California High School Exit Examination

Technical Report

July 2008–May 2009 Administrations

Prepared by

Educational Testing Service

for

the California Department of Education

August 2010

Copyright © 2010 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 2008–09 School Year *11*
Limitations of the Assessment *11*
Groups and Organizations *12*
OVERVIEW OF THIS 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 2008 29
APPENDIX 2.B: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—OCTOBER 2008 38
APPENDIX 2.C: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—NOVEMBER 2008 51
APPENDIX 2.D: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—DECEMBER 2008 65
APPENDIX 2.E: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—FEBRUARY 2009 74
APPENDIX 2.F: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MARCH 2009 89
APPENDIX 2.G: RESULTS OF TESTING VARIATIONS AND DISABILITY ANALYSES—MAY 2009 104

CHAPTER 3: ITEM DEVELOPMENT 117
1.0 Item Specifications *117*
2.0 Prepare Item Development Plan *119*
3.0 Train Item Writers *119*

4.0 Order Items based on Needs Assessment *120*
5.0 Internal Item Reviews (Educational Testing Service) *120*
6.0 External Item Reviews (California Educators and California Department of Education) *122*
7.0 Item Banking *125*
8.0 Create Field Test Sets *126*
9.0 Field Test Items *127*
10.0 Review and Process Field Test Data *128*
11.0 Create Operational Forms with California Department of Education Approval *129*
12.0 Administer Operational Test *129*

CHAPTER 4: TEST DEVELOPMENT 130
TEST LENGTH 130
RULES FOR ITEM SELECTION 130
Test Blueprint *130*
Content Rules and Item Selection *131*
Psychometric Criteria *131*
Rules for Item Sequence and Layout *136*

CHAPTER 5: TEST ADMINISTRATION 137
TEST SECURITY AND CONFIDENTIALITY 137
Educational Testing Service's Office of Testing Integrity *137*
Test Development *137*
Item Review *138*
Item Bank *138*
Transfer of Forms and Items to the California Department of Education *139*
Printing *139*
Test Administration *139*
Test Delivery *139*
Processing Security *141*
Data Management *141*
Transfer of Files via Secure Data Exchange *141*
Scoring and Analysis *142*
Reporting and Posting Results *142*
Student Confidentiality *142*
Data Security *142*
PROCEDURES TO MAINTAIN STANDARDIZATION 143
Test Administrators *143*
CAHSEE Directions for Administration *144*
CAHSEE District and Test Site Coordinator's Manual *145*
CAHSEE Online *145*
Test Booklets *145*
ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES 145
Identification of Students with Test Variations, Accommodations or Modifications *146*
Scoring *146*
DEMOGRAPHIC DATA CORRECTIONS 146
TESTING IRREGULARITIES 146
TEST ADMINISTRATION INCIDENTS 147

CHAPTER 6: ANALYSES 148
SAMPLES USED FOR THE ANALYSES 149
CLASSICAL ITEM ANALYSES 150
Summary of Item Statistics *153*
Procedures for Documenting Items That Fail to Meet the Desired Psychometric Characteristics *154*
DIFFERENTIAL ITEM FUNCTIONING ANALYSES 154
ITEM RESPONSE THEORY ANALYSES (CALIBRATION, SCALING AND EQUATING) 157
Measurement Model *157*
Item Calibration and Scaling *157*

