MG	2	1.64	1.35	0.69	2.60	0.47	0.12	0.38	0.55

[1] A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry.

Table 6.J.3: IRT Model Data Fit Distribution of Field-Test Items—ELA

IRT review category	Status	Item count	Percent[1]
A	Use	137	28
B	Use	121	25
C	Use	188	38
D	Review	16	3
F	Do not use	13	3
N/A	Dropped from Calibration	18	4
Total		493	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	1	13
C	Use	4	50
N/A	Dropped from Calibration	3	38
Total		8	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	VE045853	L665H006	1	49	C+	S	A	S	S	A	A	A	A
ELA	VE045858	L665H011	1	50	A	S	C+	S	S	B+	A	A	A
ELA	VE044566	L605J005	5	49	A	S	B-	S	S	C-	A	A	A
ELA	VE044575	L606J001	6	46	A	S	C-	S	S	C-	C-	C-	A
ELA	FM007575	L110B004	8	46	A	S	C-	S	S	C-	A	A	A
ELA	FM007579	L110B009	8	48	A	S	B-	S	S	C-	A	A	A
ELA	VE044725	L612J002	9	46	C-	S	A	S	S	A	C-	C-	A
ELA	VE044841	L616J002	15	46	C-	S	A	S	S	B-	A	A	A
ELA	VE044846	L616J007	16	50	A	S	B-	S	S	B-	C-	B-	A
ELA	VE046441	L6OSA1148	17	75	A	S	C-	S	S	C-	A	B-	A
ELA	VE045145	L627J008	22	46	A	S	C-	S	S	C-	C-	C-	B-
ELA	VC020868	L368F007	23	48	A	S	A	S	S	A	A	A	C-
ELA	VE045294	L633J007	25	50	A	S	A	S	C-	B-	A	A	A
ELA	VE045425	L639H001	27	46	A	S	C-	S	S	C-	B-	A	A
ELA	VE045432	L639H008	27	48	A	S	C+	S	S	C+	B+	A	A
ELA	VC019310	L247E004	29	50	A	S	C-	S	S	C-	A	A	A
ELA	VC017824	L162A003	35	46	B-	S	A	S	S	B-	C-	C-	B-
ELA	VC017831	L162A008	35	47	A	S	A	S	S	A	A	A	C-
ELA	VE046520	L6OSA1227	37	75	A	S	C+	S	S	C+	A	A	A
ELA	VE045479	L641H001	39	46	B-	S	B-	S	S	B-	C-	A	C-
ELA	VC172582	L396F009	40	49	A	S	C-	S	S	B-	A	A	A
ELA	VE046603	L6OSA1311	40	75	A	S	A	S	S	A	A	A	C-
ELA	VE046616	L6OSA1324	41	75	A	S	C-	S	S	C-	A	B-	A
ELA	VC017872	L168A009	43	50	A	S	C-	S	S	C-	B-	A	A
ELA	VE046516	L6OSA1223	44	75	A	S	C+	S	S	B+	A	A	A
ELA	VC020717	L351F006	48	49	A	S	C-	S	S	C-	C-	B-	C-
ELA	VC021073	L380F007	50	49	B-	S	B-	S	S	B-	C-	C-	A
ELA	VC020239	L302F003	55	47	A	S	B-	S	C-	B-	A	A	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	VE046439	L6OSA1146	58	75	A	S	A	S	S	A	A	A	C-
ELA	VE046514	L6OSA1221	62	75	A	S	C+	S	S	C+	A	A	A
ELA	VC018595	L225D001	63	46	B-	S	C+	S	S	C+	A	A	A
ELA	VC018596	L225D002	64	47	A	S	B-	S	S	B-	B-	A	C-
ELA	VE046609	L6OSA1317	65	75	A	S	C-	S	S	C-	A	A	A
ELA	VE046612	L6OSA1320	66	75	A	S	A	S	S	A	A	A	C-
ELA	VE045186	L629H007	68	46	A	S	B-	S	S	B-	A	A	C-
ELA	VE046605	L6OSA1313	68	75	A	S	C-	S	S	B-	A	B-	A
ELA	VE046606	L6OSA1314	69	75	A	S	A	S	S	A	C-	A	A
ELA	VE046611	L6OSA1319	70	75	A	S	A	S	S	A	A	A	C-
ELA	VC018172	L207E001	71	46	C-	S	A	S	S	A	A	A	A
ELA	VE046527	L6OSA1234	72	75	A	S	C+	S	S	C+	A	A	A
ELA	VC017945	L175A013	73	49	A	S	B+	S	S	C+	A	A	A
ELA	VE046617	L6OSA1325	74	75	A	S	C-	S	S	C-	B-	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.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-	3	1	0	0	13	3	0	0	2	0	13	3	9	2	5	1	10	2	33	7
B-	10	2	0	0	32	6	0	0	3	1	38	8	24	5	15	3	18	4		
A	457	93	0	0	410	83	0	0	32	6	410	83	458	93	470	95	465	94		
B+	22	4	0	0	31	6	0	0	1	0	26	5	2	0	3	1	0	0		
C+	1	0	0	0	7	1	0	0	0	0	6	1	0	0	0	0	0	0	9.	2
Small N[2]	0	0	493	100	0	0	493	100	455	92	0	0	0	0	0	0	0	0		
Total[3]	493	100	493	100	493	100	493	100	493	100	493	100	493	100	493	100	493	100	42	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-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
A	8	100	8	100	7	88	8	100	8	100	8	100	8	100	8	100	8	100		
B+	0	0	0	0	1	13	0	0	0	0	0	0	0	0	0	0	0	0		
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[3]	8	100	8	100	8	100	8	100	8	100	8	100	8	100	8	100	8	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] 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 2010

	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	0.62	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.61	0.30	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.83	0.40	0.46	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.84	0.40	0.49	0.65	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.72	0.33	0.35	0.52	0.54	1.00	-	-	-	-	-	-	-
Writing Conventions	0.75	0.39	0.37	0.51	0.51	0.47	1.00	-	-	-	-	-	-
Mathematics	0.58	0.35	0.37	0.50	0.51	0.47	0.48	1.00	-	-	-	-	-
Probability and Statistics	0.51	0.29	0.32	0.45	0.45	0.41	0.42	0.73	1.00	-	-	-	-
Number Sense	0.46	0.28	0.30	0.40	0.41	0.35	0.38	0.76	0.46	1.00	-	-	-
Algebra & Functions	0.51	0.32	0.33	0.44	0.44	0.41	0.43	0.83	0.53	0.52	1.00	-	-
Measurement & Geometry	0.46	0.27	0.29	0.40	0.41	0.41	0.38	0.79	0.46	0.49	0.54	1.00	-
Algebra 1	0.38	0.24	0.23	0.32	0.32	0.32	0.32	0.67	0.37	0.40	0.48	0.44	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	47.91	2.07	4.48	9.99	11.21	5.09	7.83	36.59	6.87	8.21	9.67	7.58	4.26
SD	12.11	0.52	1.51	3.36	3.61	2.37	2.93	10.84	2.47	2.87	3.38	3.29	2.16
Reliability	0.87	0.26	0.36	0.68	0.69	0.54	0.64	0.86	0.54	0.58	0.62	0.65	0.47
SEM	4.37	0.45	1.21	1.90	2.02	1.60	1.76	4.12	1.67	1.87	2.08	1.94	1.57

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 8,389 examinees completing the ELA section.
Correlations for the mathematics section are reported for 8,222 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 3,505 examinees taking both sections.

Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2010

	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	0.64	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.65	0.35	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.85	0.43	0.55	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.87	0.44	0.54	0.69	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.77	0.38	0.39	0.58	0.60	1.00	-	-	-	-	-	-	-
Writing Conventions	0.79	0.42	0.40	0.56	0.59	0.59	1.00	-	-	-	-	-	-
Mathematics	0.64	0.39	0.42	0.56	0.55	0.54	0.56	1.00	-	-	-	-	-
Probability and Statistics	0.59	0.36	0.41	0.53	0.51	0.49	0.50	0.78	1.00	-	-	-	-
Number Sense	0.53	0.33	0.36	0.47	0.45	0.43	0.46	0.81	0.56	1.00	-	-	-
Algebra & Functions	0.58	0.35	0.37	0.50	0.49	0.49	0.51	0.87	0.62	0.62	1.00	-	-
Measurement & Geometry	0.54	0.32	0.34	0.47	0.46	0.45	0.48	0.84	0.58	0.58	0.65	1.00	-
Algebra1	0.45	0.28	0.26	0.38	0.38	0.39	0.42	0.73	0.43	0.50	0.57	0.55	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	50.17	2.08	4.57	10.15	11.52	6.22	8.34	38.48	6.82	8.88	10.36	7.70	4.73
SD	13.65	0.58	1.62	3.52	3.92	2.67	3.15	13.03	2.73	3.20	4.00	3.65	2.43
Reliability	0.89	0.28	0.48	0.70	0.74	0.66	0.70	0.90	0.64	0.65	0.73	0.72	0.58
SEM	4.43	0.49	1.17	1.93	2.02	1.56	1.72	4.10	1.64	1.88	2.06	1.92	1.57

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 43,194 examinees completing the ELA section.
Correlations for the mathematics section are reported for 41,519 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 24,297 examinees taking both sections.

Table 6.K.3: Intercorrelations and Reliability Estimates by Section—November 2010

	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	0.67	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.69	0.38	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.86	0.47	0.58	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.86	0.47	0.56	0.68	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.80	0.42	0.48	0.62	0.62	1.00	-	-	-	-	-	-	-
Writing Conventions	0.79	0.46	0.45	0.56	0.57	0.60	1.00	-	-	-	-	-	-
Mathematics	0.67	0.46	0.45	0.56	0.58	0.56	0.61	1.00	-	-	-	-	-
Probability and Statistics	0.61	0.42	0.43	0.53	0.52	0.50	0.53	0.79	1.00	-	-	-	-
Number Sense	0.57	0.40	0.41	0.48	0.49	0.47	0.52	0.84	0.62	1.00	-	-	-
Algebra & Functions	0.62	0.42	0.42	0.51	0.54	0.52	0.55	0.87	0.64	0.64	1.00	-	-
Measurement & Geometry	0.55	0.38	0.36	0.45	0.47	0.47	0.50	0.83	0.57	0.59	0.65	1.00	-
Algebra1	0.48	0.35	0.30	0.38	0.41	0.41	0.47	0.75	0.47	0.55	0.58	0.55	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	51.58	2.13	4.65	10.72	12.39	6.43	7.81	39.68	6.94	8.67	10.85	8.42	4.81
SD	13.92	0.57	1.57	3.71	3.73	2.85	3.13	13.37	2.66	3.55	3.93	3.53	2.59
Reliability	0.90	0.32	0.46	0.74	0.72	0.71	0.70	0.91	0.65	0.73	0.73	0.70	0.64
SEM	4.34	0.47	1.15	1.89	1.98	1.54	1.72	4.06	1.57	1.84	2.04	1.94	1.56

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 102,616 examinees completing the ELA section.
Correlations for the mathematics section are reported for 98,156 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 60,723 examinees taking both sections.

Table 6.K.4: Intercorrelations and Reliability Estimates by Section—December 2010

	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	0.57	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.60	0.27	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.83	0.34	0.44	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.84	0.37	0.46	0.63	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.75	0.33	0.38	0.56	0.55	1.00	-	-	-	-	-	-	-
Writing Conventions	0.74	0.34	0.37	0.52	0.49	0.47	1.00	-	-	-	-	-	-
Mathematics	0.68	0.44	0.47	0.60	0.59	0.55	0.51	1.00	-	-	-	-	-
Probability and Statistics	0.65	0.42	0.43	0.58	0.59	0.50	0.48	0.74	1.00	-	-	-	-
Number Sense	0.54	0.36	0.40	0.48	0.48	0.44	0.40	0.78	0.50	1.00	-	-	-
Algebra & Functions	0.59	0.37	0.40	0.51	0.51	0.48	0.46	0.86	0.58	0.57	1.00	-	-
Measurement & Geometry	0.58	0.38	0.42	0.52	0.50	0.48	0.42	0.81	0.50	0.52	0.62	1.00	-
Algebra1	0.38	0.24	0.25	0.34	0.33	0.33	0.29	0.63	0.33	0.40	0.48	0.42	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	50.36	2.08	4.35	10.62	11.82	5.57	8.63	37.86	7.37	8.39	9.80	7.98	4.32
SD	11.72	0.49	1.43	3.41	3.55	2.42	2.68	11.50	2.62	3.08	3.56	3.40	2.04
Reliability	0.86	0.21	0.38	0.68	0.67	0.56	0.58	0.87	0.63	0.60	0.67	0.68	0.39
SEM	4.32	0.43	1.13	1.94	2.04	1.60	1.74	4.13	1.59	1.96	2.06	1.93	1.59

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 2,405 examinees completing the ELA section.
Correlations for the mathematics section are reported for 2,433 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 924 examinees taking both sections.

Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 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	0.74	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.79	0.50	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.91	0.57	0.69	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.90	0.58	0.70	0.79	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.88	0.57	0.65	0.77	0.74	1.00	-	-	-	-	-	-	-
Writing Conventions	0.85	0.57	0.63	0.71	0.70	0.73	1.00	-	-	-	-	-	-
Mathematics	0.80	0.57	0.62	0.74	0.72	0.72	0.70	1.00	-	-	-	-	-
Probability and Statistics	0.71	0.50	0.55	0.66	0.64	0.64	0.62	0.84	1.00	-	-	-	-
Number Sense	0.73	0.51	0.57	0.68	0.67	0.65	0.64	0.91	0.72	1.00	-	-	-
Algebra & Functions	0.77	0.54	0.59	0.72	0.69	0.70	0.67	0.93	0.75	0.81	1.00	-	-
Measurement & Geometry	0.73	0.52	0.57	0.67	0.65	0.66	0.64	0.92	0.71	0.78	0.81	1.00	-
Algebra 1	0.65	0.48	0.49	0.59	0.56	0.58	0.57	0.85	0.63	0.72	0.74	0.75	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	61.25	2.37	4.92	13.04	14.28	8.14	10.20	51.56	9.12	11.14	13.42	11.20	6.69
SD	16.64	0.68	1.76	4.14	4.21	3.04	3.25	17.30	2.79	4.31	4.81	4.24	3.11
Reliability	0.94	0.43	0.63	0.84	0.82	0.79	0.76	0.95	0.73	0.85	0.86	0.83	0.77
SEM	4.10	0.51	1.07	1.64	1.76	1.39	1.58	3.71	1.46	1.65	1.83	1.76	1.50

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 154,910 examinees completing the ELA section.
Correlations for the mathematics section are reported for 155,492 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 131,500 examinees taking both sections.

Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 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	0.73	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.79	0.50	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.91	0.56	0.70	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.89	0.56	0.69	0.78	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.86	0.54	0.63	0.75	0.71	1.00	-	-	-	-	-	-	-
Writing Conventions	0.87	0.58	0.63	0.72	0.70	0.73	1.00	-	-	-	-	-	-
Mathematics	0.79	0.56	0.61	0.70	0.68	0.68	0.71	1.00	-	-	-	-	-
Probability and Statistics	0.72	0.49	0.57	0.65	0.64	0.62	0.64	0.84	1.00	-	-	-	-
Number Sense	0.69	0.49	0.55	0.61	0.61	0.59	0.62	0.89	0.71	1.00	-	-	-
Algebra & Functions	0.74	0.52	0.58	0.66	0.64	0.64	0.67	0.93	0.74	0.78	1.00	-	-
Measurement & Geometry	0.70	0.50	0.55	0.62	0.60	0.61	0.63	0.91	0.70	0.74	0.80	1.00	-
Algebra 1	0.65	0.48	0.49	0.57	0.55	0.57	0.59	0.87	0.64	0.71	0.77	0.75	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	64.60	2.35	5.35	13.53	15.58	8.61	10.95	55.23	9.60	12.36	14.25	11.73	7.30
SD	15.04	0.60	1.57	3.78	3.56	2.77	3.30	16.28	2.80	3.56	4.40	4.25	3.25
Reliability	0.94	0.42	0.62	0.82	0.79	0.76	0.80	0.95	0.75	0.80	0.84	0.83	0.80
SEM	3.81	0.46	0.97	1.60	1.62	1.36	1.46	3.59	1.39	1.59	1.75	1.74	1.44

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 402,061 examinees completing the ELA section.
Correlations for the mathematics section are reported for 401,838 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 371,516 examinees taking both sections.

Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 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	0.75	0.42	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.88	0.50	0.63	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.90	0.51	0.65	0.78	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.82	0.45	0.57	0.66	0.67	1.00	-	-	-	-	-	-	-
Writing Conventions	0.81	0.49	0.56	0.63	0.64	0.65	1.00	-	-	-	-	-	-
Mathematics	0.70	0.48	0.54	0.62	0.63	0.63	0.62	1.00	-	-	-	-	-
Probability and Statistics	0.65	0.43	0.50	0.58	0.59	0.56	0.56	0.80	1.00	-	-	-	-
Number Sense	0.58	0.41	0.45	0.51	0.52	0.51	0.51	0.81	0.58	1.00	-	-	-
Algebra & Functions	0.64	0.43	0.49	0.57	0.57	0.57	0.56	0.90	0.67	0.64	1.00	-	-
Measurement & Geometry	0.61	0.41	0.46	0.53	0.53	0.56	0.53	0.86	0.60	0.59	0.70	1.00	-
Algebra 1	0.49	0.34	0.36	0.42	0.42	0.46	0.44	0.74	0.48	0.50	0.60	0.60	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	48.64	2.08	3.87	10.51	11.22	5.80	7.89	38.27	6.93	8.45	10.15	8.09	4.65
SD	15.95	0.64	1.82	3.97	4.34	3.05	3.34	13.75	2.90	3.35	4.26	3.64	2.42
Reliability	0.92	0.34	0.57	0.78	0.79	0.74	0.73	0.91	0.68	0.69	0.77	0.73	0.57
SEM	4.46	0.52	1.19	1.87	1.97	1.54	1.72	4.08	1.64	1.86	2.02	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 39,105 examinees completing the ELA section.
Correlations for the mathematics section are reported for 38,851 examinees completing the mathematics section.
Correlations between ELA and mathematics are reported for 20,185 examinees taking both sections.

Table 6.K.8: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—February 2011

	Total MC N=154,910		Composite (MC + Essay) N=154,910	
	R_{XX}	SEM	R_{XX}	SEM
Gender				
Male	0.95	3.46	0.94	4.21
Female	0.94	3.32	0.94	3.95
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.45	0.93	4.17
Asian	0.95	3.14	0.94	3.95
Pacific Islander	0.93	3.50	0.93	4.12
Filipino	0.93	3.14	0.92	3.84
Hispanic or Latino	0.93	3.64	0.93	4.23
African American	0.94	3.62	0.93	4.33
White (not of Hispanic origin)	0.94	3.01	0.93	3.86
Two or More Races	0.95	3.27	0.94	4.08
Language Fluency				
English Proficient Students	0.94	3.21	0.93	3.98
English Learner Students	0.88	3.93	0.87	4.44
Special Education Program Participation				
Students Receiving Services	0.92	3.84	0.91	4.60
Students Not Receiving Services	0.94	3.33	0.93	4.02

Table 6.K.9: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—February 2011

	Total N=155,492	
	R_{XX}	SEM
Gender		
Male	0.96	3.68
Female	0.95	3.73
Race/Ethnicity		
American Indian or Alaska Native	0.95	3.83
Asian	0.95	3.06
Pacific Islander	0.94	3.81
Filipino	0.95	3.41
Hispanic or Latino	0.94	3.94
African American	0.94	3.98
White (not of Hispanic origin)	0.95	3.43
Two or More Races	0.96	3.67
Language Fluency		
English Proficient Students	0.95	3.60
English Learner Students	0.91	4.09
Special Education Program Participation		
Students Receiving Services	0.92	4.06
Students Not Receiving Services	0.95	3.67

Table 6.K.10: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for ELA—March 2011

	Total MC N=402,061		Composite (MC + Essay) N=402,061	
	R_{XX}	SEM	R_{XX}	SEM
Gender				
Male	0.94	3.27	0.94	3.88
Female	0.93	3.12	0.93	3.71
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.23	0.94	3.86
Asian	0.95	2.77	0.94	3.59
Pacific Islander	0.93	3.26	0.93	3.85
Filipino	0.92	2.89	0.91	3.60
Hispanic or Latino	0.93	3.38	0.93	3.88
African American	0.94	3.43	0.94	4.00
White (not of Hispanic origin)	0.93	2.85	0.92	3.59
Two or More Races	0.95	3.10	0.94	3.80
Language Fluency				
English Proficient Students	0.93	3.03	0.92	3.69
English Learner Students	0.90	3.79	0.90	4.19
Special Education Program Participation				
Students Receiving Services	0.93	3.77	0.92	4.32
Students Not Receiving Services	0.93	3.13	0.92	3.74

Table 6.K.11: Reliabilities (R_{XX}) and Standard Errors of Measurement (SEMs) of Subgroups for Mathematics—March 2011

	Total N=401,838	
	R_{XX}	SEM
Gender		
Male	0.96	3.56
Female	0.95	3.61
Race/Ethnicity		
American Indian or Alaska Native	0.95	3.70
Asian	0.95	2.80
Pacific Islander	0.94	3.66
Filipino	0.94	3.26
Hispanic or Latino	0.94	3.76
African American	0.94	3.87
White (not of Hispanic origin)	0.95	3.32
Two or More Races	0.96	3.55
Language Fluency		
English Proficient Students	0.95	3.48
English Learner Students	0.92	4.02
Special Education Program Participation		
Students Receiving Services	0.93	4.04
Students Not Receiving Services	0.95	3.54

Appendix 6.L: Rater Agreement Analyses

Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2010

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	222	0	0	0	0	222	3
1	0	211	165	0	0	376	4
2	0	143	5,614	710	11	6,478	77
3	0	0	659	555	31	1,245	15
4	0	0	7	37	23	67	1
Total	222	354	6,445	1,302	65	8,388	100
Percent	3	4	77	16	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	6,625	79	79
1	1,745	21	100
2	18	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	2.07	2.08
Standard Deviation	0.57	0.57
Mean Absolute Difference Between First and Second Ratings: 0.21		
Correlation of First and Second Ratings: 0.67		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	142
Illegible	0
Off Topic	69
Cartoon / Inappropriate	9
Not in English	2
Total	**222**

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 2010

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	1,439	0	0	0	0	1,439	3
1	0	1,131	785	10	0	1,926	4
2	0	775	27,476	3,643	77	31,971	74
3	0	7	3,656	3,260	303	7,226	17
4	0	0	85	297	232	614	1
Total	1,439	1,913	32,002	7,210	612	43,176	100
Percent	3	4	74	17	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	33,538	78	78
1	9,459	22	100
2	179	0	100
3	0	0	100
4	0	0	100

Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2010

	First Rating	Second Rating
Mean	2.08	2.08
Standard Deviation	0.63	0.63
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.70		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	759
Illegible	0
Off Topic	633
Cartoon / Inappropriate	33
Not in English	16
Total	**1,441**

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 2010

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	2,383	0	0	0	0	2,383	2
1	0	3,217	2,197	29	1	5,444	5
2	0	2,230	60,678	9,575	234	72,717	71
3	0	23	9,871	9,318	906	20,118	20
4	0	0	306	987	619	1,912	2
Total	2,383	5,470	73,052	19,909	1,760	102,574	100
Percent	2	5	71	19	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	76,215	74	74
1	25,766	25	99
2	592	1	100
3	1	0	100
4	0	0	100

Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2010

	First Rating	Second Rating
Mean	2.13	2.13
Standard Deviation	0.63	0.63
Mean Absolute Difference Between First and Second Ratings: 0.26		
Correlation of First and Second Ratings: 0.65		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	1,490
Illegible	0
Off Topic	781
Cartoon / Inappropriate	71
Not in English	41
Total	**2,383**

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 2010

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	42	0	0	0	0	42	2
1	0	63	52	1	0	116	5
2	0	33	1,648	210	2	1,893	79
3	0	1	184	134	13	332	14
4	0	0	2	6	12	20	1
Total	42	97	1,886	351	27	2,403	100
Percent	2	4	78	15	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	1,899	79	79
1	498	21	100
2	6	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	2.07	2.09
Standard Deviation	0.53	0.54
Mean Absolute Difference Between First and Second Ratings: 0.21		
Correlation of First and Second Ratings: 0.62		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	19
Illegible	0
Off Topic	23
Cartoon / Inappropriate	0
Not in English	0
Total	**42**

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 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	2,892	0	0	0	0	2,892	2
1	0	3,351	2,636	47	3	6,037	4
2	0	2,567	65,628	16,764	841	85,800	55
3	0	70	17,012	28,301	4,946	50,329	33
4	0	1	859	5,238	3,694	9,792	6
Total	2,892	5,989	86,135	50,350	9,484	154,850	100
Percent	2	4	56	33	6	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	103,866	67	67
1	49,163	32	99
2	1,817	1	100
3	4	0	100
4	0	0	100

Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2011

	First Rating	Second Rating
Mean	2.38	2.37
Standard Deviation	0.74	0.74
Mean Absolute Difference Between First and Second Ratings: 0.34		
Correlation of First and Second Ratings: 0.67		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	2,055
Illegible	0
Off Topic	714
Cartoon / Inappropriate	91
Not in English	32
Total	**2,892**

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 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	3,644	0	0	0	0	3,644	1
1	0	8,935	5,609	64	3	14,611	4
2	0	5,494	186,333	44,438	1,592	237,857	59
3	0	52	44,741	73,473	9,708	127,974	32
4	0	4	1,607	10,094	6,126	17,831	4
Total	3,644	14,485	238,290	128,069	17,429	401,917	100
Percent	1	4	59	32	4	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	278,511	69	69
1	120,084	30	99
2	3,315	1	100
3	7	0	100
4	0	0	100

Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2011

	First Rating	Second Rating
Mean	2.35	2.35
Standard Deviation	0.67	0.66
Mean Absolute Difference Between First and Second Ratings: 0.32		
Correlation of First and Second Ratings: 0.63		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	2,365
Illegible	118
Off Topic	1,009
Cartoon / Inappropriate	98
Not in English	55
Total	**3,645**

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 2011

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	1,423	0	0	0	0	1,423	4
1	0	1,821	959	9	1	2,790	7
2	0	939	22,799	2,994	69	26,801	69
3	0	15	3,164	3,712	380	7,271	19
4	0	2	62	419	316	799	2
Total	1,423	2,777	26,984	7,134	766	39,084	100
Percent	4	7	69	18	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	30,071	77	77
1	8,855	23	100
2	155	0	100
3	3	0	100
4	0	0	100

Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2011

	First Rating	Second Rating
Mean	2.08	2.08
Standard Deviation	0.69	0.69
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.74		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	1,014
Illegible	0
Off Topic	348
Cartoon / Inappropriate	48
Not in English	15
Total	**1,425**

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 2010

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	8,298	0.54290	0.21736	66.76
Rater (R)	1	0.31233	0.00002	0.00
PR,*e*	8,298	0.10818	0.10818	33.23
Generalizability Coefficient				0.80
Dependability Coefficient				0.80

Table 6.M.2: Generalizability Results—October 2010

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	41,731	0.66493	0.27328	69.78
Rater (R)	1	0.00202	0.00000	0.00
PR,*e*	41,731	0.11837	0.11837	30.22
Generalizability Coefficient				0.82
Dependability Coefficient				0.82

Table 6.M.3: Generalizability Results—November 2010

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	99,175	0.64750	0.25455	64.78
Rater (R)	1	1.24030	0.00001	0.00
PR,*e*	99,175	0.13841	0.13841	35.22
Generalizability Coefficient				0.79
Dependability Coefficient				0.79

Table 6.M.4: Generalizability Results—December 2010

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	2,369	0.45965	0.17548	61.71
Rater (R)	1	0.52743	0.00018	0.06
PR,e	2,369	0.10868	0.10868	38.22
Generalizability Coefficient				0.76
Dependability Coefficient				0.76

Table 6.M.5: Generalizability Results—February 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	151,908	0.90275	0.35951	66.18
Rater (R)	1	1.03220	0.00000	0.00
PR,*e*	151,908	0.18372	0.18372	33.82
Generalizability Coefficient				0.80
Dependability Coefficient				0.80

Table 6.M.6: Generalizability Results—March 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	395,704	0.71202	0.27249	62.00
Rater (R)	1	0.53550	0.00000	0.00
PR,*e*	395,704	0.16704	0.16704	38.00
Generalizability Coefficient				0.77
Dependability Coefficient				0.77

Table 6.M.7: Generalizability Results—May 2011

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	37,918	0.81925	0.34859	74.06
Rater (R)	1	0.61521	0.00001	0.00
PR,*e*	37,918	0.12206	0.12206	25.93
Generalizability Coefficient				0.85
Dependability Coefficient				0.85

Appendix 6.N: Decision Classification Reliability Analyses

Table 6.N.1: ESEA Reliability Classifications—July 2010

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.00	0.01	0.01	0.01
Proficient (69-76)	0.00	0.02	0.02	0.03
Below Proficient (0-68)	0.00	0.01	0.95	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 (77-90)	0.00	0.01	0.01	0.01
Proficient (69-76)	0.00	0.02	0.02	0.03
Below Proficient (0-68)	0.00	0.02	0.94	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

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	0.95	0.96
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				1.00
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	0.96
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				1.00
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.97

Table 6.N.2: Pass/Not Pass Classifications—July 2010

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.23	0.04	0.27
0-54	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
55-90	0.22	0.05	0.27
0-54	0.08	0.65	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	0.21	0.04	0.25
0-42	0.04	0.71	0.75
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Score	Pass
43-80	0.20	0.05	0.25
0-42	0.06	0.68	0.75
Estimated Proportion Consistently Classified			0.89

Table 6.N.3: ESEA Reliability Classifications—October 2010

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.01	0.01	0.01	0.03
Proficient (69-76)	0.00	0.03	0.02	0.05
Below Proficient (0-68)	0.00	0.02	0.90	0.92
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)	0.01	0.01	0.01	0.03
Proficient (69-76)	0.00	0.03	0.02	0.05
Below Proficient (0-68)	0.00	0.03	0.88	0.92
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	0.00	0.02
Proficient (59-72)	0.00	0.05	0.01	0.06
Below Proficient (0-58)	0.00	0.01	0.91	0.93
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 (73-80)	0.01	0.01	0.00	0.02
Proficient (59-72)	0.00	0.05	0.01	0.06
Below Proficient (0-58)	0.00	0.02	0.90	0.93
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.4: Pass/Not Pass Classifications—October 2010

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.32	0.04	0.36
0-54	0.05	0.59	0.64
Estimated Proportion Correctly Classified			0.90

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.30	0.06	0.36
0-54	0.08	0.56	0.64
Estimated Proportion Consistently Classified			0.87

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
44-80	0.26	0.04	0.30
0-43	0.04	0.67	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
44-80	0.25	0.05	0.30
0-43	0.06	0.65	0.70
Estimated Proportion Consistently Classified			0.90

Table 6.N.5: ESEA Reliability Classifications—November 2010

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	0.02	0.01	0.04
Proficient (69-76)	0.00	0.04	0.02	0.06
Below Proficient (0-68)	0.02	0.02	0.87	0.90
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.95

[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	0.04	0.02	0.06
Below Proficient (0-68)	0.00	0.04	0.86	0.90
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
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)	0.00	0.01	0.01	0.02
Proficient (59-72)	0.00	0.05	0.02	0.07
Below Proficient (0-58)	0.00	0.01	0.90	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 (73-80)	0.00	0.01	0.01	0.02
Proficient (59-72)	0.00	0.05	0.02	0.07
Below Proficient (0-58)	0.00	0.02	0.89	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.6: Pass/Not Pass Classifications—November 2010

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.34	0.05	0.39
0-55	0.05	0.57	0.61
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	0.32	0.06	0.39
0-55	0.07	0.54	0.61
Estimated Proportion Consistently Classified			0.87

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
44-80	0.31	0.02	0.33
0-43	0.07	0.60	0.67
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
44-80	0.29	0.04	0.33
0-43	0.09	0.58	0.67
Estimated Proportion Consistently Classified			0.87

Table 6.N.7: ESEA Reliability Classifications—December 2010

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 (69-75)	0.00	0.02	0.02	0.03
Below Proficient (0-68)	0.00	0.01	0.94	0.95
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 (76-90)	0.00	0.01	0.01	0.02
Proficient (69-75)	0.00	0.01	0.02	0.03
Below Proficient (0-68)	0.00	0.02	0.92	0.95
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