True-Score Equating *158*
Equating Braille, Large-Print and Audio CD Forms *159*
Raw-Score to Scale-Score Conversion Tables and Conditional Standard Errors of Measurement .. *159*
Scaling Field Test Items *160*
Item Response Theory Model-Data Fit Analyses *160*
Summaries of Scaled Item Response Theory B-Values *161*
RELIABILITY ANALYSES 161
Test Reliabilities and Standard Errors of Measurement *161*
Strand Intercorrelations, Reliabilities and Standard Errors of Measurement *163*
Subgroup Reliabilities and Standard Errors of Measurement *163*
Writing Prompt and Rater Agreement Summary *164*
Decision Classification Analyses *165*
VALIDITY EVIDENCE 166
Test Purpose *167*
Constructs to Be Measured *167*
Intended Test Population *167*
Scores Generated and the Interpretations and Uses of These Scores *168*
Evidence Based on Content *168*
Evidence Based on Relations to Other Variables *170*
Evidence Based on Response Processes *172*
Evidence Based on Internal Structure *173*
Additional Validity Evidence *175*
Conclusions *175*
APPENDIX 6.A: CAHSEE ITEM REVIEW—DESCRIPTION AND EXAMPLES OF CLASSIFICATION CATEGORIES .. 176
APPENDIX 6.B: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—JULY 2008 179
APPENDIX 6.C: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—OCTOBER 2008 184
APPENDIX 6.D: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—NOVEMBER 2008 189
APPENDIX 6.E: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—DECEMBER 2008 194
APPENDIX 6.F: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—FEBRUARY 2009 199
APPENDIX 6.G: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MARCH 2009 204
APPENDIX 6.H: SUMMARY STATISTICS FOR OPERATIONAL ITEMS—MAY 2009 209
APPENDIX 6.I: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—FEBRUARY 2009 214
APPENDIX 6.J: SUMMARY STATISTICS FOR FIELD-TEST ITEMS—MARCH 2009 220
APPENDIX 6.K: INTERCORRELATIONS, RELIABILITY ESTIMATES AND STANDARD ERRORS OF MEASUREMENT 229
APPENDIX 6.L: RATER AGREEMENT ANALYSES 236
APPENDIX 6.M: GENERALIZABILITY ANALYSES 243
APPENDIX 6.N: DECISION CLASSIFICATION RELIABILITY ANALYSES 245
APPENDIX 6.O: SCORING TABLES FOR OPERATIONAL AND SPECIAL TEST VERSIONS 259
APPENDIX 6.P: STANDARD ERRORS OF THETA BASED ON WEIGHTED RAW SCORES 279

CHAPTER 7: PERFORMANCE STANDARDS 280
BACKGROUND AND PROCEDURES 280
CAHSEE PASSING SCORES 281
RESULTS 282

CHAPTER 8: SCORING AND REPORTING 284
PROCEDURES FOR MAINTAINING AND RETRIEVING INDIVIDUAL SCORES 284
Scoring and Reporting Specifications *284*
Scanning and Scoring *285*
TYPES OF SCORES AND STRAND SCORES 286
Raw Score *286*
Strand Score *286*
Scale Score *286*
Passing Scores *286*
Score for Writing Applications *286*
SCORE VERIFICATION PROCEDURES 287
Scoring Key Verification Process *287*

Monitoring and Quality Control of Writing Scoring *287*
Quality Control in Raw-to-Scale Score Conversions *290*
OVERVIEW OF SCORE AGGREGATION PROCEDURES 291
Individual Scores *291*
Group Scores *291*
REPORTS TO BE PRODUCED AND SCORES FOR EACH REPORT 292
Types of Score Reports *293*
Score Report Contents *293*
Score Report Applications *294*
CRITERIA FOR INTERPRETING TEST SCORES 294
CRITERIA FOR INTERPRETING SCORE REPORTS 294
APPENDIX 8.A: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—JULY 2008 296
APPENDIX 8.B: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—OCTOBER 2008 304
APPENDIX 8.C: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—NOVEMBER 2008 312
APPENDIX 8.D: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—DECEMBER 2008 320
APPENDIX 8.E: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—FEBRUARY 2009 328
APPENDIX 8.F: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—MARCH 2009 336
APPENDIX 8.G: FREQUENCY DISTRIBUTIONS AND DEMOGRAPHIC SUMMARIES—MAY 2009 344

CHAPTER 9: QUALITY CONTROL PROCEDURES 352
QUALITY CONTROL OF ITEM DEVELOPMENT 352
Item Specifications *352*
Item Writers *352*
Internal Contractor Reviews *352*
Content Expert Reviews *353*
Statewide Pupil Assessment Review Panel Review *354*
Data Review of Field Tested Items *354*
QUALITY CONTROL OF THE ITEM BANK 354
QUALITY CONTROL OF TEST MATERIALS 355
Collecting Test Materials *355*
Processing Test Materials *355*
QUALITY CONTROL OF SCANNING 356
Post-scanning Edits *356*
QUALITY CONTROL OF IMAGE EDITING 357
QUALITY CONTROL OF ANSWER DOCUMENT PROCESSING AND SCORING 357
Processing of Answer Documents *357*
Scoring and Reporting Specifications *357*
Matching Information on CAHSEE Answer Documents *357*
Matching Multiple-Choice and Writing Scores for English-Language Arts *357*
Storing Answer Documents *358*
QUALITY CONTROL OF PSYCHOMETRIC PROCESSES 358
Scoring Key Verification Process *358*
Quality Control of Item Analyses, Differential Item Functioning and Equating Process *358*
Score Verification Process *359*
Offloads to Test Development *359*
Independent Evaluation of the CAHSEE Program *359*
QUALITY CONTROL OF REPORTING 360
Excluding Student Scores from Summary Reports *360*

CHAPTER 10: HISTORICAL COMPARISONS 361
EXAMINEE PERFORMANCE 361
TEST CHARACTERISTICS 361
APPENDIX 10.A—HISTORICAL COMPARISONS ON STUDENT PERFORMANCE 362
APPENDIX 10.B—HISTORICAL COMPARISONS ON TEST CHARACTERISTICS 367

REFERENCES 369

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 2008)....3
Table E.3.2: CAHSEE Summary Statistics—English-Language Arts (November and December 2008, February 2009)....4
Table E.3.3: CAHSEE Summary Statistics—English-Language Arts (March and May 2009)....5
Table E.4.1: CAHSEE Summary Statistics—Mathematics (July and October 2008)....6
Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2008, February 2009)....7
Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2009)....8
Table 1.1: Testing Date for Each Administration by Content: 2008–09....11
Table 2.1: Items Included in the 2008–09 Administrations....18
Table 2.2: ELA and Mathematics Strands....19
Table 2.3: Subgroup Definitions....22
Table 2.4: CAHSEE Modification/Accommodation Table: 2008–09....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 2008....29
Table 2.A.2: Summary Statistics by Testing Variations and Disability—Mathematics, July 2008....29
Table 2.A.3: Demographic Summary for All Examinees by Testing Variations—ELA, July 2008....30
Table 2.A.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, July 2008....31
Table 2.A.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, July 2008....32
Table 2.A.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, July 2008....33
Table 2.A.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, July 2008....34
Table 2.A.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, July 2008....35
Table 2.A.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, July 2008....36
Table 2.A.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, July 2008....37
Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2008....38
Table 2.B.2: Summary Statistics by Testing Variations and Disability—Mathematics, October 2008....38
Table 2.B.3: Demographic Summary for All Examinees by Testing Variations—ELA, October 2008....39
Table 2.B.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, October 2008....40
Table 2.B.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, October 2008....41
Table 2.B.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, October 2008....42
Table 2.B.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, October 2008....43
Table 2.B.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, October 2008....45
Table 2.B.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, October 2008....47
Table 2.B.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, October 2008....49
Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2008....51
Table 2.C.2: Summary Statistics by Testing Variations and Disability—Mathematics, November 2008....51
Table 2.C.3: Demographic Summary for All Examinees by Testing Variations—ELA, November 2008....52
Table 2.C.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, November 2008....53
Table 2.C.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, November 2008....54
Table 2.C.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, November 2008....55
Table 2.C.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, November 2008....56
Table 2.C.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, November 2008....59
Table 2.C.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, November 2008....61
Table 2.C.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, November 2008....63
Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2008....65
Table 2.D.2: Summary Statistics by Testing Variations and Disability—Mathematics, December 2008....65
Table 2.D.3: Demographic Summary for All Examinees by Testing Variations—ELA, December 2008....66

Table 2.D.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, December 2008.....67
Table 2.D.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, December 2008.....68
Table 2.D.6: Percentiles of Scale Scores for Students with Testing Variations— Mathematics, December 2008 69
Table 2.D.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, December 2008.....70
Table 2.D.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, December 2008 71
Table 2.D.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, December 2008 72
Table 2.D.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, December 2008 73
Table 2.E.1: Summary Statistics by Testing Variations and Disability—ELA, February 2009 74
Table 2.E.2: Summary Statistics by Testing Variations and Disability—Mathematics, February 2009.....74
Table 2.E.3: Demographic Summary for All Examinees by Testing Variations—ELA, February 2009.....75
Table 2.E.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, February 2009 76
Table 2.E.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, February 2009 77
Table 2.E.6: Percentiles of Scale Scores for Students with Testing Variations — Mathematics, February 2009.....78
Table 2.E.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, February 2009 79
Table 2.E.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, February 2009.....82
Table 2.E.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, February 2009.....85
Table 2.E.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, February 2009 87
Table 2.F.1: Summary Statistics by Testing Variations and Disability—ELA, March 2009.....89
Table 2.F.2: Summary Statistics by Testing Variations and Disability—Mathematics, March 2009 89
Table 2.F.3: Demographic Summary for All Examinees by Testing Variations—ELA, March 2009 90
Table 2.F.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, March 2009.....91
Table 2.F.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, March 2009.....92
Table 2.F.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, March 2009 93
Table 2.F.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, March 2009.94
Table 2.F.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, March 2009.....97
Table 2.F.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, March 2009.....100
Table 2.F.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, March 2009.....102
Table 2.G.1: Summary Statistics by Testing Variations and Disability—ELA, May 2009 104
Table 2.G.2: Summary Statistics by Testing Variations and Disability—Mathematics, May 2009.....104
Table 2.G.3: Demographic Summary for All Examinees by Testing Variations—ELA, May 2009.....105
Table 2.G.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, May 2009 106
Table 2.G.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, May 2009 107
Table 2.G.6: Percentiles of Scale Scores for Students with Testing Variations—Mathematics, May 2009.....108
Table 2.G.7: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—ELA, May 2009 .109
Table 2.G.8: Summary of Scale Scores and Passing Rates by Disability and Testing Variations—Mathematics, May 2009.....111
Table 2.G.9: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—ELA, May 2009.....113
Table 2.G.10: Summary of Scale Scores and Passing Rates by Language Fluency and Testing Variations—Mathematics, May 2009.....115
Table 3.1 Field-test Percentages for the CAHSEE.....119
Table 3.2 Summary of Items and Forms Presented in 2009 CAHSEE 128
Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items.....134
Table 4.2: Number of Items and Mean B-Value Ranges by Content Strand for ELA 134
Table 4.3: Difficulty (B) and Discrimination (R-bis) Specifications for Mathematics 135
Table 4.4: Number of Items and Mean B-Value Ranges by Content Strand for Mathematics.....135
Table 6.1: Listing of Summary Tables for Items.....149
Table 6.2: Flagging Criteria for Classical Item Analyses 154
Table 6.3: DIF Categories 156
Table 6.4: Field-Test Items Recommended for Future Forms.....161
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 216
Table 6.I.6: Distribution of Field-Test Item DIF Classifications—ELA 218
Table 6.I.7: Distribution of Field-Test Item DIF Classifications—Mathematics 219