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	0.03	0.01	0.05
Below Proficient (0-58)	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 (73-80)	0.00	0.00	0.00	0.01
Proficient (59-72)	0.00	0.03	0.01	0.05
Below Proficient (0-58)	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 2010

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.30	0.05	0.35
0-54	0.06	0.58	0.65
Estimated Proportion Correctly Classified			0.89

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.29	0.07	0.35
0-54	0.09	0.56	0.65
Estimated Proportion Consistently Classified			0.84

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
44-80	0.24	0.04	0.29
0-43	0.04	0.68	0.71
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
44-80	0.23	0.06	0.29
0-43	0.06	0.65	0.71
Estimated Proportion Consistently Classified			0.88

Table 6.N.9: ESEA Reliability Classifications—February 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.19	0.04	0.00	0.23
Proficient (68-75)	0.03	0.12	0.04	0.19
Below Proficient (0-67)	0.00	0.03	0.55	0.58
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 (76-90)	0.18	0.04	0.01	0.23
Proficient (68-75)	0.04	0.09	0.05	0.19
Below Proficient (0-67)	0.00	0.04	0.53	0.58
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.91
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 (72-80)	0.12	0.02	0.00	0.15
Proficient (58-71)	0.02	0.21	0.03	0.26
Below Proficient (0-57)	0.00	0.02	0.57	0.59
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 (72-80)	0.12	0.03	0.00	0.15
Proficient (58-71)	0.03	0.19	0.04	0.26
Below Proficient (0-57)	0.00	0.04	0.56	0.59
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 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.63	0.03	0.66
0-54	0.04	0.30	0.34
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.62	0.04	0.66
0-54	0.05	0.29	0.34
Estimated Proportion Consistently Classified			0.91

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.64	0.02	0.66
0-41	0.04	0.30	0.34
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.63	0.03	0.66
0-41	0.05	0.29	0.34
Estimated Proportion Consistently Classified			0.92

Table 6.N.11: ESEA Reliability Classifications—March 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (75-90)	0.26	0.04	0.00	0.30
Proficient (67-74)	0.04	0.15	0.04	0.23
Below Proficient (0-66)	0.00	0.04	0.43	0.47
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 (75-90)	0.25	0.05	0.00	0.30
Proficient (67-74)	0.06	0.12	0.06	0.23
Below Proficient (0-66)	0.00	0.05	0.42	0.47
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)	0.15	0.03	0.00	0.19
Proficient (58-71)	0.02	0.25	0.04	0.31
Below Proficient (0-57)	0.00	0.03	0.48	0.51
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.95
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)	0.15	0.04	0.00	0.18
Proficient (58-71)	0.04	0.23	0.05	0.31
Below Proficient (0-57)	0.00	0.04	0.46	0.51
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.93
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.91

Table 6.N.12: Pass/Not Pass Classifications—March 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.74	0.03	0.76
0-54	0.03	0.21	0.24
Estimated Proportion Correctly Classified			0.95

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.72	0.04	0.76
0-54	0.04	0.20	0.24
Estimated Proportion Consistently Classified			0.92

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.74	0.02	0.76
0-41	0.03	0.21	0.24
Estimated Proportion Correctly Classified			0.95

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.73	0.03	0.76
0-41	0.04	0.20	0.24
Estimated Proportion Consistently Classified			0.93

Table 6.N.13: ESEA Reliability Classifications—May 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.03	0.02	0.01	0.05
Proficient (69-76)	0.00	0.04	0.02	0.07
Below Proficient (0-68)	0.00	0.02	0.86	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 (77-90)	0.03	0.02	0.01	0.05
Proficient (69-76)	0.01	0.04	0.02	0.07
Below Proficient (0-68)	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 (72-80)	0.00	0.01	0.01	0.02
Proficient (58-71)	0.00	0.05	0.02	0.07
Below Proficient (0-57)	0.00	0.01	0.90	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 (72-80)	0.00	0.01	0.01	0.02
Proficient (58-71)	0.00	0.05	0.02	0.07
Below Proficient (0-57)	0.00	0.02	0.88	0.90
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Table 6.N.14: Pass/Not Pass Classifications—May 2011

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.30	0.03	0.33
0-54	0.05	0.62	0.67
Estimated Proportion Correctly Classified			0.93

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.30	0.04	0.33
0-54	0.06	0.60	0.67
Estimated Proportion Consistently Classified			0.89

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.32	0.02	0.34
0-41	0.07	0.59	0.66
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.30	0.04	0.34
0-41	0.09	0.57	0.66
Estimated Proportion Consistently Classified			0.87

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 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	498.7676	450	23
88	475.7848	450	23
87	462.3210	450	23
86	452.5597	450	23
85	444.7119	445	21
84	438.0362	438	17
83	432.1522	432	17
82	426.8505	427	17
81	421.9952	422	15
80	417.5023	418	14
79	413.3119	413	14
78	409.3771	409	13
77	405.6638	406(Advanced)	13
76	402.1496	402	12
75	398.8051	399	12
74	395.6102	396	12
73	392.5481	393	11
72	389.6053	390	11
71	386.7687	387	11
70	384.0254	384	11
69	381.3651	381(Proficient)	10
68	378.7817	379	10
67	376.2668	376	10
66	373.8138	374	10
65	371.4157	371	10
64	369.0710	369	9
63	366.7709	367	9
62	364.5139	365	9
61	362.2931	362	9
60	360.1092	360	9
59	357.9570	358	9
58	355.8324	356	9
57	353.7349	354	9
56	351.6610	352	9
55	349.6068	350(Pass)	9
54	347.5734	348	9
53	345.5558	346	9
52	343.5536	344	9
51	341.5633	342	9
50	339.5830	340	8
49	337.6126	338	8
48	335.6482	336	8
47	333.6905	334	8
46	331.7388	332	8
45	329.7893	330	8
44	327.8421	328	8
43	325.8962	326	8
42	323.9494	324	9
41	322.0013	322	9
40	320.0527	320	9
39	318.1013	318	9
38	316.1488	316	9
37	314.1925	314	9
36	312.2353	312	9
35	310.2740	310	9
34	308.3114	308	9
33	306.3456	306	9
32	304.3797	304	9
31	302.4133	302	9
30	300.4476	300	9
29	298.4797	298	9
28	296.5118	297	10
27	294.5415	295	10
26	292.5678	293	10
25	290.5906	291	10
24	288.6074	289	10
23	286.6093	287	10
22	284.5856	285	10
21	282.5362	283	11
20	280.4572	280	11
19	278.3405	278	11
18	276.1759	276	11
17	273.9401	275	11
16	271.6106	275	11
15	269.1853	275	11
14	266.6456	275	11
13	263.9688	275	11
12	261.0778	275	11
11	257.9568	275	11
10	254.5672	275	11
9	250.8421	275	11
8	246.6925	275	11
7	242.0010	275	11
6	236.5594	275	11
5	230.0821	275	11
4	222.1674	275	11
3	211.9877	275	11
2	197.7047	275	11
1	173.4483	275	11
0	147.2238	275	11

Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.3485	450	19
78	469.0993	450	19
77	455.3064	450	19
76	445.3609	445	17
75	437.5207	438	16
74	431.0079	431	14
73	425.4075	425(Advanced)	13
72	420.4730	420	12
71	416.0420	416	12
70	412.0124	412	11
69	408.3025	408	11
68	404.8551	405	10
67	401.6290	402	10
66	398.5882	399	10
65	395.7043	396	10
64	392.9575	393	9
63	390.3314	390	9
62	387.8100	388	9
61	385.3800	385	9
60	383.0314	383	9
59	380.7551	381(Proficient)	9
58	378.5434	379	8
57	376.3886	376	8
56	374.2863	374	8
55	372.2293	372	8
54	370.2144	370	8
53	368.2356	368	8
52	366.2873	366	8
51	364.3720	364	8
50	362.4834	362	8
49	360.6155	361	8
48	358.7690	359	8
47	356.9391	357	8
46	355.1251	355	8
45	353.3238	353	8
44	351.5314	352	8
43	349.7475	350(Pass)	8
42	347.9717	348	8
41	346.1967	346	8
40	344.4248	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.6529	343	8
38	340.8782	341	8
37	339.1002	339	8
36	337.3127	337	8
35	335.5186	336	8
34	333.7121	334	8
33	331.8954	332	8
32	330.0604	330	8
31	328.2087	328	8
30	326.3351	326	8
29	324.4361	324	8
28	322.5131	323	8
27	320.5595	321	8
26	318.5705	319	8
25	316.5452	317	8
24	314.4763	314	8
23	312.3612	312	8
22	310.1940	310	8
21	307.9693	308	9
20	305.6799	306	9
19	303.3148	303	9
18	300.8689	301	9
17	298.3292	298	9
16	295.6837	296	9
15	292.9190	293	10
14	290.0158	290	10
13	286.9523	287	10
12	283.7020	284	10
11	280.2306	280	11
10	276.4938	276	11
9	272.4371	275	12
8	267.9823	275	12
7	263.0175	275	12
6	257.3879	275	12
5	250.8476	275	12
4	242.9821	275	12
3	233.0164	275	12
2	219.2070	275	12
1	195.9979	275	12
0	171.7079	275	12

Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	500.8976	450	23
88	477.9664	450	23
87	464.2965	450	23
86	454.2687	450	23
85	446.1457	446	21
84	439.2125	439	17
83	433.0981	433	17
82	427.6007	428	17
81	422.5824	423	15
80	417.9572	418	15
79	413.6614	414	14
78	409.6435	410	13
77	405.8641	406(Advanced)	13
76	402.2989	402	12
75	398.9145	399	12
74	395.6886	396	12
73	392.6028	393	11
72	389.6422	390	11
71	386.7930	387	11
70	384.0418	384	10
69	381.3778	381(Proficient)	10
68	378.7944	379	10
67	376.2832	376	10
66	373.8375	374	10
65	371.4501	371	10
64	369.1191	369	9
63	366.8358	367	9
62	364.5982	365	9
61	362.3997	362	9
60	360.2404	360	9
59	358.1153	358	9
58	356.0197	356	9
57	353.9527	354	9
56	351.9104	352	9
55	349.8911	350(Pass)	9
54	347.8950	348	9
53	345.9157	346	8
52	343.9547	344	8
51	342.0063	342	8
50	340.0711	340	8
49	338.1456	338	8
48	336.2306	336	8
47	334.3220	334	8
46	332.4222	332	8
45	330.5277	331	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.6368	329	8
43	326.7503	327	8
42	324.8661	325	8
41	322.9845	323	8
40	321.1035	321	9
39	319.2279	319	9
38	317.3495	317	9
37	315.4761	315	9
36	313.6016	314	9
35	311.7304	312	9
34	309.8603	310	9
33	307.9925	308	9
32	306.1259	306	9
31	304.2617	304	9
30	302.3992	302	9
29	300.5383	301	9
28	298.6750	299	9
27	296.8090	297	10
26	294.9393	295	10
25	293.0603	293	10
24	291.1668	291	10
23	289.2545	289	10
22	287.3170	287	10
21	285.3494	285	10
20	283.3432	283	11
19	281.2913	281	11
18	279.1859	279	11
17	276.9870	277	11
16	274.7016	275	11
15	272.3122	275	11
14	269.7971	275	11
13	267.1303	275	11
12	264.2719	275	11
11	261.1566	275	11
10	257.7620	275	11
9	254.0220	275	11
8	249.8454	275	11
7	245.1088	275	11
6	239.6384	275	11
5	233.1657	275	11
4	225.2456	275	11
3	215.0390	275	11
2	200.6591	275	11
1	176.0655	275	11
0	147.2238	275	11

Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	491.2375	450	19
78	467.8254	450	19
77	453.9764	450	19
76	444.0078	444	17
75	436.1723	436	16
74	429.6682	430	14
73	424.0819	424(Advanced)	13
72	419.1659	419	12
71	414.7546	415	12
70	410.7479	411	11
69	407.0606	407	11
68	403.6387	404	10
67	400.4381	400	10
66	397.4227	397	10
65	394.5642	395	9
64	391.8438	392	9
63	389.2469	389	9
62	386.7516	387	9
61	384.3500	384	9
60	382.0283	382	9
59	379.7812	380(Proficient)	8
58	377.5961	378	8
57	375.4708	375	8
56	373.3963	373	8
55	371.3680	371	8
54	369.3822	369	8
53	367.4304	367	8
52	365.5120	366	8
51	363.6268	364	8
50	361.7657	362	8
49	359.9279	360	8
48	358.1102	358	8
47	356.3101	356	8
46	354.5258	355	8
45	352.7535	353	8
44	350.9917	351(Pass)	8
43	349.2384	349	8
42	347.4918	347	8
41	345.7477	346	7
40	344.0067	344	7

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.2656	342	7
38	340.5224	341	8
37	338.7755	339	8
36	337.0203	337	8
35	335.2588	335	8
34	333.4853	333	8
33	331.7022	332	8
32	329.9000	330	8
31	328.0820	328	8
30	326.2423	326	8
29	324.3775	324	8
28	322.4898	322	8
27	320.5711	321	8
26	318.6179	319	8
25	316.6287	317	8
24	314.5968	315	8
23	312.5189	313	8
22	310.3897	310	8
21	308.2036	308	8
20	305.9535	306	9
19	303.6290	304	9
18	301.2232	301	9
17	298.7256	299	9
16	296.1221	296	9
15	293.4004	293	9
14	290.5413	291	10
13	287.5224	288	10
12	284.3177	284	10
11	280.8925	281	11
10	277.2027	277	11
9	273.1925	275	12
8	268.7849	275	12
7	263.8673	275	12
6	258.2824	275	12
5	251.7842	275	12
4	243.9550	275	12
3	234.0127	275	12
2	220.1973	275	12
1	196.8915	275	12
0	171.7079	275	12

Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	501.4126	450	23
88	478.2685	450	23
87	464.4156	450	23
86	454.2585	450	23
85	446.0489	446	21
84	439.0609	439	17
83	432.9141	433	17
82	427.3987	427	17
81	422.3714	422	15
80	417.7419	418	15
79	413.4442	413	14
78	409.4249	409	13
77	405.6447	406(Advanced)	13
76	402.0777	402	12
75	398.6907	399	12
74	395.4615	395	12
73	392.3714	392	11
72	389.4058	389	11
71	386.5505	387	11
70	383.7922	384	11
69	381.1199	381(Proficient)	10
68	378.5283	379	10
67	376.0070	376	10
66	373.5506	374	10
65	371.1515	371	10
64	368.8088	369	9
63	366.5121	367	9
62	364.2615	364	9
61	362.0486	362	9
60	359.8758	360	9
59	357.7374	358	9
58	355.6281	356	9
57	353.5474	354	9
56	351.4904	351(Pass)	9
55	349.4567	349	9
54	347.4451	347	9
53	345.4516	345	9
52	343.4752	343	8
51	341.5129	342	8
50	339.5623	340	8
49	337.6238	338	8
48	335.6934	336	8
47	333.7717	334	8
46	331.8581	332	8
45	329.9491	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.0438	328	8
43	326.1430	326	8
42	324.2433	324	8
41	322.34500	322	8
40	320.4472	320	9
39	318.5504	319	9
38	316.6521	317	9
37	314.7535	315	9
36	312.8533	313	9
35	310.9517	311	9
34	309.0488	309	9
33	307.1427	307	9
32	305.2349	305	9
31	303.3234	303	9
30	301.4105	301	9
29	299.4912	299	9
28	297.5646	298	9
27	295.6311	296	10
26	293.6855	294	10
25	291.7267	292	10
24	289.7525	290	10
23	287.7564	288	10
22	285.7331	286	10
21	283.6770	284	10
20	281.5813	282	11
19	279.4302	279	11
18	277.2096	277	11
17	274.9168	275	11
16	272.5386	275	11
15	270.0584	275	11
14	267.4561	275	11
13	264.6719	275	11
12	261.6957	275	11
11	258.4972	275	11
10	255.0286	275	11
9	251.2257	275	11
8	247.0020	275	11
7	242.2430	275	11
6	236.7195	275	11
5	230.2037	275	11
4	222.2679	275	11
3	212.0849	275	11
2	197.8147	275	11
1	173.5745	275	11
0	147.2238	275	11

Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	491.8586	450	19
78	468.5295	450	19
77	454.7026	450	19
76	444.7336	445	17
75	436.8895	437	16
74	430.3708	430	14
73	424.7669	425(Advanced)	13
72	419.8309	420	12
71	415.3990	415	12
70	411.3699	411	11
69	407.6604	408	11
68	404.2148	404	10
67	400.9906	401	10
66	397.9516	398	10
65	395.0695	395	10
64	392.3242	392	9
63	389.7029	390	9
62	387.1838	387	9
61	384.7572	385	9
60	382.4112	382	9
59	380.1390	380(Proficient)	9
58	377.9297	378	8
57	375.7790	376	8
56	373.6797	374	8
55	371.6263	372	8
54	369.6152	370	8
53	367.6386	368	8
52	365.6946	366	8
51	363.7835	364	8
50	361.8969	362	8
49	360.0330	360	8
48	358.1892	358	8
47	356.3627	356	8
46	354.5516	355	8
45	352.7524	353	8
44	350.9632	351(Pass)	8
43	349.1823	349	8
42	347.4076	347	8
41	345.6351	346	8
40	343.8651	344	8
39	342.0945	342	8
38	340.3213	340	8
37	338.5434	339	8
36	336.7573	337	8
35	334.9634	335	8
34	333.1577	333	8
33	331.3405	331	8
32	329.5049	330	8
31	327.6523	328	8
30	325.7768	326	8
29	323.8759	324	8
28	321.9518	322	8
27	319.9941	320	8
26	318.001	318	8
25	315.9716	316	8
24	313.8959	314	8
23	311.7736	312	8
22	309.5990	310	8
21	307.3653	307	9
20	305.0658	305	9
19	302.6880	303	9
18	300.2304	300	9
17	297.6752	298	9
16	295.0143	295	9
15	292.2322	292	10
14	289.3094	289	10
13	286.2250	286	10
12	282.9513	283	11
11	279.4550	279	11
10	275.6908	276	11
9	271.6058	275	12
8	267.1202	275	12
7	262.1226	275	12
6	256.4607	275	12
5	249.8886	275	12
4	241.9948	275	12
3	232.0108	275	12
2	218.2100	275	12
1	195.0982	275	12
0	171.7079	275	12

Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	494.9215	450	23
88	471.8964	450	23
87	458.8763	450	23
86	449.6635	450	23
85	442.3753	442	21
84	436.2258	436	17
83	430.8319	431	17
82	425.9598	426	17
81	421.4752	421	15
80	417.2924	417	14
79	413.3532	413	14
78	409.6133	410	13
77	406.0505	406	13
76	402.6425	403(Advanced)	12
75	399.3692	399	12
74	396.2169	396	12
73	393.1744	393	11
72	390.2328	390	11
71	387.3847	387	11
70	384.6196	385	11
69	381.9314	382(Proficient)	10
68	379.3135	379	10
67	376.7635	377	10
66	374.2737	374	10
65	371.8390	372	10
64	369.4576	369	10
63	367.1238	367	9
62	364.8339	365	9
61	362.5833	363	9
60	360.3710	360	9
59	358.1934	358	9
58	356.0453	356	9
57	353.9252	354	9
56	351.8328	352	9
55	349.7621	350(Pass)	9
54	347.7144	348	9
53	345.6838	346	9
52	343.6705	344	9
51	341.6701	342	8
50	339.6811	340	8
49	337.7025	338	8
48	335.7309	336	8
47	333.7661	334	8
46	331.8076	332	8
45	329.8512	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	327.8966	328	8
43	325.9428	326	8
42	323.9868	324	8
41	322.0282	322	8
40	320.0670	320	9
39	318.1005	318	9
38	316.1299	316	9
37	314.1518	314	9
36	312.1687	312	9
35	310.1760	310	9
34	308.1765	308	9
33	306.1679	306	9
32	304.1519	304	9
31	302.1286	302	9
30	300.0975	300	9
29	298.0562	298	9
28	296.0063	296	10
27	293.9458	294	10
26	291.8740	292	10
25	289.7908	290	10
24	287.6916	288	10
23	285.5728	286	10
22	283.4312	283	10
21	281.2618	281	11
20	279.0628	279	11
19	276.8077	277	11
18	274.4945	275	11
17	272.1181	275	11
16	269.6655	275	11
15	267.1224	275	11
14	264.4512	275	11
13	261.6093	275	11
12	258.5963	275	11
11	255.3779	275	11
10	251.9096	275	11
9	248.0759	275	11
8	243.7809	275	11
7	238.9423	275	11
6	233.3867	275	11
5	226.8504	275	11
4	218.8951	275	11
3	208.6999	275	11
2	194.4253	275	11
1	171.0681	275	11
0	147.2238	275	11

Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.3597	450	19
78	469.0720	450	19
77	455.2377	450	19
76	445.2526	445	17
75	437.3754	437	16
74	430.8285	431	15
73	425.1967	425(Advanced)	13
72	420.2338	420	12
71	415.7768	416	12
70	411.7239	412	11
69	407.9927	408	11
68	404.5266	405	10
67	401.2838	401	10
66	398.2280	398	10
65	395.3307	395	10
64	392.5729	393	9
63	389.9374	390	9
62	387.4074	387	9
61	384.9710	385	9
60	382.6166	383	9
59	380.3369	380(Proficient)	9
58	378.1217	378	8
57	375.9658	376	8
56	373.8626	374	8
55	371.8063	372	8
54	369.7931	370	8
53	367.8159	368	8
52	365.8714	366	8
51	363.9611	364	8
50	362.0767	362	8
49	360.2152	360	8
48	358.3750	358	8
47	356.5528	357	8
46	354.7470	355	8
45	352.9543	353	8
44	351.1717	351(Pass)	8
43	349.3982	349	8
42	347.6324	348	8
41	345.8691	346	8
40	344.1093	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.3500	342	8
38	340.5886	341	8
37	338.8240	339	8
36	337.0513	337	8
35	335.2725	335	8
34	333.4819	333	8
33	331.6820	332	8
32	329.8632	330	8
31	328.0287	328	8
30	326.1724	326	8
29	324.2913	324	8
28	322.3874	322	8
27	320.4524	320	8
26	318.4827	318	8
25	316.4772	316	8
24	314.4286	314	8
23	312.3342	312	8
22	310.1882	310	8
21	307.9852	308	8
20	305.7178	306	9
19	303.3752	303	9
18	300.9520	301	9
17	298.4357	298	9
16	295.8137	296	9
15	293.0730	293	10
14	290.1941	290	10
13	287.1552	287	10
12	283.9298	284	10
11	280.4834	280	11
10	276.7718	277	11
9	272.7398	275	12
8	268.3099	275	12
7	263.3695	275	12
6	257.7630	275	12
5	251.2441	275	12
4	243.3970	275	12
3	233.4435	275	12
2	219.6330	275	12
1	196.3829	275	12
0	171.7079	275	12

Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	21
89	504.9645	450	21
88	483.6953	450	21
87	469.0089	450	21
86	458.3266	450	21
85	449.7041	450	21
84	442.3611	442	17
83	435.9049	436	17
82	430.1258	430	17
81	424.8731	425	15
80	420.0511	420	15
79	415.5899	416	14
78	411.4320	411	14
77	407.5292	408	13
76	403.8570	404(Advanced)	13
75	400.3787	400	12
74	397.0687	397	12
73	393.9065	394	11
72	390.8756	391	11
71	387.9631	388	11
70	385.1532	385	11
69	382.4391	382	10
68	379.8059	380(Proficient)	10
67	377.2526	377	10
66	374.7635	375	10
65	372.3370	372	10
64	369.9668	370	9
63	367.6497	368	9
62	365.3775	365	9
61	363.1485	363	9
60	360.9581	361	9
59	358.8049	359	9
58	356.6831	357	9
57	354.5910	355	9
56	352.5248	353	9
55	350.4834	350(Pass)	9
54	348.4660	348	9
53	346.4677	346	9
52	344.4895	344	9
51	342.5253	343	8
50	340.5777	341	8
49	338.6398	339	8
48	336.7165	337	8
47	334.7993	335	8
46	332.8927	333	8
45	330.9946	331	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	329.1010	329	8
43	327.2128	327	8
42	325.3290	325	8
41	323.4484	323	9
40	321.5701	322	9
39	319.6964	320	9
38	317.8223	318	9
37	315.9524	316	9
36	314.0813	314	9
35	312.2128	312	9
34	310.3431	310	9
33	308.4741	308	9
32	306.6025	307	9
31	304.7304	305	9
30	302.8548	303	9
29	300.9776	301	9
28	299.0923	299	9
27	297.1981	297	10
26	295.2932	295	10
25	293.3697	293	10
24	291.4277	291	10
23	289.4630	289	10
22	287.4671	287	10
21	285.4337	285	10
20	283.3560	283	11
19	281.2257	281	11
18	279.0353	279	11
17	276.7562	277	11
16	274.3786	275	12
15	271.8919	275	12
14	269.2744	275	12
13	266.5013	275	12
12	263.5430	275	12
11	260.3598	275	12
10	256.8819	275	12
9	253.0521	275	12
8	248.7918	275	12
7	243.9828	275	12
6	238.4544	275	12
5	231.9416	275	12
4	224.0045	275	12
3	213.8098	275	12
2	199.4890	275	12
1	175.0617	275	12
0	147.2238	275	12

Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	493.7337	450	19
78	470.7041	450	19
77	457.0009	450	19
76	447.0981	447	17
75	439.2802	439	16
74	432.7795	433	14
73	427.1850	427	13
72	422.2579	422(Advanced)	12
71	417.8352	418	12
70	413.8031	414	11
69	410.0939	410	11
68	406.6441	407	10
67	403.4150	403	10
66	400.3720	400	10
65	397.4864	397	10
64	394.7355	395	9
63	392.1045	392	9
62	389.5816	390	9
61	387.1486	387	9
60	384.7978	385	9
59	382.5187	383	9
58	380.3054	380(Proficient)	8
57	378.1490	378	8
56	376.0448	376	8
55	373.9875	374	8
54	371.9714	372	8
53	369.9934	370	8
52	368.0476	368	8
51	366.1300	366	8
50	364.2421	364	8
49	362.3778	362	8
48	360.5324	361	8
47	358.7060	359	8
46	356.8942	357	8
45	355.0963	355	8
44	353.3092	353	8
43	351.5295	352	8
42	349.7564	350(Pass)	8
41	347.9898	348	8
40	346.2225	346	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	344.4567	344	8
38	342.6894	343	8
37	340.9175	341	8
36	339.1407	339	8
35	337.3528	337	8
34	335.5567	336	8
33	333.7463	334	8
32	331.9238	332	8
31	330.0811	330	8
30	328.2195	328	8
29	326.3339	326	8
28	324.4197	324	8
27	322.4794	322	8
26	320.5053	321	8
25	318.4923	318	8
24	316.4394	316	8
23	314.3386	314	8
22	312.1864	312	8
21	309.9776	310	9
20	307.7047	308	9
19	305.3600	305	9
18	302.9305	303	9
17	300.4119	300	9
16	297.7863	298	9
15	295.0426	295	10
14	292.1630	292	10
13	289.1240	289	10
12	285.9016	286	10
11	282.4590	282	11
10	278.7445	279	11
9	274.7060	275	12
8	270.2698	275	12
7	265.3217	275	12
6	259.7010	275	12
5	253.1622	275	12
4	245.2842	275	12
3	235.2779	275	12
2	221.3666	275	12
1	197.8728	275	12
0	171.7079	275	12

Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	17
89	511.1055	450	17
88	496.2080	450	17
87	480.1327	450	17
86	469.4781	450	17
85	460.1742	450	17
84	452.0938	450	17
83	444.9397	445	17
82	438.5129	439	17
81	432.6858	433	16
80	427.3529	427	16
79	422.4330	422	15
78	417.8625	418	14
77	413.5885	414	14
76	409.5623	410	13
75	405.7572	406(Advanced)	13
74	402.1422	402	12
73	398.6905	399	12
72	395.3827	395	12
71	392.2029	392	11
70	389.1391	389	11
69	386.1785	386	11
68	383.3103	383	10
67	380.5240	381(Proficient)	10
66	377.8167	378	10
65	375.1801	375	10
64	372.6119	373	10
63	370.1016	370	10
62	367.6467	368	9
61	365.2383	365	9
60	362.8748	363	9
59	360.5525	361	9
58	358.2687	358	9
57	356.0180	356	9
56	353.8003	354	9
55	351.6113	352(Pass)	9
54	349.4488	349	9
53	347.3114	347	9
52	345.1953	345	9
51	343.0982	343	9
50	341.0189	341	9
49	338.9524	339	9
48	336.9037	337	9
47	334.8641	335	9
46	332.8376	333	9
45	330.8215	331	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.8119	329	9
43	326.8099	327	9
42	324.8133	325	9
41	322.8220	323	9
40	320.8337	321	9
39	318.8513	319	9
38	316.8687	317	9
37	314.8893	315	9
36	312.9097	313	9
35	310.9303	311	9
34	308.9504	309	9
33	306.9670	307	9
32	304.9814	305	9
31	302.9889	303	9
30	300.9915	301	9
29	298.9827	299	10
28	296.9617	297	10
27	294.9277	295	10
26	292.8740	293	10
25	290.7999	291	10
24	288.7013	289	10
23	286.5682	287	10
22	284.3907	284	10
21	282.1644	282	11
20	279.8850	280	11
19	277.5401	278	11
18	275.1199	275	11
17	272.6133	275	11
16	270.0057	275	11
15	267.2802	275	11
14	264.4014	275	11
13	261.3346	275	11
12	258.0703	275	11
11	254.5724	275	11
10	250.7936	275	11
9	246.6754	275	11
8	242.1461	275	11
7	237.0021	275	11
6	231.1062	275	11
5	224.2122	275	11
4	215.8634	275	11
3	205.1894	275	11
2	190.1096	275	11
1	168.7055	275	11
0	147.2238	275	11

Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	494.7435	450	19
78	472.4703	450	19
77	458.6133	450	19
76	448.6439	449	17
75	440.7814	441	16
74	434.2469	434	14
73	428.6217	429	13
72	423.6599	424(Advanced)	12
71	419.2041	419	12
70	415.1419	415	11
69	411.4019	411	11
68	407.9233	408	10
67	404.6639	405	10
66	401.5911	402	10
65	398.6763	399	10
64	395.8966	396	9
63	393.2354	393	9
62	390.6799	391	9
61	388.2170	388	9
60	385.8338	386	9
59	383.5234	384	9
58	381.2764	381(Proficient)	8
57	379.0882	379	8
56	376.9492	377	8
55	374.8583	375	8
54	372.8073	373	8
53	370.7934	371	8
52	368.8124	369	8
51	366.8584	367	8
50	364.9325	365	8
49	363.0320	363	8
48	361.1478	361	8
47	359.2833	359	8
46	357.4321	357	8
45	355.5943	356	8
44	353.7669	354	8
43	351.9467	352	8
42	350.1310	350(Pass)	8
41	348.3222	348	8
40	346.5118	347	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	344.7022	345	8
38	342.8904	343	8
37	341.0736	341	8
36	339.2511	339	8
35	337.4173	337	8
34	335.5743	336	8
33	333.7166	334	8
32	331.8463	332	8
31	329.9549	330	8
30	328.0445	328	8
29	326.1090	326	8
28	324.1439	324	8
27	322.1533	322	8
26	320.1271	320	8
25	318.0615	318	8
24	315.9561	316	8
23	313.8002	314	8
22	311.5931	312	8
21	309.3283	309	9
20	306.9984	307	9
19	304.5951	305	9
18	302.1058	302	9
17	299.5268	300	9
16	296.8387	297	9
15	294.0316	294	10
14	291.0860	291	10
13	287.9790	288	10
12	284.6850	285	11
11	281.1689	281	11
10	277.3868	277	11
9	273.2825	275	12
8	268.7798	275	12
7	263.7658	275	12
6	258.0841	275	12
5	251.4893	275	12
4	243.5658	275	12
3	233.5362	275	12
2	219.6520	275	12
1	196.3394	275	12
0	171.7079	275	12

Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	501.5543	450	23
88	478.5615	450	23
87	464.7336	450	23
86	454.5172	450	23
85	446.1999	446	21
84	439.0807	439	17
83	432.7974	433	17
82	427.1544	427	17
81	422.0161	422	15
80	417.2966	417	15
79	412.9302	413	14
78	408.8626	409	13
77	405.0534	405(Advanced)	13
76	401.4732	401	12
75	398.0864	398	12
74	394.8681	395	12
73	391.7982	392	11
72	388.8604	389	11
71	386.0378	386	11
70	383.3169	383	10
69	380.6849	381(Proficient)	10
68	378.1376	378	10
67	375.6610	376	10
66	373.2512	373	10
65	370.8995	371	10
64	368.6053	369	9
63	366.3563	366	9
62	364.1542	364	9
61	361.9892	362	9
60	359.8644	360	9
59	357.7726	358	9
58	355.7097	356	9
57	353.6737	354	9
56	351.6635	352	9
55	349.6742	350(Pass)	9
54	347.7072	348	8
53	345.7569	346	8
52	343.8239	344	8
51	341.9034	342	8
50	339.9953	340	8
49	338.0966	338	8
48	336.2075	336	8
47	334.3242	334	8
46	332.4487	332	8
45	330.5773	331	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.7082	329	8
43	326.8416	327	8
42	324.9755	325	8
41	323.1094	323	8
40	321.2412	321	8
39	319.3745	319	8
38	317.5017	318	9
37	315.6295	316	9
36	313.7507	314	9
35	311.8697	312	9
34	309.9833	310	9
33	308.0928	308	9
32	306.1961	306	9
31	304.2945	304	9
30	302.3870	302	9
29	300.4736	300	9
28	298.5502	299	9
27	296.6172	297	10
26	294.6732	295	10
25	292.7140	293	10
24	290.7390	291	10
23	288.7452	289	10
22	286.7180	287	10
21	284.6546	285	10
20	282.5496	283	11
19	280.3977	280	11
18	278.1888	278	11
17	275.9108	276	11
16	273.5527	275	12
15	271.0757	275	12
14	268.4674	275	12
13	265.7114	275	12
12	262.7758	275	12
11	259.6240	275	12
10	256.2108	275	12
9	252.4587	275	12
8	248.2491	275	12
7	243.4945	275	12
6	238.0220	275	12
5	231.5642	275	12
4	223.6788	275	12
3	213.5299	275	12
2	199.2499	275	12
1	174.8677	275	12
0	147.2238	275	12

Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.8817	450	19
78	469.7640	450	19
77	456.0577	450	19
76	446.1807	446	17
75	438.4002	438	16
74	431.9423	432	14
73	426.3931	426	13
72	421.5067	422(Advanced)	12
71	417.1257	417	12
70	413.1371	413	11
69	409.4713	409	11
68	406.0652	406	10
67	402.8798	403	10
66	399.8797	400	10
65	397.0366	397	9
64	394.3312	394	9
63	391.7449	392	9
62	389.2650	389	9
61	386.8733	387	9
60	384.5630	385	9
59	382.3230	382	8
58	380.1482	380(Proficient)	8
57	378.0286	378	8
56	375.9608	376	8
55	373.9388	374	8
54	371.9571	372	8
53	370.0126	370	8
52	368.0997	368	8
51	366.2140	366	8
50	364.3572	364	8
49	362.5238	363	8
48	360.7082	361	8
47	358.9112	359	8
46	357.1280	357	8
45	355.3582	355	8
44	353.5989	354	8
43	351.8469	352	8
42	350.0999	350(Pass)	8
41	348.3597	348	7
40	346.6183	347	7

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	344.8773	345	7
38	343.1345	343	8
37	341.3868	341	8
36	339.6334	340	8
35	337.8698	338	8
34	336.0960	336	8
33	334.3089	334	8
32	332.5079	333	8
31	330.6886	331	8
30	328.8484	329	8
29	326.9853	327	8
28	325.0934	325	8
27	323.1733	323	8
26	321.2215	321	8
25	319.2307	319	8
24	317.1984	317	8
23	315.1201	315	8
22	312.9893	313	8
21	310.8014	311	8
20	308.5500	309	9
19	306.2268	306	9
18	303.8207	304	9
17	301.3223	301	9
16	298.7203	299	9
15	295.9980	296	9
14	293.1403	293	10
13	290.1237	290	10
12	286.9226	287	10
11	283.5032	284	11
10	279.8231	280	11
9	275.8254	276	12
8	271.4400	275	12
7	266.5577	275	12
6	260.9661	275	12
5	254.4540	275	12
4	246.5965	275	12
3	236.5961	275	12
2	222.6554	275	12
1	199.0201	275	12
0	171.7079	275	12

Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA— July, October and November 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	498.7676	450	23
88	475.7848	450	23
87	462.3210	450	23
86	452.5597	450	23
85	444.7119	445	21
84	438.0362	438	17
83	432.1522	432	17
82	426.8505	427	17
81	421.9952	422	15
80	417.5023	418	14
79	413.3119	413	14
78	409.3771	409	13
77	405.6638	406(Advanced)	13
76	402.1496	402	12
75	398.8051	399	12
74	395.6102	396	12
73	392.5481	393	11
72	389.6053	390	11
71	386.7687	387	11
70	384.0254	384	11
69	381.3651	381(Proficient)	10
68	378.7817	379	10
67	376.2668	376	10
66	373.8138	374	10
65	371.4157	371	10
64	369.0710	369	9
63	366.7709	367	9
62	364.5139	365	9
61	362.2931	362	9
60	360.1092	360	9
59	357.9570	358	9
58	355.8324	356	9
57	353.7349	354	9
56	351.6610	352	9
55	349.6068	350(Pass)	9
54	347.5734	348	9
53	345.5558	346	9
52	343.5536	344	9
51	341.5633	342	9
50	339.5830	340	8
49	337.6126	338	8
48	335.6482	336	8
47	333.6905	334	8
46	331.7388	332	8
45	329.7893	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	327.8421	328	8
43	325.8962	326	8
42	323.9494	324	9
41	322.0013	322	9
40	320.0527	320	9
39	318.1013	318	9
38	316.1488	316	9
37	314.1925	314	9
36	312.2353	312	9
35	310.2740	310	9
34	308.3114	308	9
33	306.3456	306	9
32	304.3797	304	9
31	302.4133	302	9
30	300.4476	300	9
29	298.4797	298	9
28	296.5118	297	10
27	294.5415	295	10
26	292.5678	293	10
25	290.5906	291	10
24	288.6074	289	10
23	286.6093	287	10
22	284.5856	285	10
21	282.5362	283	11
20	280.4572	280	11
19	278.3405	278	11
18	276.1759	276	11
17	273.9401	275	11
16	271.6106	275	11
15	269.1853	275	11
14	266.6456	275	11
13	263.9688	275	11
12	261.0778	275	11
11	257.9568	275	11
10	254.5672	275	11
9	250.8421	275	11
8	246.6925	275	11
7	242.0010	275	11
6	236.5594	275	11
5	230.0821	275	11
4	222.1674	275	11
3	211.9877	275	11
2	197.7047	275	11
1	173.4483	275	11
0	147.2238	275	11

Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July, October and November 2010

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.3485	450	19
78	469.0993	450	19
77	455.3064	450	19
76	445.3609	445	17
75	437.5207	438	16
74	431.0079	431	14
73	425.4075	425(Advanced)	13
72	420.4730	420	12
71	416.0420	416	12
70	412.0124	412	11
69	408.3025	408	11
68	404.8551	405	10
67	401.6290	402	10
66	398.5882	399	10
65	395.7043	396	10
64	392.9575	393	9
63	390.3314	390	9
62	387.8100	388	9
61	385.3800	385	9
60	383.0314	383	9
59	380.7551	381(Proficient)	9
58	378.5434	379	8
57	376.3886	376	8
56	374.2863	374	8
55	372.2293	372	8
54	370.2144	370	8
53	368.2356	368	8
52	366.2873	366	8
51	364.3720	364	8
50	362.4834	362	8
49	360.6155	361	8
48	358.7690	359	8
47	356.9391	357	8
46	355.1251	355	8
45	353.3238	353	8
44	351.5314	352	8
43	349.7475	350(Pass)	8
42	347.9717	348	8
41	346.1967	346	8
40	344.4248	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.6529	343	8
38	340.8782	341	8
37	339.1002	339	8
36	337.3127	337	8
35	335.5186	336	8
34	333.7121	334	8
33	331.8954	332	8
32	330.0604	330	8
31	328.2087	328	8
30	326.3351	326	8
29	324.4361	324	8
28	322.5131	323	8
27	320.5595	321	8
26	318.5705	319	8
25	316.5452	317	8
24	314.4763	314	8
23	312.3612	312	8
22	310.1940	310	8
21	307.9693	308	9
20	305.6799	306	9
19	303.3148	303	9
18	300.8689	301	9
17	298.3292	298	9
16	295.6837	296	9
15	292.9190	293	10
14	290.0158	290	10
13	286.9523	287	10
12	283.7020	284	10
11	280.2306	280	11
10	276.4938	276	11
9	272.4371	275	12
8	267.9823	275	12
7	263.0175	275	12
6	257.3879	275	12
5	250.8476	275	12
4	242.9821	275	12
3	233.0164	275	12
2	219.2070	275	12
1	195.9979	275	12
0	171.7079	275	12

Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—December 2010 and February 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	23
89	494.9215	450	23
88	471.8964	450	23
87	458.8763	450	23
86	449.6635	450	23
85	442.3753	442	21
84	436.2258	436	17
83	430.8319	431	17
82	425.9598	426	17
81	421.4752	421	15
80	417.2924	417	14
79	413.3532	413	14
78	409.6133	410	13
77	406.0505	406	13
76	402.6425	403(Advanced)	12
75	399.3692	399	12
74	396.2169	396	12
73	393.1744	393	11
72	390.2328	390	11
71	387.3847	387	11
70	384.6196	385	11
69	381.9314	382(Proficient)	10
68	379.3135	379	10
67	376.7635	377	10
66	374.2737	374	10
65	371.8390	372	10
64	369.4576	369	10
63	367.1238	367	9
62	364.8339	365	9
61	362.5833	363	9
60	360.3710	360	9
59	358.1934	358	9
58	356.0453	356	9
57	353.9252	354	9
56	351.8328	352	9
55	349.7621	350(Pass)	9
54	347.7144	348	9
53	345.6838	346	9
52	343.6705	344	9
51	341.6701	342	8
50	339.6811	340	8
49	337.7025	338	8
48	335.7309	336	8
47	333.7661	334	8
46	331.8076	332	8
45	329.8512	330	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	327.8966	328	8
43	325.9428	326	8
42	323.9868	324	8
41	322.0282	322	8
40	320.0670	320	9
39	318.1005	318	9
38	316.1299	316	9
37	314.1518	314	9
36	312.1687	312	9
35	310.1760	310	9
34	308.1765	308	9
33	306.1679	306	9
32	304.1519	304	9
31	302.1286	302	9
30	300.0975	300	9
29	298.0562	298	9
28	296.0063	296	10
27	293.9458	294	10
26	291.8740	292	10
25	289.7908	290	10
24	287.6916	288	10
23	285.5728	286	10
22	283.4312	283	10
21	281.2618	281	11
20	279.0628	279	11
19	276.8077	277	11
18	274.4945	275	11
17	272.1181	275	11
16	269.6655	275	11
15	267.1224	275	11
14	264.4512	275	11
13	261.6093	275	11
12	258.5963	275	11
11	255.3779	275	11
10	251.9096	275	11
9	248.0759	275	11
8	243.7809	275	11
7	238.9423	275	11
6	233.3867	275	11
5	226.8504	275	11
4	218.8951	275	11
3	208.6999	275	11
2	194.4253	275	11
1	171.0681	275	11
0	147.2238	275	11

Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— December 2010 and February 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	492.3597	450	19
78	469.0720	450	19
77	455.2377	450	19
76	445.2526	445	17
75	437.3754	437	16
74	430.8285	431	15
73	425.1967	425(Advanced)	13
72	420.2338	420	12
71	415.7768	416	12
70	411.7239	412	11
69	407.9927	408	11
68	404.5266	405	10
67	401.2838	401	10
66	398.2280	398	10
65	395.3307	395	10
64	392.5729	393	9
63	389.9374	390	9
62	387.4074	387	9
61	384.9710	385	9
60	382.6166	383	9
59	380.3369	380(Proficient)	9
58	378.1217	378	8
57	375.9658	376	8
56	373.8626	374	8
55	371.8063	372	8
54	369.7931	370	8
53	367.8159	368	8
52	365.8714	366	8
51	363.9611	364	8
50	362.0767	362	8
49	360.2152	360	8
48	358.3750	358	8
47	356.5528	357	8
46	354.7470	355	8
45	352.9543	353	8
44	351.1717	351(Pass)	8
43	349.3982	349	8
42	347.6324	348	8
41	345.8691	346	8
40	344.1093	344	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.3500	342	8
38	340.5886	341	8
37	338.8240	339	8
36	337.0513	337	8
35	335.2725	335	8
34	333.4819	333	8
33	331.6820	332	8
32	329.8632	330	8
31	328.0287	328	8
30	326.1724	326	8
29	324.2913	324	8
28	322.3874	322	8
27	320.4524	320	8
26	318.4827	318	8
25	316.4772	316	8
24	314.4286	314	8
23	312.3342	312	8
22	310.1882	310	8
21	307.9852	308	8
20	305.7178	306	9
19	303.3752	303	9
18	300.9520	301	9
17	298.4357	298	9
16	295.8137	296	9
15	293.0730	293	10
14	290.1941	290	10
13	287.1552	287	10
12	283.9298	284	10
11	280.4834	280	11
10	276.7718	277	11
9	272.7398	275	12
8	268.3099	275	12
7	263.3695	275	12
6	257.7630	275	12
5	251.2441	275	12
4	243.3970	275	12
3	233.4435	275	12
2	219.6330	275	12
1	196.3829	275	12
0	171.7079	275	12

Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—March and May 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	17
89	511.1055	450	17
88	496.2080	450	17
87	480.1327	450	17
86	469.4781	450	17
85	460.1742	450	17
84	452.0938	450	17
83	444.9397	445	17
82	438.5129	439	17
81	432.6858	433	16
80	427.3529	427	16
79	422.4330	422	15
78	417.8625	418	14
77	413.5885	414	14
76	409.5623	410	13
75	405.7572	406(Advanced)	13
74	402.1422	402	12
73	398.6905	399	12
72	395.3827	395	12
71	392.2029	392	11
70	389.1391	389	11
69	386.1785	386	11
68	383.3103	383	10
67	380.5240	381(Proficient)	10
66	377.8167	378	10
65	375.1801	375	10
64	372.6119	373	10
63	370.1016	370	10
62	367.6467	368	9
61	365.2383	365	9
60	362.8748	363	9
59	360.5525	361	9
58	358.2687	358	9
57	356.0180	356	9
56	353.8003	354	9
55	351.6113	352(Pass)	9
54	349.4488	349	9
53	347.3114	347	9
52	345.1953	345	9
51	343.0982	343	9
50	341.0189	341	9
49	338.9524	339	9
48	336.9037	337	9
47	334.8641	335	9
46	332.8376	333	9
45	330.8215	331	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.8119	329	9
43	326.8099	327	9
42	324.8133	325	9
41	322.8220	323	9
40	320.8337	321	9
39	318.8513	319	9
38	316.8687	317	9
37	314.8893	315	9
36	312.9097	313	9
35	310.9303	311	9
34	308.9504	309	9
33	306.9670	307	9
32	304.9814	305	9
31	302.9889	303	9
30	300.9915	301	9
29	298.9827	299	10
28	296.9617	297	10
27	294.9277	295	10
26	292.8740	293	10
25	290.7999	291	10
24	288.7013	289	10
23	286.5682	287	10
22	284.3907	284	10
21	282.1644	282	11
20	279.8850	280	11
19	277.5401	278	11
18	275.1199	275	11
17	272.6133	275	11
16	270.0057	275	11
15	267.2802	275	11
14	264.4014	275	11
13	261.3346	275	11
12	258.0703	275	11
11	254.5724	275	11
10	250.7936	275	11
9	246.6754	275	11
8	242.1461	275	11
7	237.0021	275	11
6	231.1062	275	11
5	224.2122	275	11
4	215.8634	275	11
3	205.1894	275	11
2	190.1096	275	11
1	168.7055	275	11
0	147.2238	275	11

Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—March and May 2011

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	19
79	494.7435	450	19
78	472.4703	450	19
77	458.6133	450	19
76	448.6439	449	17
75	440.7814	441	16
74	434.2469	434	14
73	428.6217	429	13
72	423.6599	424(Advanced)	12
71	419.2041	419	12
70	415.1419	415	11
69	411.4019	411	11
68	407.9233	408	10
67	404.6639	405	10
66	401.5911	402	10
65	398.6763	399	10
64	395.8966	396	9
63	393.2354	393	9
62	390.6799	391	9
61	388.2170	388	9
60	385.8338	386	9
59	383.5234	384	9
58	381.2764	381(Proficient)	8
57	379.0882	379	8
56	376.9492	377	8
55	374.8583	375	8
54	372.8073	373	8
53	370.7934	371	8
52	368.8124	369	8
51	366.8584	367	8
50	364.9325	365	8
49	363.0320	363	8
48	361.1478	361	8
47	359.2833	359	8
46	357.4321	357	8
45	355.5943	356	8
44	353.7669	354	8
43	351.9467	352	8
42	350.1310	350(Pass)	8
41	348.3222	348	8
40	346.5118	347	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	344.7022	345	8
38	342.8904	343	8
37	341.0736	341	8
36	339.2511	339	8
35	337.4173	337	8
34	335.5743	336	8
33	333.7166	334	8
32	331.8463	332	8
31	329.9549	330	8
30	328.0445	328	8
29	326.1090	326	8
28	324.1439	324	8
27	322.1533	322	8
26	320.1271	320	8
25	318.0615	318	8
24	315.9561	316	8
23	313.8002	314	8
22	311.5931	312	8
21	309.3283	309	9
20	306.9984	307	9
19	304.5951	305	9
18	302.1058	302	9
17	299.5268	300	9
16	296.8387	297	9
15	294.0316	294	10
14	291.0860	291	10
13	287.9790	288	10
12	284.6850	285	11
11	281.1689	281	11
10	277.3868	277	11
9	273.2825	275	12
8	268.7798	275	12
7	263.7658	275	12
6	258.0841	275	12
5	251.4893	275	12
4	243.5658	275	12
3	233.5362	275	12
2	219.6520	275	12
1	196.3394	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) 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. Assembly Bill (AB) 1609 passed in October 2001, which prohibited ninth-graders from taking the CAHSEE in order to have a census testing of all tenth-graders. 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 primarily was composed of California educators (e.g., English or mathematics teachers), while the other panel primarily was 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 2010–11 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 2010 administration, 27 percent of examinees who took the ELA examination received a raw score of 55 or higher and thereby passed the ELA. Of those who passed, 5 percent of examinees received a score of 69 or higher and passed the proficient cut point, while 1 percent scored 77 or above and passed the advanced cut point.

Among examinees who took the mathematics exam, 26 percent passed by attaining a raw score of 43 or higher. Of those who passed, 4 percent were at or above the proficient cut point with a raw score of 59 or higher, while 1 percent was at or above the advanced cut point with a raw score of 73 and 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 2010	Pass	55	27	43	26
	Proficient And Above	69	5	59	4
	Advanced And Above	77	1	73	1
October 2010	Pass	55	37	44	31
	Proficient And Above	69	8	59	7
	Advanced And Above	77	3	73	2
November 2010	Pass	56	40	44	35
	Proficient And Above	69	10	59	9
	Advanced And Above	77	4	73	2
December 2010	Pass	55	36	44	29
	Proficient And Above	69	6	59	5
	Advanced And Above	76	2	73	1
February 2011	Pass	55	67	42	68
	Proficient And Above	68	42	58	41
	Advanced And Above	76	23	72	15
March 2011	Pass	55	77	42	77
	Proficient And Above	67	53	58	50
	Advanced And Above	75	30	72	19
May 2011	Pass	55	34	42	36
	Proficient And Above	69	12	58	10
	Advanced And Above	77	5	72	2

[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 provides 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, percent 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 MC 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] it is considered discrepant, and a scoring leader provides a third score, which becomes the score of record. Otherwise, the writing score is obtained by averaging two scores.

[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.

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.

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.

[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.

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 block.
- 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 percentages of students passing each CAHSEE content area are 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 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.

[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.

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. Students who use accommodations that are specified in their IEP or Section 504 plan 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 the 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 was 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 2010

Table 8.A.1: Frequency Distributions, ELA—July 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9	0	9	99
440-449	4	0	13	99
430-439	15	0	28	99
420-429	24	0	52	99
410-419	38	0	90	99
400-409	59	1	149	98
390-399	122	1	271	97
380-389	132	2	403	95
370-379	250	3	653	92
360-369	582	7	1,235	85
350-359[1]	1,055	13	2,290	73
340-349	1,486	18	3,776	55
330-339	1,519	18	5,295	37
320-329	1,205	14	6,500	23
310-319	849	10	7,349	12
300-309	491	6	7,840	7
290-299	306	4	8,146	3
280-289	140	2	8,286	1
270-279	103	1	8,389	0

[1]Passing Score = 350

Table 8.A.2: Frequency Distributions, Mathematics—July 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	16	0	16	99
440-449	9	0	25	99
430-439	17	0	42	99
420-429	15	0	57	99
410-419	19	0	76	99
400-409	48	1	124	98
390-399	88	1	212	97
380-389	119	1	331	96
370-379	219	3	550	93
360-369	435	5	985	88
350-359[1]	1,145	14	2,130	74
340-349	1,405	17	3,535	57
330-339	1,969	24	5,504	33
320-329	1,305	16	6,809	17
310-319	923	11	7,732	6
300-309	340	4	8,072	2
290-299	109	1	8,181	1
280-289	26	0	8,207	0
270-279	15	0	8,222	0

[1]Passing Score = 350

Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9	0	9	99
440-449	4	0	13	99
430-439	15	0	28	99
420-429	24	0	52	99
410-419	38	0	90	99
403-409[1]	34	0	124	99
390-402	147	2	271	97
380-389[2]	132	2	403	95
370-379	250	3	653	92
360-369	582	7	1,235	85
350-359	1,055	13	2,290	73
340-349	1,486	18	3,776	55
330-339	1,519	18	5,295	37
320-329	1,205	14	6,500	23
310-319	849	10	7,349	12
300-309	491	6	7,840	7
290-299	306	4	8,146	3
280-289	140	2	8,286	1
270-279	103	1	8,389	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	16	0	16	99
440-449	9	0	25	99
430-439	17	0	42	99
422-429[1]	5	0	47	99
410-421	29	0	76	99
400-409	48	1	124	98
390-399	88	1	212	97
380-389[2]	119	1	331	96
370-379	219	3	550	93
360-369	435	5	985	88
350-359	1,145	14	2,130	74
340-349	1,405	17	3,535	57
330-339	1,969	24	5,504	33
320-329	1,305	16	6,809	17
310-319	923	11	7,732	6
300-309	340	4	8,072	2
290-299	109	1	8,181	1
280-289	26	0	8,207	0
270-279	15	0	8,222	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.A.5: Demographic Summary for All Examinees ELA—July 2010

	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	8,389	2,290	27	6,099	73	337	64	56	56	42	52	2.1
Grade[3]												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-	-
Twelfth	7,046	1,754	25	5,292	75	335	63	54	55	41	51	2.1
Adult Education	1,343	536	40	807	60	346	68	62	62	49	57	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	4,598	1,240	27	3,358	73	335	63	54	56	42	51	2.1
Female	3,781	1,047	28	2,734	72	338	65	57	56	43	53	2.1
Unknown	10	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	37	11	30	26	70	337	63	55	57	45	53	2.0
Asian	735	134	18	601	82	331	57	53	53	39	53	2.0
Pacific Islander	62	21	34	41	66	338	59	54	56	43	59	2.1
Filipino	171	36	21	135	79	336	56	55	57	40	54	2.1
Hispanic or Latino	5,702	1,488	26	4,214	74	336	65	55	56	42	52	2.1
African American	779	253	32	526	68	339	65	58	59	44	52	2.1
White (not of Hispanic origin)	639	277	43	362	57	349	66	62	63	49	58	2.3
Two or More Races	264	70	27	194	73	331	59	53	54	40	49	1.9
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	2,522	873	35	1,649	65	341	65	58	59	44	53	2.1
Initially Fluent English Proficient (IFEP)	287	102	36	185	64	340	65	57	58	45	55	2.1
Reclassified Fluent English Proficient (RFEP)	276	121	44	155	56	346	68	62	61	48	61	2.2
English-Learner Students	4,174	719	17	3,455	83	330	62	52	52	39	49	2.0
Unknown	1,130	475	42	655	58	347	69	62	62	49	57	2.2
Economically Disadvantaged												
No	1,297	408	31	889	69	339	64	57	57	44	54	2.1
Yes	5,313	1,203	23	4,110	77	334	63	54	54	41	51	2.0
Unknown	1,779	679	38	1,100	62	344	67	60	60	47	55	2.1
Special Education Program Participation												
Students Receiving Services	880	141	16	739	84	325	56	48	49	36	45	1.9
Students Not Receiving Services	7,509	2,149	29	5,360	71	338	65	56	57	43	53	2.1

[1] Results for groups with less 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 2010

	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	8,222	2,130	26	6,092	74	338	53	48	48	42	36
Grade[3]											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-
Twelfth	6,750	1,563	23	5,187	77	337	52	48	47	41	35
Adult Education	1,472	567	39	905	61	346	58	52	53	49	39
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	3,764	986	26	2,778	74	338	52	49	47	43	34
Female	4,447	1,139	26	3,308	74	339	54	48	49	41	36
Unknown	11	5	45	6	55	340	52	48	54	40	39
Race/Ethnicity											
American Indian or Alaska Native	45	12	27	33	73	338	54	50	48	42	33
Asian	247	118	48	129	52	354	57	59	57	51	48
Pacific Islander	57	17	30	40	70	340	53	50	49	42	37
Filipino	109	26	24	83	76	338	50	48	49	41	40
Hispanic or Latino	5,628	1,336	24	4,292	76	337	52	47	48	42	35
African American	1,126	253	22	873	78	336	53	48	47	39	35
White (not of Hispanic origin)	711	288	41	423	59	346	58	53	53	48	40
Two or More Races	299	80	27	219	73	338	51	48	48	43	35
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	3,330	875	26	2,455	74	338	54	49	48	41	35
Initially Fluent English Proficient (IFEP)	375	102	27	273	73	339	55	49	48	42	37
Reclassified Fluent English Proficient (RFEP)	585	181	31	404	69	341	56	49	51	43	37
English-Learner Students	2,691	479	18	2,212	82	334	49	46	46	39	33
Unknown	1,241	493	40	748	60	347	58	52	54	50	40
Economically Disadvantaged											
No	1,461	391	27	1,070	73	340	55	49	49	42	36
Yes	4,836	1,057	22	3,779	78	336	51	47	47	40	34
Unknown	1,925	682	35	1,243	65	344	56	51	52	47	38
Special Education Program Participation											
Students Receiving Services	677	88	13	589	87	326	44	43	40	34	29
Students Not Receiving Services	7,545	2,042	27	5,503	73	339	54	49	49	43	36

[1] Results for groups with less 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 2010

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		8,389	7,986	95	279	3	124	1	403	5
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	7,046	6,787	96	187	3	72	1	259	4
	Adult Education	1,343	1,199	89	92	7	52	4	144	11
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	4,598	4,357	95	172	4	69	2	241	5
	Female	3,781	3,619	96	107	3	55	1	162	4
	Unknown	10	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	37	37	100	0	0	0	0	0	0
	Asian	735	716	97	11	1	8	1	19	3
	Pacific Islander	62	61	98			1	2	1	2
	Filipino	171	165	96	5	3	1	1	6	4
	Hispanic or Latino	5,702	5,496	96	157	3	49	1	206	4
	African American	779	740	95	28	4	11	1	39	5
	White (not of Hispanic origin)	639	520	81	68	11	51	8	119	19
	Two or More Races	264	251	95	10	4	3	1	13	5
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	2,522	2,325	92	134	5	63	2	197	8
	Initially Fluent English Proficient (IFEP)	287	270	94	15	5	2	1	17	6
	Reclassified Fluent English Proficient (RFEP)	276	259	94	13	5	4	1	17	6
	English-Learner Students	4,174	4,135	99	34	1	5	0	39	1
	Unknown	1,130	997	88	83	7	50	4	133	12
Economically Disadvantaged	No	1,297	1,212	93	53	4	32	2	85	7
	Yes	5,313	5,171	97	107	2	35	1	142	3
	Unknown	1,779	1,603	90	119	7	57	3	176	10
Special Education Program Participation	Receiving Services	880	864	98	12	1	4	0	16	2
	Not Receiving Services	7,509	7,122	95	267	4	120	2	387	5

[1] Results for groups with less 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 2010

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		8,222	7,891	96	284	3	47	1	331	4
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	6,750	6,546	97	170	3	34	1	204	3
	Adult Education	1,472	1,345	91	114	8	13	1	127	9
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	3,764	3,574	95	165	4	25	1	190	5
	Female	4,447	4,306	97	119	3	22	0	141	3
	Unknown	11	11	100	0	0	0	0	0	0
Race/Ethnicity	American Indian or Alaskan Native	45	45	100	0	0	0	0	0	0
	Asian	247	185	75	46	19	16	6	62	25
	Pacific Islander	57	55	96	1	2	1	2	2	4
	Filipino	109	105	96	3	3	1	1	4	4
	Hispanic or Latino	5,628	5,476	97	139	2	13	0	152	3
	African American	1,126	1,109	98	17	2			17	2
	White (not of Hispanic origin)	711	633	89	64	9	14	2	78	11
	Two or More Races	299	283	95	14	5	2	1	16	5
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	3,330	3,208	96	99	3	23	1	122	4
	Initially Fluent English Proficient (IFEP)	375	361	96	12	3	2	1	14	4
	Reclassified Fluent English Proficient (RFEP)	585	575	98	10	2	0	0	10	2
	English-Learner Students	2,691	2,631	98	53	2	7	0	60	2
	Unknown	1,241	1,116	90	110	9	15	1	125	10
Economically Disadvantaged	No	1,461	1,377	94	65	4	19	1	84	6
	Yes	4,836	4,734	98	90	2	12	0	102	2
	Unknown	1,925	1,780	92	129	7	16	1	145	8
Special Education Program Participation	Receiving Services	677	671	99	5	1	1	0	6	1
	Not Receiving Services	7,545	7,220	96	279	4	46	1	325	4

[1] Results for groups with less 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 2010

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	276	297	320	336	350	379	413	337	25	8,389
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	276	295	320	334	348	374	406	335	24	7,046
Adult Education	291	306	330	344	360	399	432	346	28	1,343
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	293	318	334	350	381	413	335	27	4,598
Female	283	300	324	338	350	376	413	338	24	3,781
Unknown	293	293	338	339	356	369	369	340	23	10
Race/Ethnicity										
American Indian or Alaska Native	283	293	318	340	354	374	379	337	24	37
Asian	275	295	316	332	346	367	406	331	23	735
Pacific Islander	278	289	324	341	352	367	409	338	25	62
Filipino	295	304	320	336	346	374	399	336	21	171
Hispanic or Latino	278	297	322	336	350	374	402	336	24	5,702
African American	275	298	324	340	354	381	413	339	25	779
White (not of Hispanic origin)	275	297	324	344	369	413	445	349	35	639
Two or More Races	275	287	310	330	350	379	413	331	29	264
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	275	295	324	340	356	390	422	341	28	2,522
Initially Fluent English Proficient (IFEP)	275	295	324	342	356	384	402	340	27	287
Reclassified Fluent English Proficient (RFEP)	283	314	333	346	360	384	406	346	22	276
English-Learner Students	278	295	318	332	344	362	379	330	20	4,174
Unknown	291	306	330	344	360	402	432	347	28	1,130
Economically Disadvantaged										
No	275	295	322	338	354	387	427	339	28	1,297
Yes	276	297	320	334	348	369	396	334	23	5,313
Unknown	280	300	326	342	358	396	427	344	28	1,779
Special Education Program Participation										
Students Receiving Services	275	283	308	324	340	367	390	325	25	880
Students Not Receiving Services	280	298	322	338	352	381	413	338	25	7,509

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation.

[3] Results for groups with less than 11 students are not reported.

Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2010

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	296	308	324	337	350	374	408	338	22	8,222
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	293	306	323	336	348	370	405	337	21	6,750
Adult Education	303	314	332	343	355	390	416	346	23	1,472
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	293	306	323	336	350	381	412	338	24	3,764
Female	298	310	326	337	350	370	405	339	20	4,447
Unknown	319	319	326	336	353	362	362	340	15	11
Race/Ethnicity										
American Indian or Alaska Native	287	312	330	339	350	364	368	338	17	45
Asian	293	306	324	346	381	438	450	354	39	247
Pacific Islander	308	308	323	337	352	379	438	340	23	57
Filipino	308	314	326	336	348	374	390	338	20	109
Hispanic or Latino	296	308	324	337	348	370	396	337	20	5,628
African American	296	308	324	336	348	366	383	336	18	1,126
White (not of Hispanic origin)	293	306	328	343	359	402	445	346	29	711
Two or More Races	293	308	323	336	350	383	408	338	23	299
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	293	306	324	337	350	374	416	338	22	3,330
Initially Fluent English Proficient (IFEP)	296	308	326	337	350	374	420	339	21	375
Reclassified Fluent English Proficient (RFEP)	301	317	330	341	352	366	385	341	16	585
English-Learner Students	296	306	321	334	344	364	399	334	19	2,691
Unknown	301	314	330	343	357	396	431	347	25	1,241
Economically Disadvantaged										
No	293	308	326	337	350	383	438	340	24	1,461
Yes	296	306	323	336	346	366	393	336	19	4,836
Unknown	298	310	328	341	355	388	416	344	24	1,925
Special Education Program Participation										
Students Receiving Services	280	298	312	324	339	361	379	326	20	677
Students Not Receiving Services	296	310	326	337	350	376	408	339	21	7,545

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation.

[3] Results for groups with less than 11 students are not reported.

Appendix 8.B: Frequency Distributions and Demographic Summaries—October 2010

Table 8.B.1: Frequency Distributions, ELA—October 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	174	0	174	99
440-449	85	0	259	99
430-439	234	1	493	99
420-429	288	1	781	98
410-419	498	1	1,279	97
400-409	415	1	1,694	96
390-399	1,004	2	2,698	94
380-389	1,038	2	3,736	91
370-379	2,040	5	5,776	87
360-369	4,188	10	9,964	77
350-359[1]	6,089	14	16,053	63
340-349	7,004	16	23,057	47
330-339	6,284	15	29,341	32
320-329	4,789	11	34,130	21
310-319	4,187	10	38,317	11
300-309	2,253	5	40,570	6
290-299	1,395	3	41,965	3
280-289	723	2	42,688	1
270-279	506	1	43,194	0

[1] Passing Score = 350

Table 8.B.2: Frequency Distributions, Mathematics—October 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	204	0	204	99
440-449	91	0	295	99
430-439	234	1	529	99
420-429	171	0	700	98
410-419	419	1	1,119	97
400-409	501	1	1,620	96
390-399	551	1	2,171	95
380-389	1,156	3	3,327	92
370-379	1,338	3	4,665	89
360-369	3,443	8	8,108	80
350-359[1]	4,907	12	13,015	69
340-349	7,362	18	20,377	51
330-339	7,704	19	28,081	32
320-329	5,783	14	33,864	18
310-319	4,607	11	38,471	7
300-309	2,096	5	40,567	2
290-299	721	2	41,288	1
280-289	123	0	41,411	0
270-279	108	0	41,519	0

[1] Passing Score = 350

Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	174	0	174	99
440-449	85	0	259	99
430-439	234	1	493	99
420-429	288	1	781	98
410-419	498	1	1,279	97
403-409[1]	208	0	1,487	97
390-402	1,211	3	2,698	94
380-389[2]	1,038	2	3,736	91
370-379	2,040	5	5,776	87
360-369	4,188	10	9,964	77
350-359	6,089	14	16,053	63
340-349	7,004	16	23,057	47
330-339	6,284	15	29,341	32
320-329	4,789	11	34,130	21
310-319	4,187	10	38,317	11
300-309	2,253	5	40,570	6
290-299	1,395	3	41,965	3
280-289	723	2	42,688	1
270-279	506	1	43,194	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	204	0	204	99
440-449	91	0	295	99
430-439	234	1	529	99
422-429[1]	171	0	700	98
410-421	419	1	1,119	97
400-409	501	1	1,620	96
390-399	551	1	2,171	95
380-389[2]	1,156	3	3,327	92
370-379	1,338	3	4,665	89
360-369	3,443	8	8,108	80
350-359	4,907	12	13,015	69
340-349	7,362	18	20,377	51
330-339	7,704	19	28,081	32
320-329	5,783	14	33,864	18
310-319	4,607	11	38,471	7
300-309	2,096	5	40,567	2
290-299	721	2	41,288	1
280-289	123	0	41,411	0
270-279	108	0	41,519	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2010

	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	43,194	16,053	37	27,141	63	342	65	56	58	52	56	2.1
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	17,566	7,092	40	10,474	60	343	67	57	59	52	56	2.1
Twelfth	23,687	8,117	34	15,570	66	341	64	55	57	51	55	2.1
Adult Education	1,941	844	43	1,097	57	350	68	61	62	58	60	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	25,159	8,954	36	16,205	64	340	66	56	57	50	54	2.0
Female	17,984	7,077	39	10,907	61	345	65	57	59	54	58	2.2
Unknown	51	22	43	29	57	341	65	54	58	51	55	1.9
Race/Ethnicity												
American Indian or Alaska Native	325	118	36	207	64	340	68	57	56	50	52	2.1
Asian	2,618	929	35	1,689	65	343	59	56	57	56	59	2.0
Pacific Islander	327	135	41	192	59	345	66	58	58	54	56	2.2
Filipino	581	247	43	334	57	348	68	59	59	57	60	2.2
Hispanic or Latino	26,609	9,178	34	17,431	66	340	64	55	57	51	55	2.1
African American	4,688	1,743	37	2,945	63	340	66	57	57	49	53	2.1
White (not of Hispanic origin)	5,976	2,913	49	3,063	51	352	72	62	62	56	59	2.2
Two or More Races	2,070	790	38	1,280	62	343	64	57	58	53	56	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	17,553	7,654	44	9,899	56	347	70	60	60	53	56	2.2
Initially Fluent English Proficient (IFEP)	1,553	781	50	772	50	351	71	61	62	57	60	2.2
Reclassified Fluent English Proficient (RFEP)	2,227	1,376	62	851	38	355	73	64	65	59	64	2.2
English-Learner Students	19,377	5,148	27	14,229	73	335	60	52	54	49	53	2.0
Unknown	2,484	1,094	44	1,390	56	349	68	61	61	57	59	2.1
Economically Disadvantaged												
No	9,902	4,734	48	5,168	52	351	70	61	62	56	59	2.2
Yes	27,622	9,020	33	18,602	67	338	63	54	56	50	54	2.0
Unknown	5,670	2,299	41	3,371	59	346	67	59	59	54	57	2.1
Special Education Program Participation												
Students Receiving Services	8,248	1,442	17	6,806	83	326	58	47	48	42	45	1.9
Students Not Receiving Services	34,946	14,611	42	20,335	58	346	67	59	60	54	58	2.1

[1] Results for groups with less 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 2010

	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	41,519	13,015	31	28,504	69	342	52	52	52	43	39
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	17,053	5,911	35	11,142	65	343	53	54	53	43	41
Twelfth	22,397	6,466	29	15,931	71	341	52	51	51	42	39
Adult Education	2,069	638	31	1,431	69	343	57	51	53	45	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	20,854	6,562	31	14,292	69	341	53	53	51	43	38
Female	20,606	6,434	31	14,172	69	342	52	52	53	42	41
Unknown	59	19	32	40	68	339	48	50	49	43	38
Race/Ethnicity											
American Indian or Alaska Native	354	100	28	254	72	339	52	51	48	41	37
Asian	1,492	832	56	660	44	365	59	65	64	57	57
Pacific Islander	339	111	33	228	67	343	51	51	54	43	41
Filipino	524	210	40	314	60	348	54	55	58	47	45
Hispanic or Latino	25,461	7,096	28	18,365	72	339	51	50	50	41	38
African American	5,553	1,415	25	4,138	75	337	50	50	49	38	37
White (not of Hispanic origin)	5,899	2,527	43	3,372	57	351	58	58	56	49	44
Two or More Races	1,897	724	38	1,173	62	347	55	55	54	47	43
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	19,421	6,545	34	12,876	66	343	54	53	52	43	40
Initially Fluent English Proficient (IFEP)	1,797	726	40	1,071	60	348	57	56	56	46	42
Reclassified Fluent English Proficient (RFEP)	3,255	1,322	41	1,933	59	348	58	55	57	45	43
English-Learner Students	14,418	3,495	24	10,923	76	338	48	50	49	40	38
Unknown	2,628	927	35	1,701	65	346	57	53	54	46	40
Economically Disadvantaged											
No	10,267	4,249	41	6,018	59	350	57	57	56	47	44
Yes	25,456	6,862	27	18,594	73	338	50	50	50	40	38
Unknown	5,796	1,904	33	3,892	67	344	55	52	53	45	39
Special Education Program Participation											
Students Receiving Services	6,518	863	13	5,655	87	327	42	45	40	34	31
Students Not Receiving Services	35,001	12,152	35	22,849	65	345	54	54	54	44	41

[1] Results for groups with less 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 2010

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		43,194	39,458	91	2,249	5	1,487	3	3,736	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	17,566	15,946	91	1,027	6	593	3	1,620	9
	Twelfth	23,687	21,872	92	1,035	4	780	3	1,815	8
	Adult Education	1,941	1,640	84	187	10	114	6	301	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	25,159	23,253	92	1,233	5	673	3	1,906	8
	Female	17,984	16,160	90	1,015	6	809	4	1,824	10
	Unknown	51	45	88	1	2	5	10	6	12
Race/Ethnicity	American Indian or Alaskan	325	297	91	21	6	7	2	28	9
	Asian	2,618	2,372	91	133	5	113	4	246	9
	Pacific Islander	327	290	89	25	8	12	4	37	11
	Filipino	581	518	89	31	5	32	6	63	11
	Hispanic or Latino	26,609	25,130	94	1,062	4	417	2	1,479	6
	African American	4,688	4,324	92	253	5	111	2	364	8
	White (not of Hispanic origin)	5,976	4,711	79	593	10	672	11	1,265	21
	Two or More Races	2,070	1,816	88	131	6	123	6	254	12
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	17,553	15,121	86	1,335	8	1,097	6	2,432	14
	Initially Fluent English Proficient (IFEP)	1,553	1,312	84	148	10	93	6	241	16
	Reclassified Fluent English Proficient (RFEP)	2,227	1,971	89	186	8	70	3	256	11
	English-Learner Students	19,377	18,971	98	341	2	65	0	406	2
	Unknown	2,484	2,083	84	239	10	162	7	401	16
Economically Disadvantaged	No	9,902	8,154	82	872	9	876	9	1,748	18
	Yes	27,622	26,391	96	919	3	312	1	1,231	4
	Unknown	5,670	4,913	87	458	8	299	5	757	13
Special Education Program Participation	Receiving Services	8,248	8,116	98	109	1	23	0	132	2
	Not Receiving Services	34,946	31,342	90	2,140	6	1,464	4	3,604	10

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

Table 8.B.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—October 2010

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		41,519	38,192	92	2,627	6	700	2	3,327	8
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	17,053	15,580	91	1,175	7	298	2	1,473	9
	Twelfth	22,397	20,708	92	1,304	6	385	2	1,689	8
	Adult Education	2,069	1,904	92	148	7	17	1	165	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	20,854	18,993	91	1,474	7	387	2	1,861	9
	Female	20,606	19,144	93	1,150	6	312	2	1,462	7
	Unknown	59	55	93	3	5	1	2	4	7
Race/Ethnicity	American Indian or Alaskan	354	327	92	21	6	6	2	27	8
	Asian	1,492	1,009	68	311	21	172	12	483	32
	Pacific Islander	339	305	90	28	8	6	2	34	10
	Filipino	524	459	88	55	10	10	2	65	12
	Hispanic or Latino	25,461	24,328	96	995	4	138	1	1,133	4
	African American	5,553	5,297	95	222	4	34	1	256	5
	White (not of Hispanic	5,899	4,868	83	772	13	259	4	1,031	17
	Two or More Races	1,897	1,599	84	223	12	75	4	298	16
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	19,421	17,494	90	1,509	8	418	2	1,927	10
	Initially Fluent English Proficient (IFEP)	1,797	1,572	87	161	9	64	4	225	13
	Reclassified Fluent English Proficient (RFEP)	3,255	3,057	94	162	5	36	1	198	6
	English-Learner Students	14,418	13,757	95	531	4	130	1	661	5
	Unknown	2,628	2,312	88	264	10	52	2	316	12
Economically Disadvantaged	No	10,267	8,654	84	1,182	12	431	4	1,613	16
	Yes	25,456	24,367	96	950	4	139	1	1,089	4
	Unknown	5,796	5,171	89	495	9	130	2	625	11
Special Education Program Participation	Receiving Services	6,518	6,410	98	98	2	10	0	108	2
	Not Receiving Services	35,001	31,782	91	2,529	7	690	2	3,219	9

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

Table 8.B.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—October 2010

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	279	297	325	342	358	393	433	342	29	43,194
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	279	297	325	342	360	396	433	343	29	17,566
Twelfth	277	297	323	340	356	393	433	341	29	23,687
Adult Education	289	308	331	346	365	406	433	350	30	1,941
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	295	321	340	356	390	423	340	29	25,159
Female	283	302	327	344	358	402	439	345	29	17,984
Unknown	279	283	306	340	362	423	433	341	39	51
Race/Ethnicity										
American Indian or Alaska Native	275	291	321	338	360	393	418	340	30	325
Asian	279	299	325	340	356	402	439	343	30	2,618
Pacific Islander	283	299	325	344	360	399	433	345	30	327
Filipino	289	306	332	344	360	406	450	348	29	581
Hispanic or Latino	281	299	325	340	356	381	414	340	26	26,609
African American	275	293	321	340	358	390	423	340	29	4,688
White (not of Hispanic origin)	277	295	326	348	374	428	450	352	38	5,976
Two or More Races	275	289	321	340	360	410	446	343	34	2,070
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	275	297	325	344	365	410	446	347	33	17,553
Initially Fluent English Proficient (IFEP)	281	302	332	350	369	406	439	351	31	1,553
Reclassified Fluent English Proficient (RFEP)	295	315	340	354	367	396	428	355	24	2,227
English-Learner Students	279	295	321	336	350	369	390	335	23	19,377
Unknown	283	302	329	346	365	410	439	349	32	2,484
Economically Disadvantaged										
No	279	299	329	348	369	423	450	351	35	9,902
Yes	279	297	323	340	354	379	406	338	25	27,622
Unknown	277	299	325	342	362	406	439	346	32	5,670
Special Education Program Participation										
Students Receiving Services	275	287	310	327	342	367	387	326	24	8,248
Students Not Receiving Services	283	302	329	344	360	399	439	346	29	34,946

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

.

Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2010

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	293	306	324	339	355	392	430	342	26	41,519
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	293	306	324	341	356	392	430	343	27	17,053
Twelfth	293	306	322	339	353	392	436	341	26	22,397
Adult Education	301	313	328	339	355	389	419	343	24	2,069
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	291	304	322	339	355	395	436	341	28	20,854
Female	296	308	326	341	355	389	430	342	25	20,606
Unknown	293	301	319	333	358	384	424	339	27	59
Race/Ethnicity										
American Indian or Alaska Native	288	304	321	333	353	389	436	339	27	354
Asian	299	313	335	355	395	444	450	365	40	1,492
Pacific Islander	291	306	322	339	355	395	430	343	27	339
Filipino	296	310	332	346	360	400	430	348	27	524
Hispanic or Latino	293	306	324	339	353	378	407	339	22	25,461
African American	288	304	321	335	351	378	411	337	24	5,553
White (not of Hispanic origin)	296	308	328	346	366	419	450	351	33	5,899
Two or More Races	293	306	324	341	364	415	450	347	33	1,897
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	293	304	324	341	356	397	436	343	28	19,421
Initially Fluent English Proficient (IFEP)	299	308	330	344	360	411	450	348	29	1,797
Reclassified Fluent English Proficient (RFEP)	301	317	335	346	358	382	424	348	21	3,255
English-Learner Students	293	306	322	335	349	375	419	338	23	14,418
Unknown	296	310	328	341	358	404	436	346	28	2,628
Economically Disadvantaged										
No	296	308	328	344	364	415	450	350	32	10,267
Yes	293	306	322	337	351	375	407	338	23	25,456
Unknown	293	306	324	339	356	400	436	344	28	5,796
Special Education Program Participation										
Students Receiving Services	284	299	313	324	341	364	387	327	21	6,518
Students Not Receiving Services	296	308	328	342	356	395	436	345	27	35,001

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

Appendix 8.C: Frequency Distributions and Demographic Summaries—November 2010

Table 8.C.1: Frequency Distributions, ELA—November 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	400	0	400	99
440-449	238	0	638	99
430-439	599	1	1,237	99
420-429	840	1	2,077	98
410-419	1,003	1	3,080	97
400-409	1,828	2	4,908	95
390-399	2,163	2	7,071	93
380-389	3,885	4	10,956	89
370-379	5,986	6	16,942	83
360-369	11,571	11	28,513	72
350-359[1]	12,206	12	40,719	60
340-349	18,756	18	59,475	42
330-339	13,294	13	72,769	29
320-329	10,114	10	82,883	19
310-319	7,822	8	90,705	12
300-309	5,420	5	96,125	6
290-299	4,042	4	100,167	2
280-289	1,262	1	101,429	1
270-279	1,187	1	102,616	0

[1] Passing Score = 350

Table 8.C.2: Frequency Distributions, Mathematics—November 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	806	1	806	99
440-449	285	0	1,091	99
430-439	687	1	1,778	98
420-429	744	1	2,522	97
410-419	797	1	3,319	97
400-409	1,322	1	4,641	95
390-399	2,062	2	6,703	93
380-389	2,582	3	9,285	91
370-379	4,795	5	14,080	86
360-369	8,047	8	22,127	77
350-359[1]	11,926	12	34,053	65
340-349	17,861	18	51,914	47
330-339	17,677	18	69,591	29
320-329	12,603	13	82,194	16
310-319	9,548	10	91,742	7
300-309	4,332	4	96,074	2
290-299	1,382	1	97,456	1
280-289	418	0	97,874	0
270-279	282	0	98,156	0

[1] Passing Score = 350

Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	400	0	400	99
440-449	238	0	638	99
430-439	599	1	1,237	99
420-429	840	1	2,077	98
410-419	1,003	1	3,080	97
403-409[1]	1,188	1	4,268	96
390-402	2,803	3	7,071	93
380-389[2]	3,885	4	10,956	89
370-379	5,986	6	16,942	83
360-369	11,571	11	28,513	72
350-359	12,206	12	40,719	60
340-349	18,756	18	59,475	42
330-339	13,294	13	72,769	29
320-329	10,114	10	82,883	19
310-319	7,822	8	90,705	12
300-309	5,420	5	96,125	6
290-299	4,042	4	100,167	2
280-289	1,262	1	101,429	1
270-279	1,187	1	102,616	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	806	1	806	99
440-449	285	0	1,091	99
430-439	687	1	1,778	98
422-429[1]	361	0	2,139	98
410-421	1,180	1	3,319	97
400-409	1,322	1	4,641	95
390-399	2,062	2	6,703	93
380-389[2]	2,582	3	9,285	91
370-379	4,795	5	14,080	86
360-369	8,047	8	22,127	77
350-359	11,926	12	34,053	65
340-349	17,861	18	51,914	47
330-339	17,677	18	69,591	29
320-329	12,603	13	82,194	16
310-319	9,548	10	91,742	7
300-309	4,332	4	96,074	2
290-299	1,382	1	97,456	1
280-289	418	0	97,874	0
270-279	282	0	98,156	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2010

	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	102,616	40,719	40	61,897	60	345	66	60	62	54	52	2.1
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	67,171	28,765	43	38,406	57	347	68	61	63	55	53	2.1
Twelfth	31,946	10,451	33	21,495	67	341	64	57	59	51	50	2.1
Adult Education	3,499	1,503	43	1,996	57	350	68	63	65	58	53	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	59,195	22,290	38	36,905	62	342	67	59	60	52	50	2.1
Female	43,350	18,398	42	24,952	58	348	66	60	64	56	54	2.2
Unknown	71	31	44	40	56	345	65	61	62	53	51	2.1
Race/Ethnicity												
American Indian or Alaska Native	776	345	44	431	56	346	68	61	62	54	51	2.1
Asian	5,481	2,135	39	3,346	61	346	64	58	63	56	55	2.1
Pacific Islander	759	348	46	411	54	349	68	62	64	55	54	2.3
Filipino	1,524	742	49	782	51	353	70	63	66	58	57	2.3
Hispanic or Latino	67,143	23,680	35	43,463	65	341	65	58	60	52	51	2.1
African American	10,161	3,860	38	6,301	62	342	66	59	61	51	48	2.1
White (not of Hispanic origin)	14,946	8,690	58	6,256	42	360	74	68	68	62	59	2.3
Two or More Races	1,826	919	50	907	50	354	71	65	66	58	56	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	42,815	20,690	48	22,125	52	351	70	64	65	57	54	2.2
Initially Fluent English Proficient (IFEP)	4,037	2,223	55	1,814	45	356	72	67	67	60	58	2.3
Reclassified Fluent English Proficient (RFEP)	7,831	5,048	64	2,783	36	358	73	68	69	63	61	2.3
English-Learner Students	43,979	10,972	25	33,007	75	335	61	53	57	48	48	2.0
Unknown	3,954	1,786	45	2,168	55	351	69	63	65	58	54	2.2
Economically Disadvantaged												
No	22,561	12,022	53	10,539	47	356	72	66	67	60	58	2.3
Yes	67,744	23,173	34	44,571	66	340	64	57	60	51	50	2.1
Unknown	12,311	5,524	45	6,787	55	349	68	62	64	57	54	2.2
Special Education Program Participation												
Students Receiving Services	20,608	3,854	19	16,754	81	328	58	49	52	43	42	1.9
Students Not Receiving Services	82,008	36,865	45	45,143	55	349	69	62	64	56	55	2.2

[1] Results for groups with less 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 2010

	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	98,156	34,053	35	64,103	65	344	53	51	54	47	40
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	64,176	23,987	37	40,189	63	346	54	52	55	47	42
Twelfth	30,471	8,969	29	21,502	71	341	52	49	52	46	38
Adult Education	3,509	1,097	31	2,412	69	344	54	50	55	50	35
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	48,732	17,103	35	31,629	65	344	53	52	54	46	39
Female	49,348	16,928	34	32,420	66	345	53	50	55	47	41
Unknown	76	22	29	54	71	342	54	50	53	45	34
Race/Ethnicity											
American Indian or Alaska Native	822	275	33	547	67	342	53	50	53	46	38
Asian	3,333	1,919	58	1,414	42	369	61	65	66	60	58
Pacific Islander	695	286	41	409	59	348	55	53	56	48	44
Filipino	1,410	690	49	720	51	354	57	58	61	52	48
Hispanic or Latino	63,407	19,410	31	43,997	69	341	52	49	52	45	38
African American	11,543	2,966	26	8,577	74	337	50	47	50	42	36
White (not of Hispanic origin)	15,010	7,660	51	7,350	49	357	61	59	62	53	46
Two or More Races	1,936	847	44	1,089	56	350	58	55	58	50	42
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	46,122	17,480	38	28,642	62	346	55	53	55	47	41
Initially Fluent English Proficient (IFEP)	4,670	2,063	44	2,607	56	352	58	56	59	51	44
Reclassified Fluent English Proficient (RFEP)	9,548	4,343	45	5,205	55	350	58	55	59	50	44
English-Learner Students	33,828	8,765	26	25,063	74	338	48	47	50	44	38
Unknown	3,988	1,402	35	2,586	65	346	55	52	56	50	37
Economically Disadvantaged											
No	23,406	10,893	47	12,513	53	354	59	57	60	51	45
Yes	62,516	18,414	29	44,102	71	340	51	48	52	45	38
Unknown	12,234	4,746	39	7,488	61	348	55	53	56	49	41
Special Education Program Participation											
Students Receiving Services	16,896	2,465	15	14,431	85	328	43	41	43	37	32
Students Not Receiving Services	81,260	31,588	39	49,672	61	348	56	53	57	49	42

[1] Results for groups with less 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 2010

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		102,616	91,660	89	6,688	7	4,268	4	10,956	11
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	67,171	59,546	89	4,783	7	2,842	4	7,625	11
	Twelfth	31,946	29,174	91	1,549	5	1,223	4	2,772	9
	Adult Education	3,499	2,940	84	356	10	203	6	559	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	59,195	53,658	91	3,574	6	1,963	3	5,537	9
	Female	43,350	37,938	88	3,109	7	2,303	5	5,412	12
	Unknown	71	64	90	5	7	2	3	7	10
Race/Ethnicity	American Indian or Alaskan	776	668	86	67	9	41	5	108	14
	Asian	5,481	4,808	88	321	6	352	6	673	12
	Pacific Islander	759	645	85	66	9	48	6	114	15
	Filipino	1,524	1,269	83	132	9	123	8	255	17
	Hispanic or Latino	67,143	62,727	93	3,235	5	1,181	2	4,416	7
	African American	10,161	9,304	92	621	6	236	2	857	8
	White (not of Hispanic origin)	14,946	10,820	72	2,036	14	2,090	14	4,126	28
	Two or More Races	1,826	1,419	78	210	12	197	11	407	22
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	42,815	35,367	83	4,150	10	3,298	8	7,448	17
	Initially Fluent English Proficient (IFEP)	4,037	3,232	80	481	12	324	8	805	20
	Reclassified Fluent English Proficient (RFEP)	7,831	6,695	85	882	11	254	3	1,136	15
	English-Learner Students	43,979	43,072	98	764	2	143	0	907	2
	Unknown	3,954	3,294	83	411	10	249	6	660	17
Economically Disadvantaged	No	22,561	17,564	78	2,586	11	2,411	11	4,997	22
	Yes	67,744	63,794	94	2,975	4	975	1	3,950	6
	Unknown	12,311	10,302	84	1,127	9	882	7	2,009	16
Special Education Program Participation	Receiving Services	20,608	20,170	98	359	2	79	0	438	2
	Not Receiving Services	82,008	71,490	87	6,329	8	4,189	5	10,518	13

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

Table 8.C.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—November 2010

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		98,156	88,871	91	7,146	7	2,139	2	9,285	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	64,176	57,706	90	5,054	8	1,416	2	6,470	10
	Twelfth	30,471	27,963	92	1,828	6	680	2	2,508	8
	Adult Education	3,509	3,202	91	264	8	43	1	307	9
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	48,732	43,689	90	3,812	8	1,231	3	5,043	10
	Female	49,348	45,111	91	3,330	7	907	2	4,237	9
	Unknown	76	71	93	4	5	1	1	5	7
Race/Ethnicity	American Indian or Alaskan	822	742	90	68	8	12	1	80	10
	Asian	3,333	2,193	66	596	18	544	16	1,140	34
	Pacific Islander	695	615	88	63	9	17	2	80	12
	Filipino	1,410	1,160	82	189	13	61	4	250	18
	Hispanic or Latino	63,407	59,908	94	3,088	5	411	1	3,499	6
	African American	11,543	10,976	95	497	4	70	1	567	5
	White (not of Hispanic origin)	15,010	11,681	78	2,384	16	945	6	3,329	22
	Two or More Races	1,936	1,596	82	261	13	79	4	340	18
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	46,122	40,375	88	4,340	9	1,407	3	5,747	12
	Initially Fluent English Proficient (IFEP)	4,670	3,953	85	471	10	246	5	717	15
	Reclassified Fluent English Proficient (RFEP)	9,548	8,716	91	721	8	111	1	832	9
	English-Learner Students	33,828	32,309	96	1,236	4	283	1	1,519	4
	Unknown	3,988	3,518	88	378	9	92	2	470	12
Economically Disadvantaged	No	23,406	19,105	82	3,031	13	1,270	5	4,301	18
	Yes	62,516	59,280	95	2,847	5	389	1	3,236	5
	Unknown	12,234	10,486	86	1,268	10	480	4	1,748	14
Special Education Program Participation	Receiving Services	16,896	16,608	98	266	2	22	0	288	2
	Not Receiving Services	81,260	72,263	89	6,880	8	2,117	3	8,997	11

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

Table 8.C.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—November 2010

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	279	298	326	343	362	399	433	345	30	102,616
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	279	299	328	345	362	399	433	347	30	67,171
Twelfth	275	294	320	340	356	395	433	341	30	31,946
Adult Education	292	307	330	345	367	406	433	350	30	3,499
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	294	322	342	360	395	427	342	30	59,195
Female	286	301	328	345	364	406	439	348	30	43,350
Unknown	277	296	319	345	369	399	439	345	33	71
Race/Ethnicity										
American Indian or Alaska Native	275	294	323	345	364	406	433	346	33	776
Asian	279	298	326	343	360	413	446	346	32	5,481
Pacific Islander	284	303	326	347	367	409	439	349	32	759
Filipino	290	307	332	349	369	418	446	353	32	1,524
Hispanic or Latino	279	298	324	342	358	384	413	341	27	67,143
African American	275	294	322	342	360	389	418	342	29	10,161
White (not of Hispanic origin)	279	299	336	358	384	427	450	360	38	14,946
Two or More Races	277	298	330	351	376	422	446	354	36	1,826
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	277	298	328	349	369	413	446	351	34	42,815
Initially Fluent English Proficient (IFEP)	284	305	336	354	374	413	446	356	32	4,037
Reclassified Fluent English Proficient (RFEP)	294	319	343	358	371	399	422	358	24	7,831
English-Learner Students	279	296	319	336	349	371	389	335	23	43,979
Unknown	286	305	330	347	369	409	439	351	31	3,954
Economically Disadvantaged										
No	282	301	334	354	376	422	450	356	35	22,561
Yes	277	296	322	342	356	381	409	340	26	67,744
Unknown	277	298	326	347	367	413	446	349	34	12,311
Special Education Program Participation										
Students Receiving Services	275	286	309	328	345	369	389	328	26	20,608
Students Not Receiving Services	284	303	330	347	364	406	439	349	30	82,008

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2010

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	292	305	326	340	356	398	445	344	28	98,156
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	295	307	328	342	358	398	445	346	28	64,176
Twelfth	292	305	324	339	353	395	445	341	28	30,471
Adult Education	298	310	328	340	355	392	430	344	25	3,509
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	289	303	324	340	358	401	445	344	30	48,732
Female	298	310	328	342	356	395	437	345	26	49,348
Unknown	295	305	327	340	353	382	450	342	25	76
Race/Ethnicity										
American Indian or Alaska Native	289	305	324	339	356	392	437	342	27	822
Asian	300	314	337	356	401	450	450	369	43	3,333
Pacific Islander	295	307	328	344	362	408	437	348	29	695
Filipino	300	312	333	349	368	420	450	354	31	1,410
Hispanic or Latino	295	305	326	340	353	380	415	341	23	63,407
African American	289	303	320	335	351	378	411	337	24	11,543
White (not of Hispanic origin)	295	310	333	351	374	430	450	357	35	15,010
Two or More Races	292	305	328	346	368	415	450	350	33	1,936
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	292	305	326	342	360	408	450	346	31	46,122
Initially Fluent English Proficient (IFEP)	298	310	331	347	364	425	450	352	32	4,670
Reclassified Fluent English Proficient (RFEP)	305	316	335	347	360	390	425	350	23	9,548
English-Learner Students	295	305	322	337	351	376	420	338	23	33,828
Unknown	298	310	328	342	358	404	445	346	28	3,988
Economically Disadvantaged										
No	295	310	331	347	368	425	450	354	33	23,406
Yes	292	305	324	339	353	380	411	340	24	62,516
Unknown	292	307	328	342	362	415	450	348	32	12,234
Special Education Program Participation										
Students Receiving Services	283	298	312	326	340	364	387	328	21	16,896
Students Not Receiving Services	298	310	330	344	360	404	445	348	28	81,260

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

Appendix 8.D: Frequency Distributions and Demographic Summaries—December 2010

Table 8.D.1: Frequency Distributions, ELA—December 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	3	0	3	99
440-449	3	0	6	99
430-439	5	0	11	99
420-429	11	0	22	99
410-419	19	1	41	98
400-409	14	1	55	98
390-399	38	2	93	96
380-389	41	2	134	94
370-379	86	4	220	91
360-369	226	9	446	81
350-359[1]	417	17	863	64
340-349	462	19	1,325	45
330-339	427	18	1,752	27
320-329	283	12	2,035	15
310-319	155	6	2,190	9
300-309	83	3	2,273	5
290-299	80	3	2,353	2
280-289	30	1	2,383	1
270-279	22	1	2,405	0

[1]Passing Score = 350

Table 8.D.2: Frequency Distributions, Mathematics—December 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	4	0	4	99
440-449	0	0	4	99
430-439	7	0	11	99
420-429	11	0	22	99
410-419	6	0	28	99
400-409	20	1	48	98
390-399	32	1	80	97
380-389	48	2	128	95
370-379	91	4	219	91
360-369	144	6	363	85
350-359[1]	339	14	702	71
340-349	486	20	1,188	51
330-339	503	21	1,691	31
320-329	344	14	2,035	16
310-319	236	10	2,271	7
300-309	117	5	2,388	2
290-299	35	1	2,423	0
280-289	8	0	2,431	0
270-279	2	0	2,433	0

[1]Passing Score = 350

Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	3	0	3	99
440-449	3	0	6	99
430-439	5	0	11	99
420-429	11	0	22	99
410-419	19	1	41	98
403-409[1]	14	1	55	98
390-402	38	2	93	96
380-389[2]	41	2	134	94
370-379	86	4	220	91
360-369	226	9	446	81
350-359	417	17	863	64
340-349	462	19	1,325	45
330-339	427	18	1,752	27
320-329	283	12	2,035	15
310-319	155	6	2,190	9
300-309	83	3	2,273	5
290-299	80	3	2,353	2
280-289	30	1	2,383	1
270-279	22	1	2,405	0

[1] Advanced-Level Cut = 403
[2] Proficient-Level Cut = 380

Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	4	0	4	99
440-449	0	0	4	99
430-439	7	0	11	99
422-429[1]	8	0	19	99
410-421	9	0	28	99
400-409	20	1	48	98
390-399	32	1	80	97
380-389[2]	48	2	128	95
370-379	91	4	219	91
360-369	144	6	363	85
350-359	339	14	702	71
340-349	486	20	1,188	51
330-339	503	21	1,691	31
320-329	344	14	2,035	16
310-319	236	10	2,271	7
300-309	117	5	2,388	2
290-299	35	1	2,423	0
280-289	8	0	2,431	0
270-279	2	0	2,433	0

[1] Advanced-Level Cut = 422
[2] Proficient-Level Cut = 380

Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2010

	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	2,405	863	36	1,542	64	342	62	59	59	46	58	2.1
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	303	153	50	150	50	347	61	61	62	51	59	2.2
Twelfth	1,674	507	30	1,167	70	338	61	57	57	44	56	2.0
Adult Education	428	203	47	225	53	350	67	66	63	53	61	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	1,280	420	33	860	67	339	62	57	58	45	56	2.1
Female	1,117	440	39	677	61	344	62	61	61	48	59	2.1
Unknown	8	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	8	-	-	-	-	-	-	-	-	-	-	-
Asian	153	36	24	117	76	336	53	56	56	41	58	2.0
Pacific Islander	21	8	38	13	62	337	55	50	57	44	58	2.2
Filipino	47	15	32	32	68	344	60	61	57	46	64	2.1
Hispanic or Latino	1,648	568	34	1,080	66	341	63	59	59	46	57	2.1
African American	235	91	39	144	61	340	56	59	57	45	56	2.2
White (not of Hispanic origin)	185	101	55	84	45	355	68	66	67	56	63	2.2
Two or More Races	108	41	38	67	62	342	64	59	60	46	56	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	607	271	45	336	55	348	63	62	62	50	61	2.2
Initially Fluent English Proficient (IFEP)	43	20	47	23	53	349	73	62	65	51	58	2.2
Reclassified Fluent English Proficient (RFEP)	79	38	48	41	52	346	64	61	62	48	59	2.2
English-Learner Students	1,198	329	27	869	73	337	61	56	57	43	56	2.0
Unknown	478	205	43	273	57	345	62	62	60	50	57	2.1
Economically Disadvantaged												
No	328	148	45	180	55	349	65	62	63	51	61	2.2
Yes	1,518	493	32	1,025	68	339	61	57	58	45	57	2.1
Unknown	559	222	40	337	60	344	62	61	60	49	57	2.1
Special Education Program Participation												
Students Receiving Services	277	61	22	216	78	331	56	50	53	41	53	2.0
Students Not Receiving Services	2,128	802	38	1,326	62	343	63	60	60	47	58	2.1

[1] Results for groups with less 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 2010

	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	2,433	702	29	1,731	71	340	57	49	49	44	36
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	269	108	40	161	60	345	57	52	53	48	39
Twelfth	1,682	408	24	1,274	76	338	55	48	48	42	35
Adult Education	482	186	39	296	61	347	63	54	51	50	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	1,006	303	30	703	70	340	55	50	48	45	35
Female	1,417	396	28	1,021	72	341	58	49	49	44	37
Unknown	10	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity											
American Indian or Alaska Native	13	5	38	8	62	342	54	56	46	46	37
Asian	51	16	31	35	69	343	50	52	52	46	41
Pacific Islander	19	1	5	18	95	330	50	41	49	33	28
Filipino	25	5	20	20	80	339	56	50	46	41	38
Hispanic or Latino	1,581	422	27	1,159	73	339	56	48	49	44	35
African American	385	98	25	287	75	337	57	48	46	40	36
White (not of Hispanic origin)	231	118	51	113	49	353	66	58	57	52	41
Two and More Races	128	37	29	91	71	343	58	54	49	46	36
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	891	279	31	612	69	342	59	51	50	44	37
Initially Fluent English Proficient (IFEP)	57	14	25	43	75	340	61	49	49	41	32
Reclassified Fluent English Proficient (RFEP)	162	58	36	104	64	345	62	50	54	48	38
English-Learner Students	804	159	20	645	80	335	51	46	46	41	35
Unknown	519	192	37	327	63	345	60	51	51	49	37
Economically Disadvantaged											
No	444	155	35	289	65	345	59	53	52	47	38
Yes	1,391	342	25	1,049	75	338	55	47	48	42	35
Unknown	598	205	34	393	66	343	59	51	50	47	36
Special Education Program Participation											
Students Receiving Services	246	40	16	206	84	330	47	45	42	36	31
Students Not Receiving Services	2,187	662	30	1,525	70	342	58	50	50	45	36

[1] Results for groups with less 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 2010

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		2,405	2,271	94	79	3	55	2	134	6
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	303	264	87	21	7	18	6	39	13
	Twelfth	1,674	1,637	98	22	1	15	1	37	2
	Adult Education	428	370	86	36	8	22	5	58	14
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	1,280	1,218	95	39	3	23	2	62	5
	Female	1,117	1,045	94	40	4	32	3	72	6
	Unknown	8	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	8	-	-	-	-	-	-	-	-
	Asian	153	150	98	0	0	3	2	3	2
	Pacific Islander	21	21	100	0	0	0	0	0	0
	Filipino	47	44	94	1	2	2	4	3	6
	Hispanic or Latino	1,648	1,581	96	42	3	25	2	67	4
	African American	235	225	96	7	3	3	1	10	4
	White (not of Hispanic origin)	185	145	78	21	11	19	10	40	22
	Two or More Races	108	99	92	7	6	2	2	9	8
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	607	546	90	32	5	29	5	61	10
	Initially Fluent English Proficient (IFEP)	43	41	95	1	2	1	2	2	5
	Reclassified Fluent English Proficient (RFEP)	79	77	97	0	0	2	3	2	3
	English-Learner Students	1,198	1,186	99	11	1	1	0	12	1
	Unknown	478	421	88	35	7	22	5	57	12
Economically Disadvantaged	No	328	287	88	18	5	23	7	41	13
	Yes	1,518	1,483	98	26	2	9	1	35	2
	Unknown	559	501	90	35	6	23	4	58	10
Special Education Program Participation	Receiving Services	277	275	99	1	0	1	0	2	1
	Not Receiving Services	2,128	1,996	94	78	4	54	3	132	6

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

Table 8.D.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—December 2010

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		2,433	2,305	95	109	4	19	1	128	5
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	269	239	89	25	9	5	2	30	11
	Twelfth	1,682	1,632	97	41	2	9	1	50	3
	Adult Education	482	434	90	43	9	5	1	48	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	1,006	946	94	50	5	10	1	60	6
	Female	1,417	1,350	95	58	4	9	1	67	5
	Unknown	10	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	13	12	92	1	8	0	0	1	8
	Asian	51	49	96	0	0	2	4	2	4
	Pacific Islander	19	19	100	0	0	0	0	0	0
	Filipino	25	23	92	2	8	0	0	2	8
	Hispanic or Latino	1,581	1,511	96	61	4	9	1	70	4
	African American	385	374	97	10	3	1	0	11	3
	White (not of Hispanic origin)	231	198	86	28	12	5	2	33	14
	Two or More Races	128	119	93	7	5	2	2	9	7
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	891	839	94	43	5	9	1	52	6
	Initially Fluent English Proficient (IFEP)	57	56	98	0	0	1	2	1	2
	Reclassified Fluent English Proficient (RFEP)	162	152	94	9	6	1	1	10	6
	English-Learner Students	804	792	99	10	1	2	0	12	1
	Unknown	519	466	90	47	9	6	1	53	10
Economically Disadvantaged	No	444	402	91	32	7	10	2	42	9
	Yes	1,391	1,354	97	35	3	2	0	37	3
	Unknown	598	549	92	42	7	7	1	49	8
Special Education Program Participation	Receiving Services	246	244	99	2	1	0	0	2	1
	Not Receiving Services	2,187	2,061	94	107	5	19	1	126	6

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

Table 8.D.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—December 2010

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	281	298	328	342	354	385	417	342	25	2,405
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	275	283	330	350	367	410	442	347	35	303
Twelfth	281	300	326	340	352	369	393	338	22	1,674
Adult Education	296	310	332	348	363	403	426	350	27	428
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	279	296	326	340	354	379	417	339	25	1,280
Female	288	302	330	344	356	387	417	344	25	1,117
Unknown	296	296	313	328	366	379	379	336	30	8
Race/Ethnicity										
American Indian or Alaska Native	300	300	326	336	373	410	410	347	37	8
Asian	279	296	322	338	348	367	413	336	24	153
Pacific Islander	296	308	322	336	352	372	379	337	22	21
Filipino	304	318	332	342	350	385	436	344	23	47
Hispanic or Latino	281	298	328	342	354	377	410	341	23	1,648
African American	275	292	326	344	358	374	406	340	26	235
White (not of Hispanic origin)	286	302	332	354	374	417	450	355	34	185
Two or More Races	283	296	328	342	356	390	421	342	28	108
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	290	306	330	346	360	399	436	348	28	607
Initially Fluent English Proficient (IFEP)	294	322	336	346	363	379	417	349	21	43
Reclassified Fluent English Proficient (RFEP)	275	292	336	346	363	377	410	346	26	79
English-Learner Students	281	298	324	338	350	367	382	337	20	1,198
Unknown	275	290	328	346	358	399	426	345	30	478
Economically Disadvantaged										
No	290	304	330	348	363	410	442	349	31	328
Yes	281	298	328	340	352	372	393	339	22	1,518
Unknown	275	292	328	344	358	393	421	344	29	559
Special Education Program Participation										
Students Receiving Services	275	290	314	334	348	363	377	331	24	277
Students Not Receiving Services	281	300	328	342	356	385	417	343	25	2,128

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2010

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	296	306	326	339	353	380	412	340	23	2,433
Grade										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	284	301	326	344	358	401	431	345	30	269
Twelfth	296	306	324	337	349	372	405	338	21	1,682
Adult Education	298	312	332	344	357	393	425	347	24	482
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	293	303	324	339	353	383	412	340	24	1,006
Female	298	308	328	339	351	378	412	341	22	1,417
Unknown	318	318	322	326	357	387	387	339	24	10
Race/Ethnicity										
American Indian or Alaska Native	303	303	322	341	355	393	393	342	25	13
Asian	303	314	330	339	353	378	450	343	25	51
Pacific Islander	301	301	324	330	341	351	351	330	13	19
Filipino	308	312	322	339	346	393	395	339	22	25
Hispanic or Latino	296	306	326	339	351	376	405	339	22	1,581
African American	290	303	324	337	351	370	398	337	21	385
White (not of Hispanic origin)	293	308	333	351	368	408	431	353	28	231
Two or More Races	296	316	327	339	353	398	425	343	25	128
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	293	310	328	341	353	385	425	342	24	891
Initially Fluent English Proficient (IFEP)	293	312	324	341	349	374	425	340	20	57
Reclassified Fluent English Proficient (RFEP)	303	314	333	344	357	380	408	345	21	162
English-Learner Students	298	306	320	335	348	366	380	335	19	804
Unknown	293	306	328	342	357	393	425	345	26	519
Economically Disadvantaged										
No	298	310	328	342	357	398	437	345	27	444
Yes	298	306	324	337	349	372	395	338	20	1,391
Unknown	290	306	328	342	355	390	425	343	25	598
Special Education Program Participation										
Students Receiving Services	284	298	314	328	342	366	378	330	20	246
Students Not Receiving Services	298	308	328	341	353	380	420	342	23	2,187

[1] Mean scale scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less than 11 students are not reported.

Appendix 8.E: Frequency Distributions and Demographic Summaries—February 2011

Table 8.E.1: Frequency Distributions, ELA—February 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5,534	4	5,534	96
440-449	2,469	2	8,003	95
430-439	5,920	4	13,923	91
420-429	7,120	5	21,043	86
410-419	7,744	5	28,787	81
400-409	12,001	8	40,788	74
390-399	11,726	8	52,514	66
380-389	14,183	9	66,697	57
370-379	12,369	8	79,066	49
360-369	11,337	7	90,403	42
350-359[1]	13,394	9	103,797	33
340-349	12,899	8	116,696	25
330-339	10,856	7	127,552	18
320-329	10,169	7	137,721	11
310-319	6,212	4	143,933	7
300-309	4,876	3	148,809	4
290-299	3,279	2	152,088	2
280-289	1,689	1	153,777	1
270-279	1,133	1	154,910	0

[1] Passing Score = 350

Table 8.E.2: Frequency Distributions, Mathematics—February 2010

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	8,060	5	8,060	95
440-449	3,105	2	11,165	93
430-439	6,496	4	17,661	89
420-429	6,521	4	24,182	84
410-419	9,485	6	33,667	78
400-409	9,297	6	42,964	72
390-399	11,534	7	54,498	65
380-389	10,778	7	65,276	58
370-379	12,497	8	77,773	50
360-369	12,138	8	89,911	42
350-359[1]	15,613	10	105,524	32
340-349	13,576	9	119,100	23
330-339	14,497	9	133,597	14
320-329	9,904	6	143,501	8
310-319	7,417	5	150,918	3
300-309	3,237	2	154,155	1
290-299	835	1	154,990	0
280-289	265	0	155,255	0
270-279	237	0	155,492	0

[1] Passing Score = 350

Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5,534	4	5,534	96
440-449	2,469	2	8,003	95
430-439	5,920	4	13,923	91
420-429	7,120	5	21,043	86
410-419	7,744	5	28,787	81
403-409[1]	7,999	5	36,786	76
390-402	15,728	10	52,514	66
380-389[2]	14,183	9	66,697	57
370-379	12,369	8	79,066	49
360-369	11,337	7	90,403	42
350-359	13,394	9	103,797	33
340-349	12,899	8	116,696	25
330-339	10,856	7	127,552	18
320-329	10,169	7	137,721	11
310-319	6,212	4	143,933	7
300-309	4,876	3	148,809	4
290-299	3,279	2	152,088	2
280-289	1,689	1	153,777	1
270-279	1,133	1	154,910	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	8,060	5	8,060	95
440-449	3,105	2	11,165	93
430-439	6,496	4	17,661	89
422-429[1]	6,521	4	24,182	84
410-421	9,485	6	33,667	78
400-409	9,297	6	42,964	72
390-399	11,534	7	54,498	65
380-389[2]	10,778	7	65,276	58
370-379	12,497	8	77,773	50
360-369	12,138	8	89,911	42
350-359	15,613	10	105,524	32
340-349	13,576	9	119,100	23
330-339	14,497	9	133,597	14
320-329	9,904	6	143,501	8
310-319	7,417	5	150,918	3
300-309	3,237	2	154,155	1
290-299	835	1	154,990	0
280-289	265	0	155,255	0
270-279	237	0	155,492	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.E.5: Demographic Summary for All Examinees, ELA February 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	154,910	103,797	67	51,113	33	370	70	72	71	68	68	2.4
Grade[3]												
Tenth	112,073	92,368	82	19,705	18	384	77	79	78	75	74	2.5
Eleventh	17,702	5,118	29	12,584	71	336	53	54	55	49	53	2.0
Twelfth	22,377	5,132	23	17,245	77	333	50	53	52	46	51	2.0
Adult Education	2,758	1,179	43	1,579	57	349	59	65	60	58	57	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	82,068	51,338	63	30,730	37	365	69	70	69	65	65	2.3
Female	72,698	52,404	72	20,294	28	377	72	75	74	71	71	2.5
Unknown	144	55	38	89	62	345	55	58	60	53	57	2.1
Race/Ethnicity												
American Indian or Alaska Native	1431	969	68	462	32	367	71	71	72	66	66	2.3
Asian	14,938	11,434	77	3,504	23	385	74	80	76	76	74	2.6
Pacific Islander	1,138	742	65	396	35	367	69	71	70	66	66	2.4
Filipino	4,851	4,048	83	803	17	387	78	81	78	77	76	2.6
Hispanic or Latino	69,565	37,500	54	32,065	46	356	62	66	65	60	62	2.2
African American	13,189	7,215	55	5,974	45	355	64	65	65	59	60	2.2
White (not of Hispanic origin)	46,018	39,156	85	6,862	15	390	82	81	81	78	77	2.6
Two or More Races	3,780	2,733	72	1,047	28	376	74	74	74	70	70	2.4
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	88,775	67,432	76	21,343	24	379	76	76	76	72	72	2.5
Initially Fluent English Proficient (IFEP)	9,428	7,980	85	1,448	15	388	78	81	80	78	76	2.6
Reclassified Fluent English Proficient (RFEP)	19,755	17,468	88	2,287	12	383	76	81	79	77	75	2.5
English-Learner Students	33,258	9,120	27	24,138	73	335	50	55	53	49	53	2.0
Unknown	3,694	1,797	49	1,897	51	353	61	66	63	59	59	2.2
Economically Disadvantaged												
No	62,275	52,808	85	9,467	15	390	80	81	80	78	77	2.6
Yes	79,110	42,524	54	36,586	46	355	63	65	65	60	61	2.2
Unknown	13,525	8,465	63	5,060	37	366	69	71	69	66	66	2.3
Special Education Program Participation												
Students Receiving Services	18,511	4,683	25	13,828	75	331	51	50	52	45	49	1.9
Students Not Receiving Services	136,399	99,114	73	37,285	27	376	73	75	74	71	71	2.4

[1] Results for groups with less 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 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	155,492	105,524	68	49,968	32	373	70	65	67	62	56
Grade											
Tenth	111,752	92,567	83	19,185	17	387	76	74	75	70	63
Eleventh	18,054	5,549	31	12,505	69	340	55	45	47	44	37
Twelfth	22,601	6,228	28	16,373	72	338	55	43	46	42	37
Adult Education	3,085	1,180	38	1,905	62	346	60	47	51	48	40
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	78,437	52,651	67	25,786	33	373	69	66	67	62	56
Female	76,906	52,812	69	24,094	31	373	71	65	68	62	56
Unknown	149	61	41	88	59	347	58	51	51	48	42
Race/Ethnicity											
American Indian or Alaska Native	1,495	956	64	539	36	365	67	62	63	59	51
Asian	13,297	11,986	90	1,311	10	407	82	83	83	79	76
Pacific Islander	1,149	768	67	381	33	369	68	65	66	59	53
Filipino	4,776	4,108	86	668	14	393	78	77	78	73	67
Hispanic or Latino	69,905	39,113	56	30,792	44	359	65	57	60	55	48
African American	14,757	7,103	48	7,654	52	352	61	54	55	50	44
White (not of Hispanic origin)	46,244	38,828	84	7,416	16	390	77	76	77	72	65
Two or More Races	3,869	2,662	69	1,207	31	375	70	66	68	63	57
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	92,401	66,929	72	25,472	28	378	72	68	70	65	58
Initially Fluent English Proficient (IFEP)	9,876	8,046	81	1,830	19	390	77	74	76	71	65
Reclassified Fluent English Proficient (RFEP)	21,159	17,861	84	3,298	16	387	78	74	76	70	63
English-Learner Students	28,062	10,867	39	17,195	61	346	57	48	51	47	42
Unknown	3,994	1,821	46	2,173	54	351	62	51	55	51	43
Economically Disadvantaged											
No	63,238	52,718	83	10,520	17	391	77	76	77	72	65
Yes	78,354	44,216	56	34,138	44	360	65	58	60	55	48
Unknown	13,900	8,590	62	5,310	38	368	68	62	64	59	53
Special Education Program Participation											
Students Receiving Services	16,600	4,504	27	12,096	73	337	51	43	44	42	37
Students Not Receiving Services	138,892	101,020	73	37,872	27	378	72	68	70	65	58

[1] Results for groups with less 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 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		154,910	88,213	57	29,911	19	36,786	24	66,697	43
Grade	Tenth	112,073	47,694	43	28,410	25	35,969	32	64,379	57
	Eleventh	17,702	16,695	94	670	4	337	2	1,007	6
	Twelfth	22,377	21,498	96	569	3	310	1	879	4
	Adult Education	2,758	2,326	84	262	9	170	6	432	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	82,068	50,862	62	15,455	19	15,751	19	31,206	38
	Female	72,698	37,233	51	14,440	20	21,025	29	35,465	49
	Unknown	144	118	82	16	11	10	7	26	18
Race/Ethnicity	American Indian or Alaskan Native	1,431	866	61	297	21	268	19	565	39
	Asian	14,938	6,263	42	2,884	19	5,791	39	8,675	58
	Pacific Islander	1,138	721	63	207	18	210	18	417	37
	Filipino	4,851	1,867	38	1,207	25	1,777	37	2,984	62
	Hispanic or Latino	69,565	51,284	74	10,710	15	7,571	11	18,281	26
	African American	13,189	9,679	73	2,058	16	1,452	11	3,510	27
	White (not of Hispanic origin)	46,018	15,689	34	11,725	25	18,604	40	30,329	66
	Two or More Races	3,780	1,844	49	823	22	1,113	29	1,936	51
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	88,775	41,169	46	20,018	23	27,588	31	47,606	54
	Initially Fluent English Proficient (IFEP)	9,428	3,587	38	2,284	24	3,557	38	5,841	62
	Reclassified Fluent English Proficient (RFEP)	19,755	8,705	44	6,008	30	5,042	26	11,050	56
	English-Learner Students	33,258	31,864	96	1,150	3	244	1	1,394	4
	Unknown	3,694	2,888	78	451	12	355	10	806	22
Economically Disadvantaged	No	62,275	21,649	35	15,175	24	25,451	41	40,626	65
	Yes	79,110	58,304	74	12,359	16	8,447	11	20,806	26
	Unknown	13,525	8,260	61	2,377	18	2,888	21	5,265	39
Special Education Program Participation	Receiving Services	18,511	17,068	92	1,005	5	438	2	1,443	8
	Not Receiving Services	136,399	71,145	52	28,906	21	36,348	27	65,254	48

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

Table 8.E.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— February 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		155,492	90,216	58	41,094	26	24,182	16	65,276	42
Grade	Tenth	111,752	48,421	43	39,522	35	23,809	21	63,331	57
	Eleventh	18,054	17,123	95	722	4	209	1	931	5
	Twelfth	22,601	21,814	97	644	3	143	1	787	3
	Adult Education	3,085	2,858	93	206	7	21	1	227	7
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	78,437	44,831	57	20,501	26	13,105	17	33,606	43
	Female	76,906	45,259	59	20,572	27	11,075	14	31,647	41
	Unknown	149	126	85	21	14	2	1	23	15
Race/Ethnicity	American Indian or Alaskan Native	1,495	982	66	387	26	126	8	513	34
	Asian	13,297	3,213	24	4,255	32	5,829	44	10,084	76
	Pacific Islander	1,149	728	63	309	27	112	10	421	37
	Filipino	4,776	1,757	37	1,784	37	1,235	26	3,019	63
	Hispanic or Latino	69,905	51,886	74	13,816	20	4,203	6	18,019	26
	African American	14,757	11,778	80	2,369	16	610	4	2,979	20
	White (not of Hispanic origin)	46,244	17,712	38	17,134	37	11,398	25	28,532	62
	Two or More Races	3,869	2,160	56	1,040	27	669	17	1,709	44
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	92,401	48,295	52	27,586	30	16,520	18	44,106	48
	Initially Fluent English Proficient (IFEP)	9,876	4,110	42	3,143	32	2,623	27	5,766	58
	Reclassified Fluent English Proficient (RFEP)	21,159	9,439	45	7,429	35	4,291	20	11,720	55
	English-Learner Students	28,062	24,969	89	2,457	9	636	2	3,093	11
	Unknown	3,994	3,403	85	479	12	112	3	591	15
Economically Disadvantaged	No	63,238	24,300	38	22,078	35	16,860	27	38,938	62
	Yes	78,354	56,979	73	15,807	20	5,568	7	21,375	27
	Unknown	13,900	8,937	64	3,209	23	1,754	13	4,963	36
Special Education Program Participation	Receiving Services	16,600	15,183	91	1,137	7	280	2	1,417	9
	Not Receiving Services	138,892	75,033	54	39,957	29	23,902	17	63,859	46

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

Table 8.E.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—February 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	283	303	341	370	400	442	450	370	41	154,910
Grade[4]										
Tenth	293	318	361	385	411	442	450	384	37	112,073
Eleventh	275	293	318	337	353	382	416	336	27	17,702
Twelfth	275	291	316	333	348	375	411	333	26	22,377
Adult Education	285	305	329	344	365	408	430	349	30	2,758
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	279	299	335	363	394	430	450	365	40	82,068
Female	291	310	346	377	408	450	450	377	40	72,698
Unknown	281	297	322	340	365	411	450	345	35	144
Race/Ethnicity										
American Indian or Alaska Native	283	305	341	368	394	430	450	367	38	1,431
Asian	285	308	353	388	420	450	450	385	43	14,938
Pacific Islander	293	307	341	365	394	430	450	367	37	1,138
Filipino	299	325	363	391	416	450	450	387	36	4,851
Hispanic or Latino	281	299	331	353	380	420	442	356	36	69,565
African American	275	295	329	355	380	420	442	355	37	13,189
White (not of Hispanic origin)	291	318	368	394	416	450	450	390	38	46,018
Two or More Races	281	303	344	380	408	450	450	376	42	3,780
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	283	307	350	382	408	442	450	379	40	88,775
Initially Fluent English Proficient (IFEP)	297	323	363	391	416	450	450	388	37	9,428
Reclassified Fluent English Proficient (RFEP)	308	337	363	382	404	436	450	383	30	19,755
English-Learner Students	279	293	318	337	350	375	397	335	25	33,258
Unknown	281	301	329	348	372	416	442	353	34	3,694
Economically Disadvantaged										
No	293	320	365	394	416	450	450	390	38	62,275
Yes	281	299	329	353	380	420	442	355	36	79,110
Unknown	279	303	337	363	397	436	450	366	41	13,525
Special Education Program Participation										
Students Receiving Services	275	285	308	327	350	388	420	331	31	18,511
Students Not Receiving Services	289	312	346	377	404	442	450	376	39	136,399

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	300	314	341	370	403	450	450	373	40	155,492
Grade[4]										
Tenth	305	324	359	385	414	450	450	387	38	111,752
Eleventh	295	305	324	339	353	380	422	340	24	18,054
Twelfth	292	305	324	337	350	372	407	338	22	22,601
Adult Education	300	314	332	343	357	387	414	346	22	3,085
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	298	312	341	370	403	450	450	373	42	78,437
Female	303	318	343	368	400	447	450	373	39	76,906
Unknown	292	308	326	343	364	403	427	347	30	149
Race/Ethnicity										
American Indian or Alaska Native	292	310	337	362	390	433	450	365	37	1,495
Asian	312	334	380	414	447	450	450	407	38	13,297
Pacific Islander	300	316	341	364	395	439	450	369	36	1,149
Filipino	312	330	364	395	422	450	450	393	37	4,776
Hispanic or Latino	298	312	336	353	380	422	450	359	34	69,905
African American	292	308	328	348	372	414	447	352	33	14,757
White (not of Hispanic origin)	305	322	362	392	418	450	450	390	38	46,244
Two or More Races	298	314	341	372	407	450	450	375	41	3,869
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	300	314	346	376	407	450	450	378	41	92,401
Initially Fluent English Proficient (IFEP)	305	324	359	390	422	450	450	390	40	9,876
Reclassified Fluent English Proficient (RFEP)	314	332	359	385	414	450	450	387	36	21,159
English-Learner Students	298	308	326	343	359	400	439	346	28	28,062
Unknown	298	314	334	346	364	407	439	351	28	3,994
Economically Disadvantaged										
No	305	324	362	392	422	450	450	391	39	63,238
Yes	298	310	334	355	383	427	450	360	36	78,354
Unknown	298	312	337	361	395	447	450	368	40	13,900
Special Education Program Participation										
Students Receiving Services	286	300	316	332	352	392	433	337	29	16,600
Students Not Receiving Services	305	321	346	374	407	450	450	378	39	138,892

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 2011

Table 8.F.1: Frequency Distributions, ELA—March 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	21,724	5	21,724	95
440-449	8,297	2	30,021	93
430-439	19,558	5	49,579	88
420-429	22,657	6	72,236	82
410-419	36,871	9	109,107	73
400-409	25,808	6	134,915	66
390-399	37,067	9	171,982	57
380-389	44,350	11	216,332	46
370-379	37,110	9	253,442	37
360-369	30,510	8	283,952	29
350-359[1]	25,402	6	309,354	23
340-349	25,848	6	335,202	17
330-339	19,709	5	354,911	12
320-329	15,066	4	369,977	8
310-319	11,839	3	381,816	5
300-309	9,257	2	391,073	3
290-299	6,094	2	397,167	1
280-289	3,102	1	400,269	0
270-279	1,792	0	402,061	0

[1] Passing Score = 350

Table 8.F.2: Frequency Distributions, Mathematics—March 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	26,615	7	26,615	93
440-449	19,698	5	46,313	88
430-439	9,917	2	56,230	86
420-429	19,734	5	75,964	81
410-419	29,187	7	105,151	74
400-409	28,409	7	133,560	67
390-399	35,906	9	169,466	58
380-389	33,337	8	202,803	50
370-379	37,533	9	240,336	40
360-369	33,684	8	274,020	32
350-359[1]	36,569	9	310,589	23
340-349	27,102	7	337,691	16
330-339	26,739	7	364,430	9
320-329	17,556	4	381,986	5
310-319	10,417	3	392,403	2
300-309	7,323	2	399,726	1
290-299	1,379	0	401,105	0
280-289	412	0	401,517	0
270-279	321	0	401,838	0