Table 6.J.1: Summary of Field-Test Item Statistics—ELA....220
Table 6.J.2: Summary of Field-Test Item Statistics—Mathematics....220
Table 6.J.3: IRT Model Data Fit Distribution of Field-Test Items—ELA....221
Table 6.J.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics....221
Table 6.J.5: Field-Test Items Containing Significant DIF....222
Table 6.J.6: Distribution of Field-Test Item DIF Classifications—ELA....227
Table 6.J.7: Distribution of Field-Test Item DIF Classifications—Mathematics....228
Table 6.K.1: Intercorrelations and Reliability Estimates by Section—July 2008....229
Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2008....230
Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 2009....231
Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 2009....232
Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 2009....233
Table 6.K.8: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for ELA—February 2009....234
Table 6.K.9: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—February 2009....234
Table 6.K.10: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for ELA—March 2009 235
Table 6.K.11: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—March 2009....235
Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2008....236
Table 6.L.2: Summary Statistics for the ELA Essay Item—July 2008....236
Table 6.L.3: Agreement of First and Second Ratings on the ELA Essay Item—October 2008....237
Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2008....237
Table 6.L.5: Agreement of First and Second Ratings on the ELA Essay Item—November 2008....238
Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2008....238
Table 6.L.7: Agreement of First and Second Ratings on the ELA Essay Item—December 2008....239
Table 6.L.8: Summary Statistics for the ELA Essay Item—December 2008....239
Table 6.L.9: Agreement of First and Second Ratings on the ELA Essay Item—February 2009....240
Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2009....240
Table 6.L.11: Agreement of First and Second Ratings on the ELA Essay Item—March 2009....241
Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2009....241
Table 6.L.13: Agreement of First and Second Ratings on the ELA Essay Item—May 2009....242
Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2009....242
Table 6.M.1: Generalizability Results—July 2008....243
Table 6.M.2: Generalizability Results—October 2008....243
Table 6.M.3: Generalizability Results—November 2008....243
Table 6.M.4: Generalizability Results—December 2008....243
Table 6.M.5: Generalizability Results—February 2009....244
Table 6.M.6: Generalizability Results—March 2009....244
Table 6.M.7: Generalizability Results—May 2009....244
Table 6.N.1: ESEA Reliability Classifications—July 2008....245
Table 6.N.2: Pass/Not Pass Classifications—July 2008....246
Table 6.N.3: ESEA Reliability Classifications—October 2008....247
Table 6.N.4: Pass/Not Pass Classifications—October 2008....248
Table 6.N.5: ESEA Reliability Classifications—November 2008....249
Table 6.N.6: Pass/Not Pass Classifications—November 2008....250
Table 6.N.7: ESEA Reliability Classifications—December 2008....251
Table 6.N.8: Pass/Not Pass Classifications—December 2008....252
Table 6.N.9: ESEA Reliability Classifications—February 2009....253
Table 6.N.10: Pass/Not Pass Classifications—February 2009....254
Table 6.N.11: ESEA Reliability Classifications—March 2009....255
Table 6.N.12: Pass/Not Pass Classifications—March 2009....256
Table 6.N.13: ESEA Reliability Classifications—May 2009....257
Table 6.N.14: Pass/Not Pass Classifications—May 2009....258
Table 6.O.1: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, July 2008....259
Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2008....260
Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2008....261
Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2008....262
Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2008....263
Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2008..264
Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2008....265
Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2008..266

Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2009.....267
Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2009 ..268
Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2009 269
Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2009.....270
Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2009 271
Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2009.....272
Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—July,2008 273
Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July,2008.....274
Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—October, November 275
and December, 2008.....275
Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— October, November, and December, 2008 276
Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—February, March, and May 2009.....277
Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— February, March, and May 2009 278
Table 7.1: Passing Scores on the February 2004 CAHSEE 281
Table 7.2: Summary of Cut Scores and Passing Rates Based on Equating Samples 283
Table 8.1: Listing of Frequency Distribution and Demographic Summary Tables.....292
Table 8.2: Types of CAHSEE Reports 293
Table 8.A.1: Frequency Distributions, ELA—July 2008.....296
Table 8.A.2: Frequency Distributions, Mathematics—July 2008 296
Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2008.....297
Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2008 297
Table 8.A.5: Demographic Summary for All Examinees ELA—July 2008.....298
Table 8.A.6: Demographic Summary for All Examinees, Mathematics—July 2008 299
Table 8.A.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA— July 2008.....300
Table 8.A.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—July 2008 301
Table 8.A.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—July 2008.....302
Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2008 .303
Table 8.B.1: Frequency Distributions, ELA—October 2008 304
Table 8.B.2: Frequency Distributions, Mathematics—October 2008.....304
Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2008.....305
Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2008 305
Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2008 306
Table 8.B.6: Demographic Summary for All Examinees, Mathematics—October 2008.....307
Table 8.B.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—October 2008 308
Table 8.B.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—October 2008.....309
Table 8.B.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—October 2008.....310
Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2008 311
Table 8.C.1: Frequency Distributions, ELA—November 2008 312
Table 8.C.2: Frequency Distributions, Mathematics—November 2008.....312
Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2008.....313
Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2008 313
Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2008 314
Table 8.C.6: Demographic Summary for All Examinees, Mathematics—November 2008.....315
Table 8.C.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—November 2008 316
Table 8.C.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—November 2008.....317
Table 8.C.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—November 2008.....318
Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2008.....319
Table 8.D.1: Frequency Distributions, ELA—December 2008 320
Table 8.D.2: Frequency Distributions, Mathematics—December 2008.....320
Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2008.....321
Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2008 321
Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2008 322
.Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2008 322
Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2008 323
Table 8.D.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—December 2008 324
Table 8.D.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—December 2008.....325
Table 8.D.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—December 2008.....326

Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2008 327
Table 8.E.1: Frequency Distributions, ELA—February 2009 328
Table 8.E.2: Frequency Distributions, Mathematics—February 2009 328
Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2009 329
Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2009 329
Table 8.E.5: Demographic Summary for All Examinees, ELA February 2009 330
Table 8.E.6: Demographic Summary for All Examinees, Mathematics—February 2009 331
Table 8.E.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—February 2009 332
Table 8.E.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— February 2009 333
Table 8.E.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—February 2009 334
Table 8.E.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—February 2009 335
Table 8.F.1: Frequency Distributions, ELA—March 2009 336
Table 8.F.2: Frequency Distributions, Mathematics—March 2009 336
Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2009 337
Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2009 337
Table 8.F.5: Demographic Summary for All Examinees, ELA—March 2009 338
Table 8.F.6: Demographic Summary for All Examinees, Mathematics—March 2009 339
Table 8.F.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—March 2009 340
Table 8.F.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics—March 2009 341
Table 8.F.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—March 2009 342
Table 8.F.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—March 2009 343
Table 8.G.1: Frequency Distributions, ELA—May 2009 344
Table 8.G.2: Frequency Distributions, Mathematics—May 2009 344
Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2009 345
Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2009 345
Table 8.G.5: Demographic Summary for All Examinees, ELA—May 2009 346
Table 8.G.6: Demographic Summary for All Examinees, Mathematics—May 2009 347
Table 8.G.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—May 2009 348
Table 8.G.8: ESEA Demographic Summary for All CAHSEE Examinees, Mathematics— May 2009 349
Table 8.G.9: Examinee Demographics Showing Mean Scale Score at Each Percentile, ELA—May 2009 350
Table 8.G.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—May 2009 351
Table 10.A.1: Number of Examinees Tested, Scale Score Means and Standard Deviations of CAHSEE across 2006–07, 2007–08, and 2008–09 362
Table 10.A.2: Percentage of Students Passing Each Content Area across 2006–07, 2007–08, and 2008–09 362
Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced across 2006–07, 2007–08, and 2008–09 .362
Table 10.A.4: Table 10.A.4: Scale Score Distributions across 2006–07, 2007–08, and 2008–09 for ELA (July to December) 363
Table 10.A.5: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for ELA (February to May) 364
Table 10.A.6: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for Mathematics (July to December) 365
Table 10.A.7: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for Mathematics (February to May) 366
Table 10.B.1: Average Proportion Correct of Operational Test Items across 2006–07, ...367
2007–08, and 2008–09 367
Table 10.B.2: Average IRT b-values of Operational Test Items across 2006–07, ...367
2007–08, and 2008–09 367
Table 10.B.3 Average Point-Biserial Correlation of Operational Test Items across 368
2006–07, 2007–08, and 2008–09 368
Table 10.B.4 Reliabilities and Standard Error of Measurement (SEM) of Operational Test Forms across 2006–07, 2007–08, and 2008–09 368