[1] Passing Score = 350

Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	21,724	5	21,724	95
440-449	8,297	2	30,021	93
430-439	19,558	5	49,579	88
420-429	22,657	6	72,236	82
410-419	36,871	9	109,107	73
403-409[1]	12,885	3	121,992	70
390-402	49,990	12	171,982	57
380-389[2]	44,350	11	216,332	46
370-379	37,110	9	253,442	37
360-369	30,510	8	283,952	29
350-359	25,402	6	309,354	23
340-349	25,848	6	335,202	17
330-339	19,709	5	354,911	12
320-329	15,066	4	369,977	8
310-319	11,839	3	381,816	5
300-309	9,257	2	391,073	3
290-299	6,094	2	397,167	1
280-289	3,102	1	400,269	0
270-279	1,792	0	402,061	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	26,615	7	26,615	93
440-449	19,698	5	46,313	88
430-439	9,917	2	56,230	86
422-429[1]	19,734	5	75,964	81
410-421	29,187	7	105,151	74
400-409	28,409	7	133,560	67
390-399	35,906	9	169,466	58
380-389[2]	33,337	8	202,803	50
370-379	37,533	9	240,336	40
360-369	33,684	8	274,020	32
350-359	36,569	9	310,589	23
340-349	27,102	7	337,691	16
330-339	26,739	7	364,430	9
320-329	17,556	4	381,986	5
310-319	10,417	3	392,403	2
300-309	7,323	2	399,726	1
290-299	1,379	0	401,105	0
280-289	412	0	401,517	0
270-279	321	0	401,838	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.F.5: Demographic Summary for All Examinees, ELA—March 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	402,061	309,354	77	92,707	23	381	76	75	78	72	73	2.4
Grade												
Tenth	354,450	294,067	83	60,383	17	387	79	78	80	75	76	2.4
Eleventh	28,021	9,782	35	18,239	65	340	58	55	61	51	53	1.9
Twelfth	16,669	4,161	25	12,508	75	334	53	51	56	47	50	1.9
Adult Education	2,921	1,344	46	1,577	54	352	64	63	65	59	58	2.0
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	207,977	152,315	73	55,662	27	376	76	73	76	69	71	2.3
Female	193,945	156,962	81	36,983	19	386	77	78	79	74	76	2.4
Unknown	139	77	55	62	45	360	67	64	68	62	63	2.1
Race/Ethnicity												
American Indian or Alaska Native	2,881	2,191	76	690	24	379	77	74	78	70	71	2.3
Asian	33,843	29,386	87	4,457	13	402	83	83	84	81	83	2.6
Pacific Islander	2,485	1,872	75	613	25	377	74	73	77	70	73	2.3
Filipino	10,371	9,247	89	1,124	11	399	82	83	84	80	81	2.6
Hispanic or Latino	219,781	155,350	71	64,431	29	371	72	71	74	68	69	2.2
African American	28,496	18,671	66	9,825	34	366	70	68	73	64	66	2.2
White (not of Hispanic origin)	97,337	87,228	90	10,109	10	399	86	83	85	80	81	2.5
Two or More Races	6,867	5,409	79	1,458	21	386	79	76	79	73	75	2.4
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	204,712	169,767	83	34,945	17	389	81	78	81	75	76	2.4
Initially Fluent English Proficient (IFEP)	33,045	30,060	91	2,985	9	399	83	83	84	80	81	2.6
Reclassified Fluent English Proficient (RFEP)	81,186	75,515	93	5,671	7	391	80	82	83	79	79	2.4
English-Learner Students	76,526	30,031	39	46,495	61	343	58	57	61	54	54	1.9
Unknown	6,592	3,981	60	2,611	40	365	70	68	71	64	65	2.2
Economically Disadvantaged												
No	138,845	124,622	90	14,223	10	400	85	83	85	80	81	2.6
Yes	229,200	159,082	69	70,118	31	370	71	70	74	67	68	2.2
Unknown	34,016	25,650	75	8,366	25	382	77	75	78	71	73	2.3
Special Education Program Participation												
Students Receiving Services	39,875	12,988	33	26,887	67	337	56	52	59	48	50	1.9
Students Not Receiving Services	362,186	296,366	82	65,820	18	386	79	78	80	74	76	2.4

[1] Results for groups with less 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 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	401,838	310,589	77	91,249	23	382	74	73	71	65	61
Grade											
Tenth	352,493	293,794	83	58,699	17	388	77	75	74	68	64
Eleventh	28,757	10,450	36	18,307	64	343	55	53	49	44	36
Twelfth	17,372	5,041	29	12,331	71	340	53	51	46	42	33
Adult Education	3,216	1,304	41	1,912	59	348	58	55	52	49	34
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	203,032	156,108	77	46,924	23	383	74	73	71	66	60
Female	198,637	154,396	78	44,241	22	381	74	72	71	64	61
Unknown	169	85	50	84	50	357	61	62	57	51	45
Race/Ethnicity											
American Indian or Alaska Native	2,936	2,165	74	771	26	376	72	70	68	62	56
Asian	32,181	30,275	94	1,906	6	416	85	87	86	83	82
Pacific Islander	2,485	1,894	76	591	24	378	72	71	70	63	60
Filipino	10,298	9,263	90	1,035	10	400	80	81	80	75	73
Hispanic or Latino	219,620	156,968	71	62,652	29	373	70	69	67	60	55
African American	29,985	18,279	61	11,706	39	362	66	64	61	53	50
White (not of Hispanic origin)	97,330	86,441	89	10,889	11	397	81	79	79	74	68
Two or More Races	7,003	5,304	76	1,699	24	383	74	73	71	66	60
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	207,750	166,990	80	40,760	20	386	76	74	73	67	62
Initially Fluent English Proficient (IFEP)	33,608	29,779	89	3,829	11	398	81	80	79	73	70
Reclassified Fluent English Proficient (RFEP)	82,949	74,830	90	8,119	10	392	80	78	77	71	68
English-Learner Students	70,556	35,037	50	35,519	50	353	57	58	56	49	45
Unknown	6,975	3,953	57	3,022	43	363	65	63	61	56	46
Economically Disadvantaged											
No	139,491	123,504	89	15,987	11	398	81	80	79	74	70
Yes	227,885	161,716	71	66,169	29	373	70	68	67	60	56
Unknown	34,462	25,369	74	9,093	26	381	73	72	70	64	59
Special Education Program Participation											
Students Receiving Services	36,261	12,356	34	23,905	66	342	52	52	48	42	37
Students Not Receiving Services	365,577	298,233	82	67,344	18	386	76	75	74	67	63

[1] Results for groups with less 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 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		402,061	185,729	46	94,340	23	121,992	30	216,332	54
Grade	Tenth	354,450	141,879	40	91,856	26	120,715	34	212,571	60
	Eleventh	28,021	25,698	92	1,588	6	735	3	2,323	8
	Twelfth	16,669	15,784	95	557	3	328	2	885	5
	Adult Education	2,921	2,368	81	339	12	214	7	553	19
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	207,977	105,232	51	48,376	23	54,369	26	102,745	49
	Female	193,945	80,410	41	45,935	24	67,600	35	113,535	59
	Unknown	139	87	63	29	21	23	17	52	37
Race/Ethnicity	American Indian or Alaskan	2,881	1,386	48	696	24	799	28	1,495	52
	Asian	33,843	8,842	26	6,460	19	18,541	55	25,001	74
	Pacific Islander	2,485	1,246	50	642	26	597	24	1,239	50
	Filipino	10,371	2,829	27	2,510	24	5,032	49	7,542	73
	Hispanic or Latino	219,781	125,718	57	51,828	24	42,235	19	94,063	43
	African American	28,496	17,596	62	6,059	21	4,841	17	10,900	38
	White (not of Hispanic origin)	97,337	25,316	26	24,600	25	47,421	49	72,021	74
	Two or More Races	6,867	2,796	41	1,545	22	2,526	37	4,071	59
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	204,712	75,819	37	50,680	25	78,213	38	128,893	63
	Initially Fluent English Proficient (IFEP)	33,045	9,029	27	8,772	27	15,244	46	24,016	73
	Reclassified Fluent English Proficient (RFEP)	81,186	27,605	34	27,689	34	25,892	32	53,581	66
	English-Learner Students	76,526	69,067	90	6,074	8	1,385	2	7,459	10
	Unknown	6,592	4,209	64	1,125	17	1,258	19	2,383	36
Economically Disadvantaged	No	138,845	36,422	26	34,191	25	68,232	49	102,423	74
	Yes	229,200	133,959	58	52,863	23	42,378	18	95,241	42
	Unknown	34,016	15,348	45	7,286	21	11,382	33	18,668	55
Special Education Program	Receiving Services	39,875	35,116	88	3,280	8	1,479	4	4,759	12
	Not Receiving Services	362,186	150,613	42	91,060	25	120,513	33	211,573	58

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

Table 8.F.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—March 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		401,838	199,035	50	126,839	32	75,964	19	202,803	50
Grade	Tenth	352,493	152,570	43	124,526	35	75,397	21	199,923	57
	Eleventh	28,757	26,947	94	1,445	5	365	1	1,810	6
	Twelfth	17,372	16,569	95	649	4	154	1	803	5
	Adult Education	3,216	2,949	92	219	7	48	1	267	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	203,032	97,574	48	63,836	31	41,622	21	105,458	52
	Female	198,637	101,341	51	62,966	32	34,330	17	97,296	49
	Unknown	169	120	71	37	22	12	7	49	29
Race/Ethnicity	American Indian or Alaskan Native	2,936	1,635	56	899	31	402	14	1,301	44
	Asian	32,181	5,501	17	9,435	29	17,245	54	26,680	83
	Pacific Islander	2,485	1,296	52	855	34	334	13	1,189	48
	Filipino	10,298	3,037	29	4,032	39	3,229	31	7,261	71
	Hispanic or Latino	219,620	131,724	60	64,522	29	23,374	11	87,896	40
	African American	29,985	21,143	71	6,884	23	1,958	7	8,842	29
	White (not of Hispanic origin)	97,330	31,278	32	38,146	39	27,906	29	66,052	68
	Two or More Races	7,003	3,421	49	2,066	30	1,516	22	3,582	51
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	207,750	92,655	45	70,636	34	44,459	21	115,095	55
	Initially Fluent English Proficient (IFEP)	33,608	11,314	34	11,796	35	10,498	31	22,294	66
	Reclassified Fluent English Proficient (RFEP)	82,949	31,848	38	33,159	40	17,942	22	51,101	62
	English-Learner Students	70,556	58,206	82	10,032	14	2,318	3	12,350	18
	Unknown	6,975	5,012	72	1,216	17	747	11	1,963	28
Economically Disadvantaged	No	139,491	44,755	32	51,617	37	43,119	31	94,736	68
	Yes	227,885	136,699	60	65,362	29	25,824	11	91,186	40
	Unknown	34,462	17,581	51	9,860	29	7,021	20	16,881	49
Special Education Program Participation	Receiving Services	36,261	32,102	89	3,335	9	824	2	4,159	11
	Not Receiving Services	365,577	166,933	46	123,504	34	75,140	21	198,644	54

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

Table 8.F.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—March 2011

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	287	309	354	383	410	450	450	381	40	402,061
Grade[4]										
Tenth	293	317	363	389	414	450	450	387	38	354,450
Eleventh	275	293	321	341	358	389	427	340	30	28,021
Twelfth	275	289	315	333	349	381	418	334	29	16,669
Adult Education	287	305	329	347	370	410	439	352	32	2,921
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	284	305	347	378	406	445	450	376	41	207,977
Female	295	317	358	389	414	450	450	386	39	193,945
Unknown	275	284	325	361	392	427	450	360	43	139
Race/Ethnicity										
American Indian or Alaska Native	287	307	352	381	406	445	450	379	40	2,881
Asian	293	321	378	410	439	450	450	402	41	33,843
Pacific Islander	287	311	352	378	402	439	450	377	38	2,485
Filipino	303	331	375	402	427	450	450	399	36	10,371
Hispanic or Latino	287	307	345	373	395	433	450	371	37	219,781
African American	278	299	339	368	392	427	450	366	39	28,496
White (not of Hispanic origin)	295	329	378	402	427	450	450	399	37	97,337
Two or More Races	282	309	358	389	418	450	450	386	43	6,867
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	289	315	363	392	418	450	450	389	40	204,712
Initially Fluent English Proficient (IFEP)	305	337	378	402	427	450	450	399	35	33,045
Reclassified Fluent English Proficient (RFEP)	321	345	373	389	410	445	450	391	29	81,186
English-Learner Students	278	295	323	343	361	389	414	343	29	76,526
Unknown	278	301	335	363	395	439	450	365	41	6,592
Economically Disadvantaged										
No	299	331	378	402	427	450	450	400	37	138,845
Yes	284	305	343	370	395	433	450	370	38	229,200
Unknown	282	305	352	386	414	450	450	382	43	34,016
Special Education Program Participation										
Students Receiving Services	275	287	311	333	358	399	427	337	34	39,875
Students Not Receiving Services	295	321	361	386	414	450	450	386	38	362,186

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	302	320	352	381	411	450	450	382	39	401,838
Grade[4]										
Tenth	307	324	361	386	415	450	450	388	38	352,493
Eleventh	294	307	328	343	356	386	429	343	25	28,757
Twelfth	291	305	324	339	352	379	415	340	23	17,372
Adult Education	305	314	332	345	359	393	429	348	24	3,216
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	300	316	352	384	415	450	450	383	41	203,032
Female	307	322	352	379	408	450	450	381	38	198,637
Unknown	285	307	328	350	384	424	449	357	37	169
Race/Ethnicity										
American Indian or Alaska Native	300	316	348	375	405	441	450	376	38	2,936
Asian	318	345	393	424	450	450	450	416	35	32,181
Pacific Islander	302	320	350	379	405	449	450	378	37	2,485
Filipino	312	336	375	402	429	450	450	400	36	10,298
Hispanic or Latino	302	316	345	371	396	441	450	373	36	219,620
African American	297	312	336	359	386	429	450	362	35	29,985
White (not of Hispanic origin)	309	332	371	399	424	450	450	397	37	97,330
Two or More Races	300	316	350	381	415	450	450	383	42	7,003
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	302	320	357	386	415	450	450	386	39	207,750
Initially Fluent English Proficient (IFEP)	312	334	371	399	429	450	450	398	38	33,608
Reclassified Fluent English Proficient (RFEP)	320	339	369	391	415	450	450	392	34	82,949
English-Learner Students	297	309	332	348	371	411	449	353	31	70,556
Unknown	300	314	336	354	386	449	450	363	38	6,975
Economically Disadvantaged										
No	309	332	371	399	429	450	450	398	38	139,491
Yes	302	316	345	371	399	441	450	373	37	227,885
Unknown	300	316	348	379	415	450	450	381	42	34,462
Special Education Program Participation										
Students Receiving Services	288	302	318	336	359	402	441	342	31	36,261
Students Not Receiving Services	309	326	357	386	415	450	450	386	38	365,577

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 2011

Table 8.G.1: Frequency Distributions, ELA—May 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	229	1	229	99
440-449	130	0	359	99
430-439	304	1	663	98
420-429	435	1	1,098	97
410-419	517	1	1,615	96
400-409	924	2	2,539	94
390-399	973	2	3,512	91
380-389	1,394	4	4,906	87
370-379	1,645	4	6,551	83
360-369	2,793	7	9,344	76
350-359[1]	4,004	10	13,348	66
340-349	4,878	12	18,226	53
330-339	4,948	13	23,174	41
320-329	4,527	12	27,701	29
310-319	4,354	11	32,055	18
300-309	3,059	8	35,114	10
290-299	2,166	6	37,280	5
280-289	1,061	3	38,341	2
275-279	764	2	39,105	0

[1] Passing Score = 350

Table 8.G.2: Frequency Distributions, Mathematics—May 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	237	1	237	99
440-449	135	0	372	99
430-439	268	1	640	98
420-429	302	1	942	98
410-419	312	1	1,254	97
400-409	730	2	1,984	95
390-399	653	2	2,637	93
380-389	1,343	3	3,980	90
370-379	1,762	5	5,742	85
360-369	2,817	7	8,559	78
350-359[1]	5,298	14	13,857	64
340-349	6,895	18	20,752	47
330-339	5,870	15	26,622	31
320-329	5,274	14	31,896	18
310-319	4,210	11	36,106	7
300-309	1,890	5	37,996	2
290-299	619	2	38,615	1
280-289	114	0	38,729	0
270-279	122	0	38,851	0

[1] Passing Score = 350

Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	229	1	229	99
440-449	130	0	359	99
430-439	304	1	663	98
420-429	435	1	1,098	97
410-419	517	1	1,615	96
403-409	593	2	2,208	94
390-402	1,304	3	3,512	91
380-389	1,394	4	4,906	87
370-379	1,645	4	6,551	83
360-369	2,793	7	9,344	76
350-359	4,004	10	13,348	66
340-349	4,878	12	18,226	53
330-339	4,948	13	23,174	41
320-329	4,527	12	27,701	29
310-319	4,354	11	32,055	18
300-309	3,059	8	35,114	10
290-299	2,166	6	37,280	5
280-289	1,061	3	38,341	2
270-279	764	2	39,105	0

[1] Advanced-Level Cut = 403

[2] Proficient-Level Cut = 380

Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2011

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	237	1	237	99
440-449	135	0	372	99
430-439	268	1	640	98
422-429	302	1	942	98
410-421	312	1	1,254	97
400-409	730	2	1,984	95
390-399	653	2	2,637	93
380-389	1,343	3	3,980	90
370-379	1,762	5	5,742	85
360-369	2,817	7	8,559	78
350-359	5,298	14	13,857	64
340-349	6,895	18	20,752	47
330-339	5,870	15	26,622	31
320-329	5,274	14	31,896	18
310-319	4,210	11	36,106	7
300-309	1,890	5	37,996	2
290-299	619	2	38,615	1
280-289	114	0	38,729	0
270-279	122	0	38,851	0

[1] Advanced-Level Cut = 422

[2] Proficient-Level Cut = 380

Table 8.G.5: Demographic Summary for All Examinees, ELA—May 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	39,105	13,348	34	25,757	66	340	55	58	56	48	53	2.1
Grade												
Tenth	9,385	5,543	59	3,842	41	359	66	67	67	59	62	2.3
Eleventh	13,360	4,211	32	9,149	68	337	54	57	55	46	51	2.0
Twelfth	13,742	2,534	18	11,208	82	329	49	52	49	42	47	2.0
Adult Education	2,618	1,060	40	1,558	60	348	61	65	60	55	55	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	22,436	7,055	31	15,381	69	336	54	57	54	46	51	2.0
Female	16,585	6,263	38	10,322	62	345	57	60	59	52	55	2.2
Unknown	84	30	36	54	64	339	56	59	55	48	55	2.0
Race/Ethnicity												
American Indian or Alaska Native	340	132	39	208	61	344	59	62	60	49	54	2.1
Asian	1,885	589	31	1,296	69	340	51	58	54	51	56	2.1
Pacific Islander	244	82	34	162	66	342	55	58	57	48	55	2.2
Filipino	499	194	39	305	61	349	58	63	59	55	58	2.3
Hispanic or Latino	24,297	7,059	29	17,238	71	336	53	57	54	46	50	2.0
African American	3,736	1,145	31	2,591	69	334	52	55	54	44	49	2.0
White (not of Hispanic origin)	6,904	3,611	52	3,293	48	356	64	66	64	57	61	2.2
Two or More Races	1,200	536	45	664	55	348	60	62	61	53	57	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Language Fluency												
English-Only Students	16,629	7,184	43	9,445	57	346	59	61	60	51	56	2.2
Initially Fluent English Proficient (IFEP)	1,362	663	49	699	51	352	63	65	63	54	58	2.3
Reclassified Fluent English Proficient (RFEP)	2,137	1,258	59	879	41	357	65	68	66	59	62	2.3
English-Learner Students	15,893	2,937	18	12,956	82	328	48	52	49	42	47	1.9
Unknown	3,084	1,306	42	1,778	58	348	61	65	61	55	55	2.2
Economically Disadvantaged												
No	7,868	3,941	50	3,927	50	355	63	65	64	56	60	2.3
Yes	25,356	7,025	28	18,331	72	334	52	55	53	45	50	2.0
Unknown	5,881	2,382	41	3,499	59	345	59	62	59	52	54	2.1
Special Education Program Participation												
Students Receiving Services	8,100	1,166	14	6,934	86	320	44	46	45	36	42	1.8
Students Not Receiving Services	31,005	12,182	39	18,823	61	345	58	62	59	52	55	2.2

[1] Results for groups with less 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 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	38,851	13,857	36	24,994	64	344	53	50	51	45	39
Grade											
Tenth	9,233	5,162	56	4,071	44	359	61	59	60	53	47
Eleventh	13,251	4,554	34	8,697	66	342	52	48	50	43	38
Twelfth	13,400	3,042	23	10,358	77	336	49	45	45	40	34
Adult Education	2,967	1,099	37	1,868	63	346	57	49	52	49	38
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	19,403	6,423	33	12,980	67	342	52	49	49	44	37
Female	19,374	7,410	38	11,964	62	346	55	50	52	46	41
Unknown	74	24	32	50	68	344	53	51	50	45	38
Race/Ethnicity											
American Indian or Alaska Native	378	131	35	247	65	344	55	51	51	44	37
Asian	985	523	53	462	47	361	58	59	60	56	52
Pacific Islander	233	93	40	140	60	344	53	50	51	45	40
Filipino	445	212	48	233	52	354	56	55	58	52	45
Hispanic or Latino	23,598	7,447	32	16,151	68	341	52	47	49	43	37
African American	4,793	1,308	27	3,485	73	338	50	46	46	41	35
White (not of Hispanic origin)	7,126	3,622	51	3,504	49	356	60	58	58	52	44
Two or More Races	1,293	521	40	772	60	348	56	52	53	48	40
Unknown	-	-	-	-	-	-	-	-	-	-	-
Language Fluency											
English-Only Students	19,140	7,393	39	11,747	61	346	55	52	52	46	39
Initially Fluent English Proficient (IFEP)	1,633	725	44	908	56	350	58	54	55	47	41
Reclassified Fluent English Proficient (RFEP)	2,968	1,521	51	1,447	49	354	61	55	58	50	44
English-Learner Students	11,775	2,873	24	8,902	76	337	47	44	46	41	37
Unknown	3,335	1,345	40	1,990	60	348	57	51	53	49	39
Economically Disadvantaged											
No	8,598	4,143	48	4,455	52	354	59	56	57	51	43
Yes	23,846	7,255	30	16,591	70	340	51	47	48	42	37
Unknown	6,407	2,459	38	3,948	62	346	55	51	52	47	39
Special Education Program Participation											
Students Receiving Services	6,854	995	15	5,859	85	328	41	42	39	35	31
Students Not Receiving Services	31,997	12,862	40	19,135	60	348	56	51	53	47	40

[1] Results for groups with less 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 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		39,105	34,199	87	2,698	7	2,208	6	4,906	13
Grade	Tenth	9,385	6,263	67	1,616	17	1,506	16	3,122	33
	Eleventh	13,360	12,402	93	567	4	391	3	958	7
	Twelfth	13,742	13,310	97	285	2	147	1	432	3
	Adult Education	2,618	2,224	85	230	9	164	6	394	15
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	22,436	20,066	89	1,359	6	1,011	5	2,370	11
	Female	16,585	14,057	85	1,334	8	1,194	7	2,528	15
	Unknown	84	76	90	5	6	3	4	8	10
Race/Ethnicity	American Indian or Alaskan Native	340	289	85	23	7	28	8	51	15
	Asian	1,885	1,683	89	95	5	107	6	202	11
	Pacific Islander	244	209	86	26	11	9	4	35	14
	Filipino	499	412	83	42	8	45	9	87	17
	Hispanic or Latino	24,297	22,405	92	1,273	5	619	3	1,892	8
	African American	3,736	3,425	92	201	5	110	3	311	8
	White (not of Hispanic origin)	6,904	4,835	70	893	13	1,176	17	2,069	30
	Two or More Races	1,200	941	78	145	12	114	10	259	22
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	16,629	13,343	80	1,654	10	1,632	10	3,286	20
	Initially Fluent English Proficient (IFEP)	1,362	1,054	77	180	13	128	9	308	23
	Reclassified Fluent English Proficient (RFEP)	2,137	1,625	76	346	16	166	8	512	24
	English-Learner Students	15,893	15,652	98	201	1	40	0	241	2
	Unknown	3,084	2,525	82	317	10	242	8	559	18
Economically Disadvantaged	No	7,868	5,736	73	956	12	1,176	15	2,132	27
	Yes	25,356	23,556	93	1,206	5	594	2	1,800	7
	Unknown	5,881	4,907	83	536	9	438	7	974	17
Special Education Program	Receiving Services	8,100	7,902	98	140	2	58	1	198	2
	Not Receiving Services	31,005	26,297	85	2,558	8	2,150	7	4,708	15

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

Table 8.G.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— May 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		38,851	34,871	90	3,038	8	942	2	3,980	10
Grade	Tenth	9,233	6,631	72	1,875	20	727	8	2,602	28
	Eleventh	13,251	12,497	94	604	5	150	1	754	6
	Twelfth	13,400	13,010	97	349	3	41	0	390	3
	Adult Education	2,967	2,733	92	210	7	24	1	234	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	19,403	17,314	89	1,563	8	526	3	2,089	11
	Female	19,374	17,493	90	1,465	8	416	2	1,881	10
	Unknown	74	64	86	10	14	0	0	10	14
Race/Ethnicity	American Indian or Alaskan Native	378	343	91	28	7	7	2	35	9
	Asian	985	723	73	138	14	124	13	262	27
	Pacific Islander	233	212	91	17	7	4	2	21	9
	Filipino	445	359	81	61	14	25	6	86	19
	Hispanic or Latino	23,598	22,085	94	1,295	5	218	1	1,513	6
	African American	4,793	4,556	95	213	4	24	1	237	5
	White (not of Hispanic origin)	7,126	5,481	77	1,156	16	489	7	1,645	23
	Two or More Races	1,293	1,112	86	130	10	51	4	181	14
	Unknown	-	-	-	-	-	-	-	-	-
Language Fluency	English-Only Students	19,140	16,601	87	1,892	10	647	3	2,539	13
	Initially Fluent English Proficient (IFEP)	1,633	1,387	85	186	11	60	4	246	15
	Reclassified Fluent English Proficient (RFEP)	2,968	2,509	85	365	12	94	3	459	15
	English-Learner Students	11,775	11,386	97	320	3	69	1	389	3
	Unknown	3,335	2,988	90	275	8	72	2	347	10
Economically Disadvantaged	No	8,598	6,837	80	1,213	14	548	6	1,761	20
	Yes	23,846	22,335	94	1,290	5	221	1	1,511	6
	Unknown	6,407	5,699	89	535	8	173	3	708	11
Special Education Program	Receiving Services	6,854	6,686	98	149	2	19	0	168	2
	Not Receiving Services	31,997	28,185	88	2,889	9	923	3	3,812	12

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

Table 8.G.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—May 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	291	316	336	358	405	439	340	34	39,105
Grade[4]										
Tenth	275	291	327	360	392	433	450	359	43	9,385
Eleventh	275	291	318	336	354	389	427	337	30	13,360
Twelfth	275	289	310	329	344	371	405	329	26	13,742
Adult Education	291	306	327	342	364	409	433	348	31	2,618
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	287	312	334	354	401	433	336	34	22,436
Female	280	297	321	340	362	413	446	345	34	16,585
Unknown	275	289	320	340	355	392	433	339	31	84
Race/Ethnicity										
American Indian or Alaska Native	275	289	321	338	364	415	446	344	36	340
Asian	275	293	318	336	354	409	450	340	34	1885
Pacific Islander	278	297	316	338	360	401	439	342	33	244
Filipino	287	300	327	342	364	417	450	349	34	499
Hispanic or Latino	275	291	316	334	352	389	422	336	30	24,297
African American	275	283	310	332	354	392	417	334	32	3,736
White (not of Hispanic origin)	275	291	323	352	389	433	450	356	43	6,904
Two or More Races	275	289	321	344	373	417	446	348	38	1,200
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	275	289	318	342	371	422	450	346	39	16,629
Initially Fluent English Proficient (IFEP)	283	299	325	348	376	417	446	352	36	1,362
Reclassified Fluent English Proficient (RFEP)	278	302	336	356	378	413	439	357	33	2,137
English-Learner Students	275	291	312	329	344	366	386	328	23	15,893
Unknown	278	300	325	344	369	413	439	348	34	3,084
Economically Disadvantaged										
No	275	295	325	350	383	433	450	355	41	7,868
Yes	275	289	314	332	352	389	417	334	30	25,356
Unknown	275	293	321	342	366	413	446	345	36	5,881
Special Education Program Participation										
Students Receiving Services	275	280	300	318	338	366	397	320	27	8,100
Students Not Receiving Services	276	297	323	340	362	413	446	345	34	31,005

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 2011

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	293	306	325	341	357	400	438	344	28	38,851
Grade[4]										
Tenth	293	306	329	355	382	432	450	359	37	9,233
Eleventh	293	306	325	341	355	382	422	342	24	13,251
Twelfth	293	306	321	336	348	368	397	336	21	13,400
Adult Education	306	317	333	343	355	389	417	346	22	2,967
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	290	304	321	338	355	403	446	342	30	19,403
Female	299	311	329	343	357	397	438	346	26	19,374
Unknown	276	306	327	341	357	389	406	344	25	74
Race/Ethnicity										
American Indian or Alaska Native	293	306	327	341	355	397	438	344	27	378
Asian	296	311	334	350	385	446	450	361	40	985
Pacific Islander	290	306	327	343	357	389	432	344	26	233
Filipino	299	313	333	347	368	422	450	354	32	445
Hispanic or Latino	293	306	325	340	354	385	417	341	24	23,598
African American	284	304	321	336	350	378	406	338	23	4,793
White (not of Hispanic origin)	296	309	331	350	376	426	450	356	36	7,126
Two or More Races	293	309	327	343	363	413	450	348	31	1,293
Unknown	-	-	-	-	-	-	-	-	-	-
Language Fluency										
English-Only Students	293	306	325	341	361	409	446	346	31	19,140
Initially Fluent English Proficient (IFEP)	299	309	329	345	364	409	450	350	31	1,633
Reclassified Fluent English Proficient (RFEP)	301	315	336	350	366	409	446	354	27	2,968
English-Learner Students	293	306	321	336	348	372	406	337	22	11,775
Unknown	299	313	331	345	359	397	432	348	26	3,335
Economically Disadvantaged										
No	296	309	331	348	372	426	450	354	34	8,598
Yes	293	306	323	338	354	385	417	340	24	23,846
Unknown	293	309	329	343	359	403	446	346	28	6,407
Special Education Program Participation										
Students Receiving Services	280	299	313	325	340	366	394	328	22	6,854
Students Not Receiving Services	299	311	329	343	361	403	446	348	28	31,997

[1] Mean Scale Scores are reported at each percentile.

[2] SD — Standard Deviation

[3] Results for groups with less 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 that 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 their products. ETS established an Office of Quality Assurance that 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 copy editing 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 the 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 very important. 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 of the batches 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 ones 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 are 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 it 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: 2008–09, 2009–10, and 2010-11. The indicators of examinee performance include the mean and standard deviation of scale scores, the percentage of examinees classified into the passing-, 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

Given in Table 10.A.1 for the ELA and mathematics tests are the number of examinees assessed and the means and standard deviations of examinees' scale scores in 2008–09, 2009–10, and 2010-11.

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 percentage 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. 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 the average proportion correct, point-biserial correlations and reliabilities are affected by both item characteristics and student characteristics.

[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).