List of Figures

Figure 3.1 The ETS Item Development Process for CAHSEE 117
Figure 4.1: Example of Test Characteristic Curves 133
Figure 6.1: CAHSEE Item Analysis — Mathematics Item 153
Figure 6.2: CAHSEE Item Analysis — ELA Item 153
Figure 6.3: Samples of CAHSEE Item-Fit Rating Categories 178

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 2008 and February, March, and May 2009.

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 2008 and February, March, and May 2009 administrations. The majority of examinees in February and March were first-time examinees. Grade 10 students are only allowed to take the test in 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 were mostly 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 2008	20,616	5,994	7,379	7,243
October 2008	62,978	25,935	18,020	19,023
November 2008	151,860	66,307	42,415	43,138
December 2008	5,720	1,368	2,176	2,176
February 2009	223,090	168,388	26,703	27,999
March 2009	405,262	339,149	33,558	32,555
May 2009	63,173	22,154	20,627	20,392

[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 2008	13,373	3,953 (30%)	13,237	3,814 (29%)
October 2008	43,955	16,369 (37%)	44,958	14,947 (33%)
November 2008	108,722	43,566 (40%)	109,445	41,271 (38%)
December 2008	3,544	1,073 (30%)	3,544	1,259 (36%)
February 2009	195,091	129,841(67%)	196,387	131,009 (67%)
March 2009	372,707	270,600 (73%)	371,704	274,415 (74%)
May 2009	42,781	14,049 (33%)	42,546	13,233 (31%)

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 gridded fields were excluded for the purposes of the analyses reported in this technical report. Students having blank sections were excluded from the equating samples.

Highlights of the results for 2008–09 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 2008)

Administration	July 2008	October 2008
Scale Score Information		
Number of Examinees	13373	43955
Mean	337	342
SD	26	29
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	336	340
Raw Score Information		
Number of Examinees	13373	43955
Mean	47.09	49.80
SD[1]	12.44	14.04
Possible Range	0-90	0-90
Obtained Range	0-89	0-90
Median	47	50
Test Information		
Reliability	0.87	0.89
Raw Score Standard Error of Measurement	4.40	4.63
Mean Omits	0.60	0.63
SD Omits	4.62	4.82
Percent Responding to:		
All Items	88	88
All Items - 1 Item	97	97
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[2]		
Number of Items	72	72
Mean Observed Average Item Score (AIS)	0.53	0.56
Equated Mean Rasch B-Value	-0.01	-0.05
Mean R-biserial	0.41	0.45
SD R-biserial	0.09	0.10
Examinee Information		
ELA-Only Examinees		
Number of Examinees	7379	18020
Mean Scale Score	338	343
SD Scale Score	22	25
Median Scale Score	338	344
ELA and Mathematics Examinees		
Number of Examinees	5994	25935
Mean Scale Score	337	342
SD Scale Score	29	32
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 2008, February 2009)

Administration	November 2008	December 2008	February 2009
Scale Score Information			
Number of Examinees	108722	3544	195091
Mean	343	338	371
SD	33	29	41
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	342	336	371
Raw Score Information			
Number of Examinees	108722	3544	195091
Mean	52.14	50.07	62.04
SD	14.79	13.29	16.20
Possible Range	0-90	0-90	0-90
Obtained Range	0-90	3-90	0-90
Median	53	50	65
Test Information			
Reliability	0.87	0.89	0.94
Raw Score Standard Error of Measurement	4.40	4.33	3.99
Mean Omits	0.60	0.39	0.36
SD Omits	4.73	3.68	3.63
Percent Responding to:			
All Items	89	90	93
All Items - 1 Item	97	97	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)	0.59	0.56	0.71
Equated Mean Rasch B-Value	-0.19	-0.18	-0.08
Mean R-biserial	0.50	0.45	0.56
SD R-biserial	0.09	0.10	0.09
Examinee Information			
ELA-Only Examinees			
Number of Examinees	42415	2176	26703
Mean Scale Score	344	338	341
SD Scale Score	28	26	28
Median Scale Score	344	338	339
ELA and Mathematics Examinees			
Number of Examinees	66307	1368	168388
Mean Scale Score	343	337	376
SD Scale Score	37	34	41
Median Scale Score	340	332	379

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

Table E.3.3: CAHSEE Summary Statistics—English-Language Arts (March and May 2009)

Administration	March 2009	May 2009
Scale Score Information		
Number of Examinees	372707	42781
Mean	374	340
SD	39	33
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	373	337
Raw Score Information		
Number of Examinees	372707	42781
Mean	63.58	50.25
SD	15.12	15.01
Possible Range	0-90	0-90
Obtained Range	0-90	0-90
Median	66	50
Test Information		
Reliability	0.93	0.91
Raw Score Standard Error of Measurement	3.98	4.38
Mean Omits	0.29	0.76
SD Omits	3.11	5.44
Percent Responding to:		
All Items	93	88
All Items - 1 Item	98	96
All Items - 2 Items	99	98
All Items - 3 Items	99	98
All Items - 4 Items	99	98
All Items - 5 Items	99	99
Item Information[1]		
Number of Items	72	72
Mean Observed Average Item Score (AIS)	0.73	0.56
Equated Mean Rasch B-Value	-0.09	-0.12
Mean R-biserial	0.53	0.49
SD R-biserial	0.11	0.09
Examinee Information		
ELA-Only Examinees		
Number of Examinees	33558	20627
Mean Scale Score	342	341
SD Scale Score	29	30
Median Scale Score	340	339
ELA and Mathematics Examinees		
Number of Examinees	339149	22154
Mean Scale Score	377	340
SD Scale Score	38	36
Median Scale Score	378	335

[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 2008)

Administration	July 2008	October 2008
Scale Score Information		
Number of Examinees	13237	44958
Mean	340	343
SD	22	26
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	338	341
Raw Score Information		
Number of Examinees	13237	44958
Mean	37.06	38.44
Standard Deviation	11.23	12.22
Possible Range	0-80	0-80
Obtained Range	1-80	0-80
Median	36	38
Test Information		
Reliability	0.87	0.89
Raw Score Standard Error of Measurement	4.12	4.04
Mean Omits	0.40	0.49
SD Omits	3.05	3.72
Percent Responding to:		
All Items	86	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)	0.46	0.48
Equated Mean Rasch B-Value	-0.23	-0.23
Mean R-biserial	0.38	0.42
SD R-biserial	0.11	0.10
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	7243	19023
Mean Scale Score	342	344
SD Scale Score	19	20
Median Scale Score	342	343
ELA and Mathematics Examinees		
Number of Examinees	5994	25935
Mean Scale Score	337	341
SD Scale Score	26	29
Median Scale Score	333	338

Table E.4.2: CAHSEE Summary Statistics—Mathematics (November and December 2008, February 2009)

Administration	November 2008	December 2008	February 2009
Scale Score Information			
Number of Examinees	109445	3544	196387
Mean	345	345	373
SD	27	22	40
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	341	344	367
Raw Score Information			
Number of Examinees	109445	3544	196387
Mean	38.90	37.76	51.77
Standard Deviation	13.12	10.84	16.83
Possible Range	0-80	0-80	0-80
Obtained Range	0-80	5-79	0-80
Median	37	37	52
Test Information			
Reliability	0.87	0.85	0.95
Raw Score Standard Error of Measurement	4.12	4.13	3.71
Mean Omits	0.41	0.28	0.26
SD Omits	3.33	2.14	2.56
Percent Responding to:			
All Items	87	87	91
All Items - 1 Item	96	97	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	100	100
Item Information			
Number of Items	80	80	80
Mean Observed Average Item Score (AIS)	0.48	0.47	0.64
Equated Mean Rasch B-Value	-0.19	-0.11	-0.22
Mean R-biserial	0.44	0.37	0.57
SD R-biserial	0.09	0.10	0.09
Examinee Information			
Mathematics-Only Examinees			
Number of Examinees	43138	2176	27999
Mean Scale Score	346	346	345
SD Scale Score	21	18	22
Median Scale Score	345	345	343
ELA and Mathematics Examinees			
Number of Examinees	66307	1368	168388
Mean Scale Score	345	344	378
SD Scale Score	31	26	40
Median Scale Score	339	340	376