Appendix 10.A—Historical Comparisons on Student Performance

Table 10.A.1: Number of Examinees Tested, Scale Score Means, and Standard Deviations of CAHSEE across 2008–09, 2009–10 and 2010–11

Subject	Admin	Number of Students (with valid scores)			Scale Score Mean and Standard Deviation (SD)					
					2008–09		2009–10		2010–11	
		2008–09	2009–10	2010–11	Mean	SD	Mean	SD	Mean	SD
English–Language Arts	July	13,373	12,912	8,389	337	26	335	25	337	25
	October	43,955	41,570	43,194	342	29	340	31	342	29
	November	108,722	107,342	102,616	343	33	343	30	345	30
	December	3,544	2,754	2,405	338	29	338	28	342	25
	February	195,091	171,323	154,910	371	41	372	41	370	41
	March	372,707	393,121	402,061	374	39	376	40	381	40
	May	42,781	40,769	39,105	340	33	341	33	340	34
Mathematics	July	13,237	12,388	8,222	340	22	343	21	338	22
	October	44,958	40,423	41,519	343	26	344	26	342	26
	November	109,445	103,221	98,156	345	27	348	27	344	28
	December	3,544	2,556	2,433	345	22	342	21	340	23
	February	196,387	168,363	155,492	373	40	373	40	373	40
	March	371,704	388,590	401,838	378	39	381	40	382	39
	May	42,546	38,939	38,851	342	28	341	28	344	28

Table 10.A.2: Percentage of Students Passing Each Content Area across 2008–09, 2009–10 and 2010–11

Admin	Percent Passing					
	English–Language Arts			Mathematics		
	2008–09	2009–10	2010–11	2008–09	2009–10	2010–11
July	30	24	27	29	33	26
October	37	34	37	33	34	31
November	40	39	40	38	42	35
December	30	30	36	36	30	29
February	67	69	67	67	67	68
March	73	74	77	74	75	77
May	33	35	34	31	32	36

Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced across 2008–09, 2009–10 and 2010–11

Subject	Admin	Below Proficient			Proficient			Advanced		
		2008–09	2009–10	2010–11	2008–09	2009–10	2010–11	2008–09	2009–10	2010–11
English–Language Arts	July	95	95	95	3	3	4	2	2	1
	October	91	91	92	5	5	5	4	4	3
	November	88	91	90	7	6	6	5	3	4
	December	93	92	94	4	5	4	3	3	2
	February	58	56	58	17	20	19	25	24	23
	March	55	53	47	22	20	23	23	27	30
	May	89	88	88	6	6	7	5	6	5
Mathematics	July	96	96	96	4	3	3	1	1	1
	October	93	92	93	5	6	5	2	2	2
	November	91	90	91	7	7	7	2	3	2
	December	94	95	95	5	4	4	1	1	1
	February	59	58	59	26	28	26	15	14	15
	March	53	52	50	30	30	31	17	18	19
	May	90	91	90	8	7	8	2	2	2

Table 10.A.4: Scale Score Distributions across 2008, 2009 and 2010 for ELA (July to December)

Scale Score Distribution	July			October			November			December		
	2008	2009	2010	2008	2009	2010	2008	2009	2010	2008	2009	2010
450	15	15	9	16	237	174	637	327	400	8	12	3
440 – 449	10	20	4	180	229	85	695	186	238	9	3	3
430 – 439	35	12	15	109	122	234	440	526	599	18	16	5
420 – 429	47	35	24	300	320	288	1,002	649	840	20	14	11
410 – 419	52	53	38	571	537	498	1,954	1,370	1,003	26	15	19
400 – 409	96	102	59	420	380	415	1,577	1,100	1,828	32	28	14
390 – 399	198	134	122	1,070	744	1,004	2,744	2,722	2,163	55	48	38
380 – 389	214	228	132	1,048	1,312	1,038	3,614	2,867	3,885	90	75	41
370 – 379	470	383	250	2,776	1,999	2,040	6,817	7,312	5,986	158	100	86
360 – 369	1,055	631	582	3,632	3,043	4,188	9,430	9,475	11,571	241	167	226
350 – 359	1,761	1,482	1,055	6,102	5,398	6,089	14,656	15,412	12,206	416	361	417
340 – 349	2,218	2,164	1,486	6,733	6,092	7,004	15,210	16,627	18,756	552	403	462
330 – 339	2,241	2,298	1,519	6,096	5,839	6,284	13,541	14,403	13,294	559	467	427
320 – 329	1,769	1,935	1,205	5,684	4,923	4,789	10,920	11,190	10,114	482	364	283
310 – 319	1,344	1,612	849	3,447	3,975	4,187	8,496	9,893	7,822	336	279	155
300 – 309	1,074	904	491	2,866	3,403	2,253	6,813	6,120	5,420	247	219	83
290 – 299	451	519	306	1,660	1,577	1,395	5,682	4,024	4,042	164	99	80
280 – 289	169	226	140	699	863	723	2,583	1,702	1,262	78	51	30
275 – 279	154	159	103	401	577	506	1,911	1,437	1,187	53	33	22

Table 10.A.5: Scale Score Distributions across 2009, 2010 and 2011 for ELA (February to May)

Scale Score Distribution	February			March			May		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
450	7,532	6,750	5,534	13,087	16,863	21,724	254	194	229
440 – 449	3,505	2,735	2,469	5,998	7,069	8,297	117	256	130
430 – 439	8,335	6,873	5,920	15,260	17,570	19,558	306	170	304
420 – 429	9,677	7,819	7,120	9,057	20,527	22,657	424	434	435
410 – 419	10,095	8,523	7,744	29,937	22,493	36,871	545	719	517
400 – 409	15,349	13,006	12,001	21,199	34,700	25,808	911	819	924
390 – 399	14,517	12,874	11,726	31,916	33,524	37,067	1,014	881	973
380 – 389	13,347	16,087	14,183	40,500	31,876	44,350	1,240	1,445	1,394
370 – 379	15,879	14,445	12,369	36,498	37,678	37,110	1,957	1,841	1,645
360 – 369	17,951	13,545	11,337	32,263	33,007	30,510	3,503	3,214	2,793
350 – 359	13,654	16,190	13,394	34,885	35,279	25,402	3,778	4,464	4,004
340 – 349	16,214	14,355	12,899	29,126	29,868	25,848	5,721	5,224	4,878
330 – 339	14,670	11,657	10,856	23,128	23,624	19,709	6,965	6,093	4,948
320 – 329	11,753	8,858	10,169	16,789	17,163	15,066	4,889	4,216	4,527
310 – 319	10,438	6,798	6,212	14,345	12,151	11,839	3,985	3,734	4,354
300 – 309	5,830	4,899	4,876	8,379	8,616	9,257	3,008	3,274	3,059
290 – 299	3,444	3,197	3,279	5,609	6,563	6,094	2,169	2,085	2,166
280 – 289	1,673	1,662	1,689	2,813	2,675	3,102	1,161	973	1,001
275 – 279	1,228	1,050	1,133	1,918	1,875	1,792	834	733	764

Table 10.A.6: Scale Score Distributions across 2008, 2009 and 2010 for Mathematics (July to December)

Scale Score Distribution	July			October			November			December		
	2008	2009	2010	2008	2009	2010	2008	2009	2010	2008	2009	2010
450	28	31	16	231	218	204	680	940	806	14	6	4
440 – 449	14	6	9	166	95	91	319	349	285	5	1	0
430 – 439	22	14	17	95	239	234	681	750	687	13	5	7
420 – 429	31	24	15	240	228	171	807	778	744	8	11	11
410 – 419	46	63	19	391	382	419	1,285	1,217	797	17	18	6
400 – 409	81	74	48	454	470	501	1,392	1,331	1,322	42	14	20
390 – 399	121	127	88	477	534	551	2,111	2,129	2,062	41	24	32
380 – 389	216	182	119	1,133	1,101	1,156	2,832	2,765	2,582	61	57	48
370 – 379	443	448	219	1,503	1,902	1,338	5,416	5,552	4,795	129	61	91
360 – 369	834	1,195	435	4,284	3,251	3,443	9,208	9,859	8,047	379	205	144
350 – 359	1,978	1,909	1,145	5,973	5,398	4,907	16,540	17,607	11,926	550	368	339
340 – 349	2,746	2,972	1,405	9,158	8,202	7,362	16,689	19,955	17,861	831	565	486
330 – 339	2,366	2,184	1,969	7,384	6,700	7,704	20,525	15,201	17,677	675	515	503
320 – 329	1,999	1,712	1,305	6,136	6,291	5,783	14,690	11,989	12,603	458	448	344
310 – 319	1,419	1,047	923	4,544	2,930	4,607	10,617	8,456	9,548	242	143	236
300 – 309	614	297	340	1,921	1,743	2,096	4,208	3,080	4,332	62	95	117
290 – 299	225	65	109	589	562	721	965	891	1,382	11	13	35
280 – 289	32	15	26	159	99	123	282	129	418	5	4	8
275 – 279	22	23	15	120	78	108	198	243	282	1	3	2

Table 10.A.7: Scale Score Distributions across 2009, 2010 and 2011 for Mathematics (February to May)

Scale Score Distribution	February			March			May		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
450	9,981	9,423	8,060	25,088	33,829	26,615	292	201	237
440 – 449	3,833	3,390	3,105	15,389	9,242	19,698	142	106	135
430 – 439	8,066	7,176	6,496	7,868	9,432	9,917	308	215	268
420 – 429	7,961	7,007	6,521	15,392	27,208	19,734	294	240	302
410 – 419	11,820	7,140	9,485	24,046	17,686	29,187	392	303	312
400 – 409	11,493	10,308	9,297	24,184	25,766	28,409	594	533	730
390 – 399	10,994	10,034	11,534	24,349	32,878	35,906	975	820	653
380 – 389	17,347	15,728	10,778	38,479	31,367	33,337	1,096	964	1,343
370 – 379	13,280	11,891	12,497	28,657	36,091	37,533	1,745	1,616	1,762
360 – 369	19,254	14,075	12,138	39,875	39,142	33,684	2,488	2,436	2,817
350 – 359	16,980	16,876	15,613	31,088	29,700	36,569	4,907	5,094	5,298
340 – 349	21,278	17,137	13,576	33,834	32,305	27,102	7,358	5,774	6,895
330 – 339	16,246	12,968	14,497	23,294	22,761	26,739	7,070	7,591	5,870
320 – 329	13,318	11,034	9,904	18,599	18,806	17,556	6,705	5,414	5,274
310 – 319	9,284	8,442	7,417	13,416	14,291	10,417	5,129	4,456	4,210
300 – 309	3,716	3,872	3,237	5,714	5,879	7,323	2,371	2,110	1,890
290 – 299	1,155	1,369	835	1,834	1,458	1,379	425	768	619
280 – 289	186	297	265	324	434	412	127	146	114
275 – 279	195	196	237	274	315	321	128	152	122

Appendix 10.B—Historical Comparisons on Test Characteristics

Table 10.B.1: Average Proportion Correct of Operational Test Items across 2008–09, 2009–10 and 2010–11

Subject	Admin	Average *p*-value		
		2008–09	2009–10	2010–11
English–Language Arts	July	0.53	0.53	0.54
	October	0.56	0.56	0.56
	November	0.59	0.56	0.58
	December	0.56	0.57	0.57
	February	0.71	0.71	0.70
	March	0.73	0.74	0.75
	May	0.56	0.53	0.54
Mathematics	July	0.46	0.45	0.45
	October	0.48	0.48	0.47
	November	0.48	0.50	0.49
	December	0.47	0.48	0.47
	February	0.64	0.65	0.64
	March	0.67	0.67	0.69
	May	0.47	0.47	0.47

Table 10.B.2: Average IRT b-values of Operational Test Items across 2008–09, 2009–10 and 2010–11

Subject	Admin	Average IRT b-value		
		2008–09	2009–10	2010–11
English–Language Arts	July	-0.01	-0.09	-0.07
	October	-0.05	-0.10	-0.05
	November	-0.19	-0.03	-0.06
	December	-0.18	-0.17	-0.07
	February	-0.08	-0.06	-0.04
	March	-0.09	-0.12	-0.06
	May	-0.12	0.05	-0.04
Mathematics	July	-0.23	-0.10	-0.24
	October	-0.23	-0.21	-0.25
	November	-0.19	-0.17	-0.26
	December	-0.11	-0.27	-0.25
	February	-0.22	-0.28	-0.18
	March	-0.22	-0.16	-0.18
	May	-0.23	-0.24	-0.18

Table 10.B.3 Average Point-Biserial Correlation of Operational Test Items across 2008–09, 2009–10 and 2010–11

Subject	Admin	Average Point-Biserial Correlation		
		2008–09	2009–10	2010–11
English–Language Arts	July	0.32	0.31	0.31
	October	0.35	0.36	0.35
	November	0.38	0.36	0.36
	December	0.35	0.34	0.31
	February	0.45	0.45	0.45
	March	0.42	0.43	0.44
	May	0.38	0.39	0.40
Mathematics	July	0.29	0.28	0.28
	October	0.32	0.33	0.33
	November	0.34	0.34	0.35
	December	0.28	0.28	0.30
	February	0.46	0.46	0.47
	March	0.46	0.46	0.46
	May	0.35	0.34	0.35

Table 10.B.4 Reliabilities and Standard Errors of Measurement (SEMs) of Operational Test Forms across 2008–09, 2009–10 and 2010–11

Subject	Admin	Reliability			SEM		
		2008–09	2009–10	2010–11	2008–09	2009–10	2010–11
English–Language Arts	July	0.87	0.87	0.87	4.40	4.38	4.37
	October	0.89	0.91	0.89	4.63	4.35	4.43
	November	0.87	0.90	0.90	4.40	4.38	4.34
	December	0.89	0.89	0.86	4.33	4.42	4.32
	February	0.94	0.92	0.94	3.99	4.38	4.10
	March	0.93	0.93	0.94	3.98	3.91	3.81
	May	0.91	0.92	0.92	4.38	4.49	4.46
Mathematics	July	0.87	0.85	0.86	4.12	4.11	4.12
	October	0.89	0.89	0.90	4.04	4.07	4.10
	November	0.87	0.90	0.91	4.12	4.08	4.06
	December	0.85	0.85	0.87	4.13	4.15	4.13
	February	0.95	0.95	0.95	3.71	3.68	3.71
	March	0.95	0.95	0.95	3.63	3.61	3.59
	May	0.91	0.91	0.91	4.08	4.06	4.08

Chapter 11: Alternative Means Pilot Study

Introduction

A pilot study was initiated at the request of the CDE on a proposed alternative means to the CAHSEE for eligible SWDs. The purpose of the pilot study was to investigate and report on alternative means by which eligible SWDs may demonstrate comparable levels of academic achievement in the content standards in ELA and mathematics required for passing the CAHSEE. This chapter summarizes the conceptual framework for the pilot study, methods, results, conclusions, and recommendations to the CDE and the SBE. The full report (CDE, 2011a) may be downloaded from the CDE Independent Evaluations Web page at http://www.cde.ca.gov/ta/tg/hs/evaluations.asp.

CAHSEE Alternative Means Background

The background on the development of CAHSEE alternative means is described in this section. The current conception of the two-tiered process and research results from previous studies are discussed.

Assembly Bill 2040

In May 2009, AB 2040 enacted a statute that established California *EC* sections 60852.1 and 60852.2, requiring a panel of educators and others with experience working with SWDs and/or assessment to make recommendations regarding alternative means to the CAHSEE. These alternative means would provide eligible SWDs an opportunity to demonstrate the same level of academic achievement in the ELA and mathematics content standards required for passing the CAHSEE. In November 2009, the AB 2040 Panel's findings and recommendations for alternative means to the CAHSEE were presented to the SSPI and the SBE.

Based on research, data analysis, and panel discussions, the AB 2040 Panel recommended a two-tiered CAHSEE Performance Validation Process (PVP) in lieu of a new assessment. The CAHSEE PVP Tier I would require validation of student performance through scores on other assessments. If a student is unable to earn the required score in Tier I, the student would move on to Tier II. Tier II would require validation of student performance through work samples and collection of other evidence.

Tier I and Tier II

The recommended two-tiered alternative means process is represented by the flowchart in Figure 11.1 from the CDE. The process begins with the identification of eligible SWDs. According to statute (*EC* Section 60852.2[a]), in order to be eligible for alternative means the student has to have:

- An operative IEP or Section 504 plan.

- Not passed the high school exit examination, even with accommodations and/or modifications.

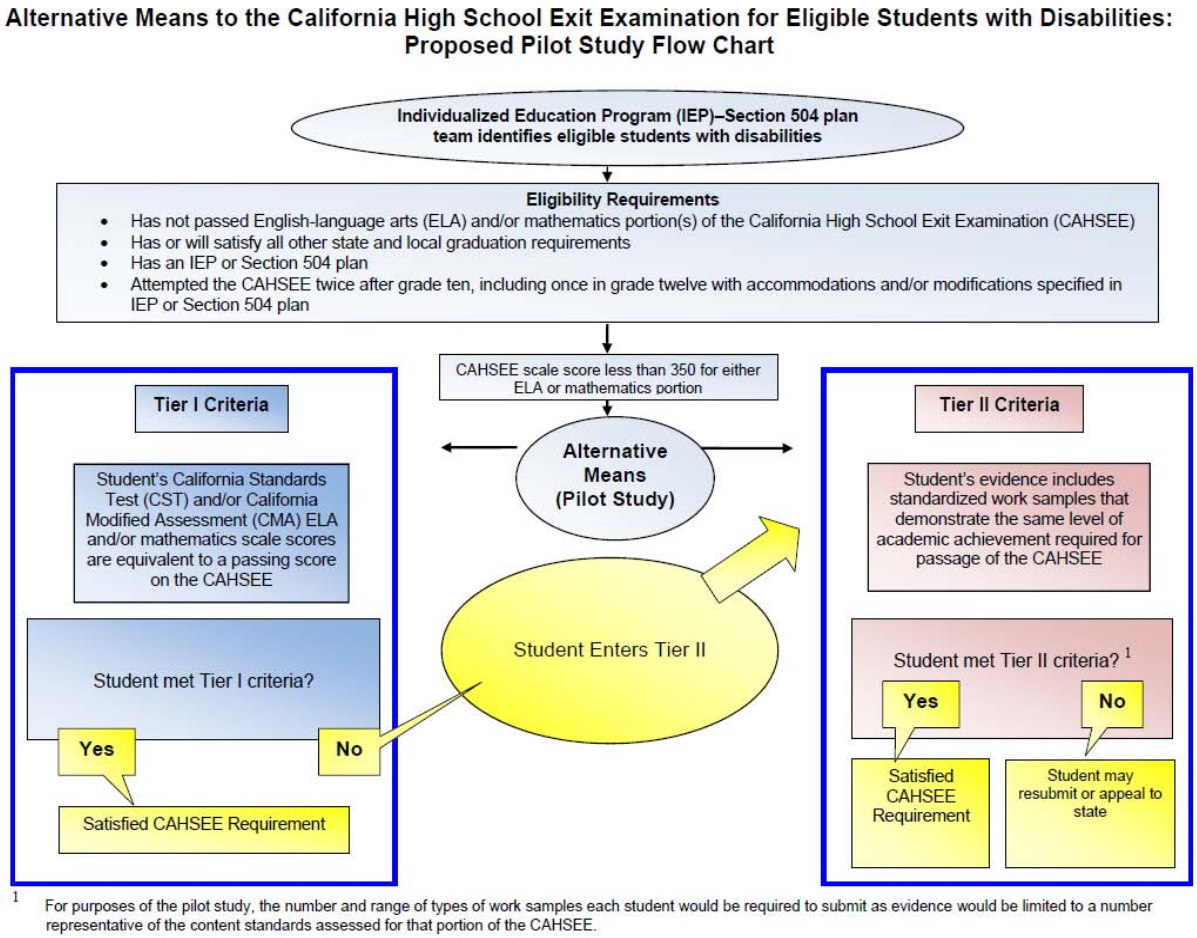
- Satisfied or will satisfy all other state and local graduation requirements.

- Attempted to pass those sections not yet passed of the high school exit examination at least twice after grade ten, including at least one attempt in grade twelve.

Students who meet these requirements will participate in the Tier I process. Tier I consists of a state-level screening of eligible SWDs who have achieved a scale score of 300 on the STAR Program's CST for ELA grade ten and/or the CST for Algebra I.

Eligible SWDs who do not achieve a minimum score in the Tier I screening would then participate in the Tier II process. Tier II requires that eligible SWDs submit a collection of evidence (COE) that would demonstrate the same level of achievement in the ELA and mathematics content standards required for passing the CAHSEE. The pilot study described here focused on the operationalization of the COE concept for Tier II.

Figure 11.1 Flowchart representing the Proposed Two-Tiered CAHSEE Alternative Means Process



Studies Related to Alternative Means

At the request of the CDE, two studies were conducted on CAHSEE alternative means that are directly relevant to the pilot investigation. The first study was undertaken by the AIR (AIR, 2010) to investigate a subgroup of high school students who had taken the CAHSEE with modifications and/or accommodations, who had not passed the CAHSEE, but who would satisfy all other requirements for graduation.

AIR conducted in-depth individual assessments based on adaptations of released CAHSEE MC items. The goals were to enhance accessibility and employ one-on-one probes for each item built on cognitive interviewing techniques. The students' responses to the scripted probes provided additional evidence about their mastery of CAHSEE content and were used to adjudicate responses that were not clearly accurate. Responses could receive a partial score if they indicated partial mastery. The study found that the item and form adaptations are feasible for the CAHSEE since current items and proficiency standards can be used. Item and form adaptations would need to be developed and field-tested.

A summary of approaches taken by states other than California that provide alternative means for students to satisfy their high school graduation requirements

was incorporated into the study. These approaches included task-specific assessment, concordant scores, and the COE currently under consideration. AIR recommended the individual assessment strategy and urged caution in adopting alternative means that include a COE due to concerns over verifying the independence of student work. AIR noted that many students simply did not have mastery of the content assessed by the CAHSEE. The report may be downloaded from the CDE Independent Evaluations Web page at http://www.cde.ca.gov/ta/tg/hs/evaluations.asp.

The second study was conducted by HumRRO (Hardoin, Wise & Buckland, 2010) and submitted as part of the independent evaluation of the CAHSEE program. The goals of the study were to collect information about (a) the feasibility of the proposed Tier I alternative means, and (b) how the level of academic achievement demonstrated by the alternative means compares to the level of academic achievement required for passing the CAHSEE.

Using available test score data on SWDs, approximately 19,000 students may be eligible for Tier I screening annually. Compared to all SWDs, Tier I eligible students were somewhat more likely to be Hispanic or African American and considerably more likely to be EL or students from socio-economically disadvantaged backgrounds. Nearly 70 percent of the eligible students were classified as having a specific learning disability.

Their approach to Tier I screening examined an average of scores from several CST across grade levels and subject areas which, taken together, covered the content included in the CAHSEE. The overall Tier I passing rate using this approach was found to be 0.2 percent, or less than 100 students per year. Although the impact is a relatively small number of students, the authors concluded that the Tier I screen would be a feasible process and recommended that the process be automated and performed by the CDE.

For Tier II, HumRRO collected feedback from LEA personnel on options for eligibility, administration, type and amount of evidence, and scoring.

Eligibility—Over 75 percent of respondents agreed that it was feasible to identify students eligible for the CAHSEE PVP by the start of the second semester of their senior year.

Administration—Forty-nine percent disagreed or strongly disagreed that the responsibilities for collecting and reviewing evidence could be implemented “fairly easily,” expressing concern about the amount of time required for PVP training and evidence collection.

Type and Amount of Evidence—More than half of the respondents agreed or strongly agreed that work samples should be focused on individual standards rather than at the strand level and about two-thirds agreed or strongly agreed that work samples would allow students to demonstrate the same level of academic achievement that the CAHSEE requires. More than three-fourths of the respondents chose a “streamlined option” that would include 50 to 75 percent as many work samples as CAHSEE MC items (58 samples for mathematics, 37 for ELA).

Scoring—Respondents were asked to consider a scoring rubric with a 0 to 4 point scale ranging from no evidence to ample evidence as recommended by the AB 2040 panel. About three-fourths of the respondents agreed or strongly agreed that this model rubric could provide consistent evaluation of any type of student evidence. Estimated median time to score work samples for the streamlined option, following initial training, was two hours for mathematics and four hours for ELA.

HumRRO concluded that while a Tier II COE might be feasible, additional research would be needed to develop specific requirements. They recommended that a pilot study be conducted before a system becomes operational. Copies of the report can be found on the CDE Independent Evaluations Web page at http://www.cde.ca.gov/ta/tg/hs/evaluations.asp.

Conceptual Framework

The following section describes the purpose and conceptual framework of the pilot study as a proof of concept. The concept is the utility of a COE as an alternative means for SWDs to meet the CAHSEE requirement.

Proof of Concept

The CAHSEE Alternative Means Pilot Study was intended to explore methodologies and challenges faced by gathering a variety of work samples from LEAs, assigning consistent scores, and drawing conclusions that might better inform development of a full COE. The proof of concept activity was focused on various work sample types to capture student work across the range of California academic standards measured on the CAHSEE. The pilot was designed to avoid a large sample from any individual student to prevent the impression that a valid alternative was in place or that the student's work submission might meet the CAHSEE requirement. While the pilot itself was focused on exploring possible task types, practitioners from the field were surveyed about the likely benefits and challenges of a full COE beyond the task-level efforts of the pilot. The intent was to inform possible next steps in fully implementing an alternative means process.

Collection of Evidence

Tier II is envisioned to represent a broad, structured collection of student work samples. One goal of the pilot was to explore task types. Each student could complete up to three work samples using five types of evidence:

- On-demand writing samples or mathematics tasks
- On-demand classroom tasks (e.g. tests or quizzes)
- Classroom prepared tasks (e.g. reports or projects)
- Computer presentations (e.g. slideshow or graphics)
- Video or audio presentation

The on-demand writing and mathematics tasks were created by ETS for the pilot study. All on-demand tasks were performance assessments.

Methods

The research methodology for the pilot study is discussed with regard to recruitment and sampling, directions for administration, evaluation procedures, survey methodology, and focus group activities.

Recruitment

All California LEAs with high school students were invited to participate in the alternative means pilot study. A letter encouraging LEA and high school student participation was sent by the CDE to district and county superintendents, charter school administrators, and high school principals. The letter was followed by an email from ETS to all CAHSEE coordinators. Students eligible to participate in the pilot study include:

- SWDs currently enrolled in grades eleven or twelve who have taken the CAHSEE, whether or not they have passed.

- Students without disabilities enrolled in grades eleven or twelve who have taken the CAHSEE, whether or not they have passed, and achieved a CAHSEE scale score within the range of 325 to 375[14].

Initially, 76 LEAs expressed an interest in participating in the pilot and 66 LEAs were sent pilot study materials during the first week of May 2011. Nine of these LEAs submitted student work samples for evaluation in June. Although original expectations were to collect between 4,400 and 6,500 work samples, a total of 508 work samples were submitted. The low return rate may be due to the request coming late in the school year.

Matrix Sampling of Strands

In order to ensure adequate coverage of the content while minimizing the effort required by teachers and students, each participating LEA was assigned CAHSEE content strands according to a sampling matrix. Each of the six ELA strands was matched with one of the five mathematics strands, for a total of 30 different ELA and mathematics strand combinations. For example, the first participating LEA was assigned word analysis and probability and statistics, the second LEA was assigned reading comprehension and number sense, and so on. At least one of the three work samples from each student was required to be an “on-demand” task. Other samples could consist of regular classroom work or a media presentation.

[14] For purposes of the pilot study only, general education students were included in the sample in order to provide a wider range of performance. In a statewide operational implementation of CAHSEE alternative means, eligibility would be restricted to SWDs as defined by statute.

Directions for Administration

The CAHSEE Directions for Administration–Alternative Means Pilot Study (CDE, 2011b) provides information to participating California teachers, Site Coordinators, and LEA CAHSEE Coordinators about identifying, administering, collecting, and submitting work samples for the study participants.

Included in the manual is the Student Work Sample Submission Form which collects information on each work sample, including date the work sample was completed, Statewide Student Identifier (SSID), LEA name, course in which the student completed the work, content area, strand and standard code of which the work sample is aligned, type of work sample, and a description of the work sample.

The Student Information/Signature Form gathered demographic information about the student. In addition, teachers were required to certify that each work sample was the student's original work.

Evaluation Procedures

This section describes the procedures developed to evaluate the student work samples submitted for the pilot study. ETS personnel facilitated the scoring. Content-related inquiries were answered and scoring conflicts were resolved as they arose.

ETS staff conducted a meeting in order to fine-tune scoring rubrics and procedures and to select work samples to be used as anchor papers for training evaluators. Current high school teachers who had been screened and who had participated in previous CAHSEE item review sessions were recruited for participation. Twenty ELA and 14 mathematics teachers participated in evaluating student work samples and focus group discussions. The generic rubric was used to score each student submission. The four-score point rubric consisted of scores ranging from 0 for no evidence to 3 for adequate evidence, as described in Figure 11.2.

Figure 11.2: CAHSEE Alternative Means Work Sample Evaluation Rubric

3. There is *adequate evidence* that the student has demonstrated the skills and knowledge stated in the standard being addressed.

    - Completes most or all of the task or approximately 70% or more completed.
    - Demonstrates ability to master task; work attempted is displayed.
    - Correct answer or correct logic, equations and assumptions, but may display minor errors.
    - Displays consistent clarity and facility in the usage of language with minor grammatical errors.

2. There is *some evidence* that the student has demonstrated the skills and knowledge stated in the standard being addressed.

- Partially completes task or approximately 50% or more completed.

- Demonstrates partial ability to master task; some attempt at work is displayed.

- May have correct answer with incorrect logic, equations or assumptions, or incorrect answer with correct logic, equations or assumptions; will display errors.

- Displays some facility in the usage of language and grammatical errors may or may not affect clarity or understanding.

1. There is *little evidence* that the student has demonstrated the skills and knowledge stated in the standard being addressed.

    - Incomplete task or approximately 25% or more completed.

    - Demonstrates little ability to complete or master task and little or no attempt at work is displayed.

    - May display a correct or incorrect answer without displaying little or any attempt at logic or equations or may display an incorrect attempt at work.

    - Displays inadequate facility in the usage of language and grammatical errors affect clarity or understanding.

0. N*o evidence* that the student has the skills and knowledge stated in the standard being addressed.

    - N/S = non-scorable or not assessed
    - Blank
    - Non responsive
    - Illegible
    - Other language
    - Not aligned to standard

In order to train the teachers, the selected materials covered the spectrum of different student response types. Evaluators were given an orientation and introduction to the task. The ELA and mathematics evaluators received instruction on the training materials specific to their subject area.

Survey Methodology

Survey data collection was part of the pilot study design. An email was sent to all CAHSEE coordinators inviting them to participate in an online survey regarding

alternative means. All interested high school staff in their LEA, particularly staff who worked with SWDs, were invited to participate.

The survey consisted of four sections: *respondent profile, experience with CAHSEE*, *CAHSEE alternative means*, *and alternative means pilot study*. The profile section asked for current job responsibilities (e.g. classroom teacher, school administrator); the experience with CAHSEE section asked about familiarity with content standards and experience with SWDs; the alternative means section asked about the feasibility of implementation, and the alternative means pilot study section asked for study participants' reactions. The majority of the survey items presented statements to which respondents could select the level to which they agreed or disagreed on a four-point scale. Respondents also were provided opportunities to add open-ended comments to the CAHSEE alternative means and alternative means pilot study sections.

Focus Groups

The teacher evaluators participated in focus groups following the scoring. The evaluators were presented with these topics, prior to scoring, so that they would be able to keep notes and formulate opinions as they participated in the scoring process. For the scoring of the pilot assessment, the groups were divided by content area. The focus group topics included:

- Rubrics and evaluation procedures
- Alternative means survey
- Directions for administration and submission forms
- Operationalized submission of work samples by strand

Each group was led by a facilitator who was familiar with the alternative means study. Following the verbal feedback portion, participants were invited to write any additional observations and comments anonymously.

Results

The results include the findings for the quantitative and qualitative data.

Quantitative Data

The quantitative data include a description of participating student characteristics and various analyses of the student work samples submitted, the scores for the work samples, correlations with CAHSEE scores, and calculations of rater agreement.

Student Characteristics

In total, 179 students participated in the pilot study, many of whom provided evidence for both content areas. There were 149 participants for ELA and 133 for mathematics from nine LEAs and 13 high schools. The number of students per school ranged from 1 to 35 for ELA and 2 to 34 for mathematics.

Table 11.1 shows the demographic characteristics of the students by content area. There were more male than female participants for both content areas. Approximately half of the students were identified as having Hispanic ethnicity, and the most frequently reported race was white followed by African American. Note that a high percent of responses were missing for the Hispanic and race characteristics, 20 and 50 percent, respectively. Approximately 44 percent of the ELA and mathematics participants were from socioeconomically disadvantaged backgrounds, but three in ten teachers did not provide a response for their students. About 84 percent of the ELA students and 86 percent of the mathematics students were in either grade eleven or twelve. Slightly more than 60 percent of students were English only speakers and slightly less than 20 percent were EL students for each content area.

Table 11.1 Student Demographic Characteristics by Content Area

Demographic Group		ELA		Mathematics	
		Frequency	Percent	Frequency	Percent
Gender	Female	55	36.91	47	35.34
	Male	78	52.35	72	54.14
	Missing	16	10.74	14	10.53
Hispanic Ethnicity	No	46	30.87	35	26.32
	Yes	71	47.65	69	51.88
	Missing	32	21.48	29	21.80
Race	African American	21	14.09	13	9.77
	Other Asian	1	0.67	1	0.75
	Filipino	2	1.34	4	3.01
	Samoan	1	0.67	0	0.00
	White	37	24.83	28	21.05
	White/Filipino	1	0.67	1	0.75
	Korean	0	0.00	1	0.75
	American Indian	8	5.37	10	7.52
	Chinese	1	0.67	1	0.75
	Asian Indian	1	0.67	1	0.75
	Missing	76	51.01	73	54.89
Socioeconomic Disadvantage	No	40	26.85	35	26.32
	Yes	65	43.62	58	43.61
	Missing	44	29.53	40	30.08
Grade	Ten	10	6.71	10	7.52
	Eleven	74	49.66	65	48.87
	Twelve	51	34.23	49	36.84
	Missing	14	9.40	9	6.77
English Proficiency	English Learner	27	18.12	24	18.05
	English Only	93	62.42	83	62.41
	Initially Fluent	2	1.34	3	2.26
	Reclassified Fluent	9	6.04	9	6.77
	Missing	18	12.08	14	10.53
Total		**149**		**133**	

Table 11.2 shows the special educational considerations by content area. A high percent of participants, 78 for ELA and 77 for mathematics, had an IEP. None of the participating students were reported to have an active Section 504 plan. The most frequently reported disability was the specific learning disabilities category, which described 67 percent of the ELA and 63 percent of the mathematics participants. Approximately ten percent in each content area had no physical or learning disability. The majority of students in the sample used accommodations for testing.

Table 11.2. Special Education Considerations by Content Area

Special Education Considerations		ELA		Mathematics	
		Frequency	Percent	Frequency	Percent
Education Plan	IEP	116	77.85	103	77.44
	504 Plan	0	0.00	0	0.00
	Neither IEP or 504	15	10.07	15	11.28
	Missing	18	12.08	15	11.28
Special Education Services	None	15	10.07	15	11.28
	Autism	2	1.34	3	2.26
	Deaf	4	2.68	4	3.01
	Emotional Disturbance	1	0.67	1	0.75
	Hard of Hearing	3	2.01	4	3.01
	Intellectual Disabilities	1	0.67	1	0.75
	Other Health Impairment	3	2.01	5	3.76
	Specific Learning Disability	100	67.11	84	63.16
	Specific Learning Disability/Autism	2	1.34	2	1.50
	Speech or Language Impairment	3	2.01	2	1.50
	Missing	15	10.07	12	9.02
Accommodations	No	51	34.23	39	29.32
	Yes	82	55.03	80	60.15
	Missing	16	10.74	14	10.53
Total		**149**		**133**	

Teachers were asked to report CAHSEE scale score information for the individual participants. The descriptive statistics for the most recent CAHSEE score and the corresponding percentile ranks of 10, 25, 50, 75, and 90 are summarized in Table 11.3. The mean ELA scale score was below the passing performance level of 350; whereas, the mean mathematics scale score was slightly above the passing score.

Table 11.3 Teacher Reported "Most Recent" CAHSEE Test Results: Descriptive Statistics

Content	N	Mean	Std Dev	Min	Max	P10	P25	P50	P75	P90
ELA	134	341.64	26.66	275	412	309	323	343	359	374
Math	119	353.22	28.16	279	433	316	331	353	371	389

Note: There were 15 ELA and 14 mathematics students with missing teacher reported scale scores.

The number of times the participants had taken the CAHSEE is shown in Table 11.4. Most of the students had taken the CAHSEE more than once; approximately 44 percent two to three times and eleven percent four or more times. Table 11.5 provides the CAHSEE passing classification based on teacher reported CAHSEE scores by disability status for ELA and mathematics participants. The majority of the ELA participants had not passed the CAHSEE; whereas, the majority of mathematics participants had passed the CAHSEE.

Table 11.4 Teacher Reported Number of CAHSEE Attempts by Content Area

Number of Attempts	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
1	45	30.20	47	35.34
2	34	22.82	28	21.05
3	37	24.83	30	22.56
4	4	2.68	5	3.76
5	3	2.01	1	0.75
More than 5	10	6.71	8	6.02
Missing	16	10.74	14	10.53
Total	**149**	**100.00**	**133**	**100.00**

Table 11.5 provides the CAHSEE passing classification based on teacher reported CAHSEE scores by disability status for ELA and mathematics participants. The majority of the ELA participants had not passed the CAHSEE; whereas, the majority of mathematics participants had passed the CAHSEE.

Table 11.5 Teacher Reported CAHSEE Pass Status by Content Area and Student Sample

Content Area	Pass Status	SWD		non-SWD		Missing		Total
		N	Percent	N	Percent	N	Percent	
ELA	**Fail**	63	56.25	8	53.33	4	57.14	75
	Pass	49	43.75	7	46.67	3	42.86	59
	Total	112	100.00	15	100.00	7	100.00	134
Math	**Fail**	38	38.78	7	46.67	2	33.33	47
	Pass	60	61.22	8	53.33	4	66.67	72
	Total	98	100.00	15	100.00	6	100.00	119

Descriptive Results

Summary information about the work samples submitted, the scores assigned by the raters, and the correlations between the Alternative Means ratings and the CAHSEE strand scores are presented in this section. Tables 11.6 and 11.7 provide the number of student submissions by work sample and strand for ELA and mathematics, respectively. Evidence of student work was submitted for each work sample type for ELA and for three work sample types for mathematics. The vast majority of student work samples came from the Classroom Prepared Tasks and the On Demand Classroom Performance standardized tasks. Only the Classroom Prepared Tasks had student work samples representing all strands for both content areas.

Each student submission was scored independently by three raters. The rater scores are summarized by the p-value, which is standardized on a scale of 0 to 1, and calculated as the average rater score divided by the maximum score of three. Tables 11.6 and 11.7 show a range of difficulty across strand and work sample combinations. For ELA, the most difficult strand-task combination was Word Analysis submitted under the Classroom Prepared Task (p-value = 0.12). The least difficult strand-task combination was Writing Conventions submitted under the On Demand Classroom Performance Task (p-value = 0.77). For mathematics, the most difficult strand was Number Sense (p-value = 0.30) and the least difficult strand was Probability and Statistics (p-value = 0.82), both submitted under the On Demand Classroom Performance Task. Statistics were not provided for samples with less than 10 observations.

Table 11.6. Task Difficulty by Work Sample and Strand for ELA

Work Sample	Strand	ELA	
		N	p-value
Audio-Visual Presentation	RL	8	-
Computer Presentation	RW	2	-
	WS	15	0.39
Classroom Prepared Task	RC	40	0.33
	RL	23	0.37
	RW	10	0.12
	WA	18	0.33
	WC	1	-
	WS	25	0.41
On Demand Classroom Performance	RC	100	0.61
	RL	94	0.53
	RW	14	0.47
	WC	58	0.77
	WS	49	0.58
On Demand Writing Prompt	WA	5	-
	WC	1	-
	WS	4	-

Note: RC = Reading Comprehension, RL = Literary Responses & Analysis, RW = Word Analysis, WA = Writing Application, WC = Writing Conventions, WS = Writing Strategies; Statistics were not reported for samples of less than 10 observations.

Table 11.7. Task Difficulty by Work Sample and Strand for Mathematics

Work Sample	Strand	Mathematics	
		N	p-value
Computer Presentation	PS	3	-
Classroom Prepared Task	A1	78	0.49
	AF	36	0.67
	MG	4	-
	NS	16	0.48
	PS	28	0.65
On Demand Classroom Performance	A1	28	0.34
	AF	66	0.48
	MR	6	-
	NS	96	0.30
	PS	32	0.82

Notes: A1 = Algebra 1, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning, NS = Number Sense, PS = Probability & Statistics; Statistics were not reported for samples of less than 10 observations.

Since the constructs being measured by the Alternate Means assessment are the same as the ones being measured by the CAHSEE, one would expect the correlations to be significantly different from zero, even though the means of displaying competence are different. To investigate this question, the CAHSEE performance data were matched to the Alternate Means by student identification. The information was collapsed across work sample type. The correlations between the mean of the three rater scores and the corresponding CAHSEE strand raw scores are presented in Table 11.8. The correlations for ELA ranged from 0.50 for Writing Conventions to -0.20 for Writing Strategies. The correlations for mathematics ranged from 0.32 for Algebra and Functions to -0.11 for Algebra 1. Three of the ELA strands and two of the mathematics strands had correlations that were significantly different from zero.

Table 11.8. Correlations and Standard Errors between Rater Scores and CAHSEE Strand Scores

Content	Strand	N	Correlation	SE
ELA	Word Analysis	26	0.40*	0.19
	Reading Comprehension	63	0.34**	0.12
	Literary Responses and Analysis	64	0.23	0.12
	Writing Strategies	71	-0.20	0.12
	Writing Conventions	44	0.50**	0.13
	Writing Application	21	0.04	0.23
Mathematics	Probability and Statistics	56	0.15	0.13
	Number Sense	95	0.20*	0.10
	Algebra and Functions	65	0.32**	0.12
	Algebra 1	78	-0.11	0.11
	Measurement and Geometry	4	-	-

Note: Correlations are not reported for samples of less than 10 observations.

* $p \leq 0.05$; ** $p \leq 0.01$

Rater Agreement

Each strand by work sample was scored by three raters independently on a 0 to 3 scale. Rater agreement is summarized using the indices of rater consensus and intraclass correlations by work sample and strand. The results are shown in Tables 11.9 and 11.10 for ELA and mathematics, respectively. The consensus ratings express the percent of perfect agreement among the three independent raters. The consensus values ranged from 20 to 83 for ELA and from 52 to 78 for mathematics.

The intraclass correlation is a general measurement of agreement between multiple raters on the same set of observations, which are scores from student work samples in this case. The intraclass correlation assesses agreement by comparing the variability of ratings of the same response to the total variation across all ratings and responses. When raters score an item response the same way, this indicates that the total variation depends largely on the response attributes. The intraclass correlation ignores mean rater differences.

A higher intraclass coefficient reflects greater consistency of rater scores than a low intraclass correlation. Perfect agreement would yield a value of 1 and no agreement would yield a value of 0. The majority of the ELA and all of the mathematics work sample-strand combinations intraclass correlation coefficients were significant. The intraclass correlations ranged from 0.12 to 0.81 for ELA. The intraclass correlations indicate that the rater scores were most consistent for the Writing Application strand of the Classroom Prepared Task and for the Writing Strategies strand of the On Demand Classroom Performance. The rater scores were least consistent for the Writing Strategies strand for the Classroom Prepared Task. The intraclass correlations ranged from 0.65 to 0.93 for mathematics. The intraclass correlations indicate that the rater scores were most consistent for the Algebra and Functions strand and least consistent for the Probability and Statistics strand of the Classroom Prepared Task. Overall, the rater agreement was higher for mathematics than for ELA.

Table 11.9. Summary Statistics of Rater Agreement for ELA

Work Sample	Strand	N	Consensus	Intraclass Correlation	Standard Error
Audio-Visual Presentation	RL	8	-	-	-
Computer Presentation	RW	2	-	-	-
	WS	15	46.67	0.49	0.20
Classroom Prepared Task	RC	40	57.50	0.66*	0.09
	RL	23	60.87	0.75*	0.09
	RW	10	70.00	0.50	0.24
	WA	18	66.67	0.81*	0.08
	WC	1	-	-	-
	WS	25	20.00	0.12	0.20
On Demand Classroom Performance	RC	100	25.00	0.48*	0.08
	RL	94	40.43	0.56*	0.07
	RW	14	21.43	0.16	0.27
	WC	58	82.76	0.54*	0.09
	WS	49	48.98	0.80*	0.05
On Demand Writing Prompt	WA	5	-	-	-
	WC	1	-	-	-
	WS	4	-	-	-

Note: RC = Reading Comprehension, RL = Literary Responses & Analysis, RW = Word Analysis, WA = Writing Application, WC = Writing Conventions, WS = Writing Strategies; Statistics are not reported for samples of less than 10 observations.

* $p \leq 0.001$

Table 11.10. Summary Statistics of Rater Agreement for Mathematics

Work sample	Strand	N	Consensus	Intraclass Correlation	Standard Error
Computer Presentation	PS	3	-	-	-
Classroom Prepared Task	A1	78	78.21	0.91*	0.02
	AF	36	77.78	0.93*	0.02
	MG	4	-	-	-
	NS	16	56.25	0.89*	0.05
	PS	28	60.71	0.65*	0.11
On Demand Classroom Performance	A1	28	64.29	0.79*	0.07
	AF	66	51.52	0.80*	0.04
	MR	6	-	-	-
	NS	96	62.50	0.78*	0.04
	PS	32	78.13	0.88*	0.04

Note: A1 = Algebra 1, AF = Algebra & Functions, MG = Measurement & Geometry, MR = Mathematical Reasoning, NS = Number Sense, PS = Probability & Statistics; Statistics are not reported for samples of less than 10 observations.
* $p \leq 0.001$

Qualitative Data

The qualitative data include results from the survey and focus groups. General themes from the results are presented.

Survey Results

The survey results are summarized according to the four survey sections: respondent profile, experience with CAHSEE, CAHSEE alternative means, and alternative means pilot study.

Respondent Profile

The survey sample included 371 respondents who identified themselves as high-school teachers (62%), school or district administrators (27%), and other (11%). Nearly all respondents (97%) were directly involved with high school students.

Respondents were asked about the subject area currently taught. The most frequently endorsed response was not applicable (32%), likely from administrators. The other selections were special education (18%), ELA (16%), mathematics (13%), and multiple subject areas (10%), science (4%), social studies/history (3%). The majority of respondents (53%) reported having more than ten years of teaching experience, and about one in five reported more than 20 years teaching experience.

Experience with CAHSEE

Respondents were asked about their experience with the CAHSEE and SWDs. Almost all of the respondents indicated that they were familiar with CAHSEE administration procedures (98%), were familiar with the content standards measured

by CAHSEE (97%), that SWDs in their school or district were administered the CAHSEE with appropriate accommodations and modifications (95%), and were aware of SWDs who were likely to meet all other graduation requirements, but may not pass the CAHSEE (97%).

CAHSEE Alternative Means

The Alternative Means section addressed aspects of the process. A high proportion of respondents (85%) agreed or strongly agreed that SWDs who had not passed CAHSEE would be able to demonstrate high-school competency by alternative means. Close to 4 of 5 respondents agreed that a collection of work samples would accurately reflect what students know and are able to do.

Respondents were asked whether they collected student work samples during the school year that target standards measured by CAHSEE. When the “not applicable” group was excluded from the sample, 77 percent of respondents agreed or strongly agreed while 23 percent disagreed or strongly disagreed. Respondents endorsed that providing an alternative means to meet the CAHSEE requirement would increase academic expectations for SWDs (74%). Responses were evenly split between opinions about whether compiling a COE would place an undue burden on teachers. Nine of 10 respondents indicated that professional development would be required to implement CAHSEE alternative means.

There was an opportunity to provide additional comments or suggestions at the end of the survey. The responses fell into three categories.

- Some questioned whether many SWDs were capable of demonstrating the same level of academic achievement in the standards as required for passing the CAHSEE.

- Some felt that SWDs should be held to the same expectations as non-disabled students or continue to be exempt, questioning the need at all for an alternative means to the CAHSEE. The comments seemed evenly split between those who supported an alternative means and those who did not.

- Some questioned the mechanics of the CAHSEE alternative means as currently envisioned, particularly the details of compiling and scoring work samples.

Alternative Means Pilot Study

The last section of the survey was completed by those who participated in the pilot study and teachers to participate in the evaluation and focus group session. There were 77 respondents for this section. Respondents were asked to rate their satisfaction with the DFA, the ease of implementing each work sample type, the amount of time to compile the evidence, and whether the COE is an effective means of demonstrating academic achievement of the content standards.

Overall satisfaction was quite high, with 91 percent expressing that they were satisfied or very satisfied with the DFA. The highest ranking section of the DFA was the *Student/Information Signature Form*, with 86 percent being satisfied or very satisfied.

This was followed by the on-demand performance task for ELA included in the appendix of the DFA, with a satisfaction rating of 82 percent. The lowest satisfaction ratings were for the on-demand performance task for mathematics and the *Student Work Sample Submission Form*, which both rated 75 percent satisfied or very satisfied.

The easiest work sample types to implement was the on-demand classroom performance task followed by the on-demand writing prompt, where 61 and 58 percent endorsed easy or very easy, respectively. The most difficult work sample types were the audio/visual presentation and the computer presentation, where 24 and 30 percent endorsed difficult or very difficult, respectively. Note that there were between 29 and 45 percent endorsement of not applicable across the questions.

Respondents were asked to estimate the amount of time it took them to compile the COE submitted for each student. The responses here are referenced to a single student with no more than three work samples covering two CAHSEE content strands. The most popular response to this item was “less than one hour,” with 63 percent. Combined results indicate that 90 percent of respondents took two hours or less to compile evidence for a single student. Eighty percent agreed or strongly that the COE process is an effective means of demonstrating academic achievement of the content standards assessed by CAHSEE.

Focus Group Results

This section provides a synthesis of the outcomes derived from focus group discussions with 34 teachers. Follow-up interviews were conducted with 12 special education teachers and nine students who participated in the pilot, and three additional teachers who participated were interviewed one-on-one. Focus group discussions were structured around four topics:

- Rubrics and evaluation procedures
- Alternative means survey
- Directions for administration and submission forms
- Submission of work samples

<u>Student Input</u>

Students were asked for their general impressions of the CAHSEE examination, their experience in taking the examination, and their opinions of how CAHSEE alternative means would impact them and their fellow students. The students interviewed attended a high-performing charter high school, were SWDs, and although reported difficulty with the CAHSEE, most had passed.

Most of the students agreed that the CAHSEE was a fair measure of student knowledge and ability and felt the requirement should remain as a condition for graduation. One student expressed that the CAHSEE alternative means would affect students differently, depending on what subjects they had difficulty learning. A student, speaking American Sign Language through an interpreter, expressed that although she passed the mathematics portion of the CAHSEE on the first attempt, she struggled with the ELA portion, particularly with the essay. She thought that this was not due to her knowledge or ability to express herself, but due to discrepancies

between the English language and the American Sign Language. This student submitted a video recording of herself signing a response to literature for the pilot study and felt that the possibilities of the alternative means assessment would be beneficial to other students who communicate as she does.

Rubrics/Evaluation Procedures

Discussion in these sessions was focused on the appropriateness of the rubric used to score student work samples submitted for the pilot study and any changes or improvements that could be made.

Teachers were satisfied with the zero-to-three point holistic rubric used to score the pilot-study work samples, with some of them indicating that the use of multiple rubrics would be too cumbersome given the varied work sample types submitted. Some felt that the addition of a fourth score point was needed in order to give additional credit to students who go "*above and beyond*" adequate performance. The group stated that a score point of three should be interpreted as a passing score that met the CAHSEE requirement. The group believed that local scoring of the assessment would be less expensive, but may be subject to bias and possible inaccuracies.

Some evaluators expressed concern that some teachers who submitted work samples seemed unaware of the CAHSEE standards or had difficulty aligning the work samples to the standards. For example, some tasks were aligned to the wrong standard and others measured standards not assessed on the CAHSEE. There was a consensus that the implementation would require a great deal of training and support.

Teachers discussed what would entail a complete COE and how it should be evaluated. Two competing viewpoints emerged. One stance was that all the work samples should be as standardized as possible. For example, prescribed performance tasks would be administered during extended periods in the classroom. The other position was to allow the diversity of classroom work samples to function as appropriate accommodations for the target population.

Considerable discussion ensued regarding the appropriate focus of compiling the COE and evaluating it. All acknowledged that a focus on individual standards would result in excessive and non-equivalent demands on teachers and students. Competency by strand was deemed the more equitable option, with flexibility in the standards covered by the performance constrained by the judgment of adequacy used in evaluating the work.

Alternative Means Survey

Discussions with the teachers regarding the survey results confirmed the general findings and served to explicate some of the responses.

One item on the survey asked respondents to indicate the extent to which they agreed with whether they *collect student work samples (e.g. classroom quizzes, student essays, class projects) that target standards measured by CAHSEE and could be submitted as part of a collection of evidence for alternative means.* While the ELA committee felt that teachers would most likely agree that they could collect work samples in their regular classes, the mathematics committee felt that this could be a difficult undertaking because the CAHSEE measures middle school rather than high

school mathematics. Several other teachers related how they would not collect student work unless they knew well ahead of time that students were eligible for alternative means.

Another survey item asked respondents the extent to which they agreed that: *Providing an alternative means to meet the CAHSEE requirement would increase academic expectations for SWDs.* A relatively large proportion of respondents, 26 percent, disagreed with this statement. The focus group participants that that some respondents may believe an alternative to the CAHSEE would result in lower expectations, while others may believe that respondents may have high expectations for all students and an alternative measure would not alter these expectations.

Responses were split for the survey item that asked the extent of agreement with the statement that: *Compiling a collection of evidence for CAHSEE alternative means will place an undue burden on teachers of eligible students*. Fifty percent of respondents agreed, while 50 percent disagreed. Focus group participants felt that this response may have been influenced by administrators in the sample, who would likely disagree with the statement, while teachers were likely to agree. There was also a split between content areas, with ELA teachers tending to disagree and mathematics teachers tending to agree. This was primarily due to the misalignment of high school mathematics with CAHSEE content.

Submission of Work Samples

This topic focused on how teachers might select and compile appropriate student work samples for submission as a CAHSEE Tier II COE.

Two primary concerns emerged from the discussion. First, concern was expressed over the misalignment of content taught in typical high school mathematics classes and the content assessed by CAHSEE. Participants wanted to minimize any extra work required of teachers to the extent possible. The COE process could extend beyond the purview of the classroom.

A second concern was over the timeline for submission. Currently, students must take the CAHSEE at least once during their senior year in order to be eligible for alternative means. Participants expressed some doubt as to whether this eligibility requirement would provide adequate time to compile an adequate COE and receive a score in time for graduation decisions.

Summary of Findings

The quantitative data produced by the pilot study was too small to generalize the results. The results do, however, suggest several high level conclusions. The qualitative data from practitioners and experts produced by the survey and focus group activities provided a rich source of information to further inform the development of the CAHSEE alternative means process.

Collection of Evidence

The types of student work samples as defined in the pilot study seem practical for compiling a COE that demonstrates students' competency in the content standards

assessed on CAHSEE. Of the five work sample types, the vast majority of submitted work came from classroom prepared tasks (i.e.; an assignment, unit quiz, or chapter test completed in the classroom) and on-demand classroom performance (i.e.; a performance task provided by ETS and completed in the classroom). There were relatively few samples of audio-visual or computer presentations, and these were rated as most difficult to implement, although the samples submitted suggest that these types of tasks may work well for particular types of students.

High school ELA teachers affirmed that their regular classroom work was generally aligned with CAHSEE standards and would be amenable for collection, but high school mathematics teachers expressed concern about the alignment of their regular classroom work with CAHSEE content. With the exception of CAHSEE remediation classes, the content taught in junior or senior year mathematics classes does not align with the middle school mathematics content assessed by CAHSEE. This could pose a considerable challenge to high school mathematics teachers attempting to compile a COE.

The data show a split in opinion on whether a COE would place an undue burden on teachers of eligible students. Strong consensus was reached that teachers and other staff responsible for collecting evidence needed early notification of students' eligibility, specific guidelines for compiling evidence, and clear examples of appropriate work in order to ensure an effective and reasonably efficient collection process.

It was noted that numerous work samples were incorrectly coded to CAHSEE content strands, while other work samples measured standards not assessed on CAHSEE. This finding seems to indicate that detailed knowledge of the CAHSEE content standards is inconsistent. Some practitioners may lack a full understanding of the concept of alignment to standards. A robust professional development agenda and support system targeting educators who work with eligible students will be required to ensure that submitted work samples tightly align with the required CAHSEE content.

Scoring Procedures

The four-point holistic rubric used during the evaluation session functioned well for the calibration and evaluation sessions. The score points used were as follows: 0 = no evidence, 1 = little evidence, 2 = some evidence, and 3 = adequate evidence. Evaluators decided that the use of a holistic approach to evaluating Tier II screening was appropriate given the range of materials and work sample types represented in the pilot. The perception was that the holistic rubric could result in a somewhat quicker and uniform scoring process than the use of task-specific analytic rubrics.

Given that assessment of the overall performance is key to the determination of CAHSEE pass/fail scoring, it was also decided that an overall score of "3" was required to pass and adequately meet the CAHSEE requirement with a COE. One disadvantage to this approach is that only limited feedback can be provided to the student in the event that students are allowed to resubmit a collection that did not pass initially.

Evaluators felt that centralized state-level scoring by highly trained raters was preferable to regional LEA-level scoring in order to ensure equitable and unbiased

scoring procedures. The consensus ratings and correlations across multiple raters used to evaluate the limited student work submitted for the pilot were not high enough to produce reliable scores on a large-scale assessment. Additional work in developing evaluation procedures and materials will be needed.

Passing Rates

The vast majority of students who participated in the pilot study were receiving Special Education services and had IEPs. Nearly two-thirds of the students had taken the CAHSEE repeatedly and not passed. Estimated correlations between alternative means and CAHSEE strand scores were generally low-to-moderate, possibly indicating related, but differentiated performance. These findings, while not definitive, hint at the possibility of increased performance and passing rates for eligible students provided an alternative means to CAHSEE.

Opinions were not uniform with regards to whether SWDs who have not passed CAHSEE would be able to demonstrate high-school competency through a COE. Some questioned the functional equivalence of anything alternative, while others expressed concern over the independence of student work, how much work is enough to meet the requirements of a full collection, about who does the scoring, and how passing criteria would be set.

Timeline

The current regulations and timeline for CAHSEE alternative means establishes July 2012 as the implementation date for an operational assessment. The status of assessment development suggests that this date is not realistic. The intent of the pilot study was to inform California's possible next steps in fully implementing a COE for a Tier II process. Following traditional development procedures for large-scale CR assessments, a practical next step would be a field test designed to

- Establish guidelines for administering and compiling work samples
- Fine-tune scoring procedures and rubrics
- Establish validity and reliability for a full COE
- Develop methods to report the results in a timely fashion and in ways that are meaningful to all stakeholders

There is simply not enough time remaining before the July 2012 deadline to accomplish these essential tasks.

Recommendations

The following recommendations are not presented in order of importance, but follow thoughts based on the available evidence.

Recommendation 1: A full-scale census field test should take place prior to operational implementation of Tier II.

A much more representative sample of student work than that collected for the pilot study is needed in order to fully understand how teachers will respond to the complete COE process. A field test will allow for further improvements in the directions for

administration, training materials, support services, performance level descriptors, and scoring procedures. In addition, a standard-setting procedure should take place prior to operational implementation of Tier II, requiring a much larger and varied sample of student work. Materials and procedures can be re-validated once CAHSEE alternative means become operational.

The pilot study results reported here are limited in scope, due to the short timeframe available for its implementation and the low response rate from participating LEAs. The study design did not envision the collection and submission of a complete COE; rather, only pieces of the process were explored. The pilot work was approved in March and begun in April, with materials distributed in May and collected in early June. The brief window came at the end of the school year immediately following state testing, leading to limited participation. A full-scale census field test to resolve these limitations of the pilot will need to be timed to the seasonal demands of California high schools.

Recommendation 2: The CDE should consider earlier identification of Tier II eligible students, prior to the commencement of a student's senior year.

There was concern expressed over the eligibility requirements for participation in Tier II as currently conceptualized. In practice, the requirement that students have to take the CAHSEE at least once during their senior year likely will mean that students will not be identified as eligible until winter of their senior year. This results in a very restricted time period in which to develop, assemble, submit, and score an appropriate COE prior to graduation decisions. Furthermore, it does not allow for any type of feedback loop or resubmission process should the submitted collection fall short of CAHSEE requirements.

It is recommended that students be identified as eligible for CAHSEE alternative means no later than the end of their junior year, and that initial submission of an alternative means COE occurs in the fall of their senior year, at the latest. This would allow adequate time for collecting evidence and possibly a second submission; if needed, to get evaluative feedback with regard to missing evidence and the adequacy of the overall performance prior to graduation. This would also encourage collection of materials over a longer period, which may be necessary for many eligible students in this population.

Recommendation 3: Evaluating the Tier II collection of evidence should be approached holistically and conducted at the state level.

Holistic evaluation is a procedure for scoring varied student work samples in which the evaluator makes a single judgment of the overall quality of the response, instead of awarding points separately for different features. Evaluation of student work with a holistic rubric is especially appropriate for complex learning tasks or for types of tasks that integrate content from more than one area, such as those found in the current study. Trained evaluators would use a scoring guide that describes a typical response at each score level, along with exemplar responses that serve as illustrations of each score level, and would be calibrated with continual monitoring of scoring and inter-rater reliability calculations.

Concerns for potential bias and inconsistent reliability of scoring at the local level lead to the recommendation that the collections of evidence for CAHSEE alternative means be submitted to the state for centralized scoring by highly trained and monitored evaluators.

Recommendation 4: The focus of student work samples collected for CAHSEE alternative means should be at the strand level.

The current threshold for a passing score on the CAHSEE is approximately 60 percent correct for ELA and 55 percent correct for mathematics, and the goal of an alternative means assessment is to create an alternate pathway for eligible students to demonstrate the same level of achievement as students who pass the CAHSEE. It should not be a requirement that all CAHSEE standards be addressed individually in this alternative assessment, as this is more than is required for the passing of the regular CAHSEE. The CAHSEE alternative means COE need not address every standard on the CAHSEE assessment.

Required work samples could be sophisticated enough to cover multiple standards within a strand or even across strands. A judgment can then be made as to whether the evidence is adequate to support a passing score on each strand. The state will have to make a decision as to how many strands students should be required to perform adequately in order to meet the overall CAHSEE requirement. One consideration for an overall holistic score is in terms of a compensatory model that combines strand scores, where strong performance in one strand could compensate for weaker performance in another strand. This would require the reinstatement of the “4 - ample evidence” rating not utilized in the scoring rubric for the current study. In this manner, various combinations of strand scores could be averaged to produce an overall “3 - adequate evidence” rating required for passing CAHSEE alternative means.

Recommendation 5: The state should consider providing a bank of prescribed on-demand performance tasks for each CAHSEE strand, and allow other work sample types as supplements.

In order to maximize the effectiveness of CAHSEE alternative means, a reasonable balance must be struck between constraints placed on acceptable work sample types and total autonomy to select whatever may be judged suitable for a particular student. While the educators involved with the pilot study felt that the sample types provided were appropriate for the assessment, they also strongly supported the idea of “structure with flexibility” when delineating the contents of an alternative means COE.

A well-developed and field-tested bank of prescribed on-demand performance tasks of the type utilized for this pilot study could provide the desired structure and could mitigate several issues. Submission of a performance task for each strand would ensure adequate and equitable coverage of the CAHSEE content standards, provide for clear exemplars of performance to users, and make the scoring and evaluation process much more efficient and accurate. Selection of the particular performance task to be submitted and the selection of additional work as supplements allows for balanced flexibility in adapting submissions for a particular student. Although this

approach would add additional development time and effort, these would be balanced by decreasing the time and effort needed to score and evaluate the work.

Recommendation 6: Make guidance to the field as simple and specific as possible.

The directions to the field for collecting and submitting an adequate Tier II COE should be both simple and specific. While the pilot was intentionally designed to be open to many different submission options, the consistent feedback from most of the educators involved in the study is to make manuals and support materials released to the field very specific and easy to understand. Detailed checklists dictating how each strand should be addressed with regard to both accuracy and quantity must be provided, including multiple examples of an adequate performance for each required strand.

Practitioners are more likely to embrace this style of assessment if they have a very clear understanding of what is expected of them and why. In addition to reducing the frustration level of teachers and setting clear expectations for students, increasing standardization of submission will also simplify receiving, scoring, and reporting on this assessment.

Recommendation 7: Provide for a robust professional development program focused on the required CAHSEE content and guidance toward appropriate student work samples for submission.

Pilot study participants indicated the need for training in order to understand the standards assessed by CAHSEE. While 95 percent of the survey respondents said they were familiar with the standards assessed through the CAHSEE, the evidence submitted for the pilot often did not match the specific standard the teacher stated they were attempting to address. Ongoing professional development for the content assessed by the CAHSEE alternative means, particularly targeting educators who work with eligible students, will be necessary.

One suggestion for increasing teachers' understanding of the content standards would be to create a set of online tutorials on how to address each required strand for the CAHSEE alternative means COE. Professional development such as this would be an effective way to share standardized information with a large group of teachers. These tutorials do not necessarily need to be restricted to teachers working on the CAHSEE alternative means submission. They could be helpful to teachers working with any student having trouble with the CAHSEE, ultimately reducing the number of students who need to participate in an alternative means.

References

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Institutes for Research (2010). *Independent evaluation study of certain students who used modifications and/or accommodations on the California High School Exit Examination (CAHSEE).* Washington, DC: American Institutes for Research.

California Department of Education. (2006). *Mathematics Framework for California Public Schools*. Sacramento, CA: California Department of Education.

California Department of Education. (2010a). *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. (2010b). *California High School Exit Examination Directions for Administration (July 2010–May 2011).* Sacramento, CA; California Department of Education.

California Department of Education. (2010c). *California Content Standards*. http://www.cde.ca.gov/be/st/ss/*.*

California Department of Education. (2011a). *California High School Exit Examination Alternative Means Pilot Study*. Sacramento, CA: California Department of Education.

California Department of Education. (2011b). *California High School Exit Examination Directions for Administration – Alternative Means Pilot Study*. Sacramento, CA: California Department of Education.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

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.

Hardoin, M. M., Wise, L. L., & Buckland, W. W. (2010). Exploration of alternative means for students with disabilities to meet the CAHSEE requirement. In D. E. Becker, L. L. Wise & C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination: 2010 Evaluation Report.* (FR-10-56). Alexandria, VA: Human Resources Research Organization.

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., Wise, L. L., Thacker, A. & Hardoin-Mandeville, M. (2008) 2008 review of CAHSEE test quality. In D. E. Becker, L. L. Wise, and C. Watters (Eds.). *Independent Evaluation of the California High School Exit Examination (CAHSEE): 2008 Evaluation Report.* (FR-08-100). Alexandria, VA: Human Resources Research Organization.

Taylor, L., & Wise, L. L. (2009). 2009 review of CAHSEE test quality. 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.

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.

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.