Table E.4.3: CAHSEE Summary Statistics—Mathematics (March and May 2009)

Administration	March 2009	May 2009
Scale Score Information		
Number of Examinees	371704	42546
Mean	378	342
SD	39	28
Possible Range	275-450	275-450
Obtained Range	275-450	275-450
Median	376	338
Raw Score Information		
Number of Examinees	371704	42546
Mean	54.21	38.19
Standard Deviation	16.26	13.55
Possible Range	0-80	0-80
Obtained Range	0-80	0-80
Median	56	36
Test Information		
Reliability	0.95	0.91
Raw Score Standard Error of Measurement	3.63	4.08
Mean Omits	0.22	0.58
SD Omits	2.17	4.35
Percent Responding to:		
All Items	91	86
All Items - 1 Item	98	96
All Items - 2 Items	99	98
All Items - 3 Items	99	99
All Items - 4 Items	100	99
All Items - 5 Items	100	99
Item Information		
Number of Items	80	80
Mean Observed Average Item Score (AIS)	0.67	0.47
Equated Mean Rasch B-Value	-0.22	-0.23
Mean R-biserial	0.57	0.45
SD R-biserial	0.10	0.09
Examinee Information		
Mathematics-Only Examinees		
Number of Examinees	32555	20392
Mean Scale Score	347	343
SD Scale Score	24	25
Median Scale Score	345	341
ELA and Mathematics Examinees		
Number of Examinees	339149	22154
Mean Scale Score	381	341
SD Scale Score	39	31
Median Scale Score	380	334

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 demonstrate grade-level competency 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, 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 2008–09, 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 2008 and February, March, and May 2009.

Test Purpose

The primary 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 grade-level competency in reading, writing, and mathematics. The CAHSEE helps identify students who are not developing 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, 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 students with disabilities, meeting the exemption requirement pursuant to California *Education Code (EC)* Section 60852.3, or 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 students with disabilities 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 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 the 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 not necessary to use this form if test security was maintained.

The ELA examination measures reading and writing skills as defined by the State Board of Education (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 CR item. The CR item is a written response to a writing prompt.

The CAHSEE mathematics examination measures standards adopted by the SBE; that is grades 6 and 7 mathematics and Algebra I[2]. 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 10, 11, or 12 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 CAHSEE are used primarily to identify students who are not developing minimum grade level competencies in reading, writing, and mathematics that are contained in the California Content Standards. All California public school students, except eligible students with disabilities, 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 10 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 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 2008–09 school year on dates that were determined by the State Superintendent of Public Instruction (SSPI). Schools must administer the CAHSEE on the designated date, as shown in Table 1.1.

[2] The blueprints for CAHSEE ELA and 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: 2008–09

Administration	English-Language Arts (Tuesday)	Mathematics (Wednesday)
July 2008	July 29, 2008	July 30, 2008
October 2008	October 7, 2008	October 8, 2008
November 2008	November 4, 2008	November 5, 2008
December 2008	December 6, 2008*	December 13, 2008*
February 2009	February 3, 2009	February 4, 2009
March 2009	March 17, 2009	March 18, 2009
May 2009	May 12, 2009	May 13, 2009

* Saturday administration

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

- Grade 10 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 11 students may take the part(s) of CAHSEE not previously passed up to two times per school year and may test in consecutive administrations (e.g., October and November).
- Grade 12 students may take the part(s) of 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 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 2008–09 School Year

There were no new or unusual changes in examination content and test administration in the 2008–09 school year. The test blueprint remained unchanged as did the demographic groups that were evaluated for Adequate Yearly Progress (AYP). Equating and scoring methodologies were unchanged from the previous year.

Limitations of the Assessment

Score Interpretation

A school district uses 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 SBE is the state education agency that sets education policy for kindergarten through grade 12 in the areas of standards, instructional materials, assessment, and accountability. The SBE adopts textbooks for kindergarten through grade eight, adopts regulations to implement legislation, and has the authority to grant waivers of the *EC*. In 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 the California state'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 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 This Technical Report

This technical report describes the procedures applied to the CAHSEE for the four 2008 and the three 2009 administrations in the 2008–09 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 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 2008–09 administrations.

- Chapter 5 presents the processes involved in the actual 2008–09 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 2008–09 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 CAHSEE score use and interpretation.

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

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

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

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

Each chapter contains summary tables in the body of the text. Extended appendices that give more detailed information are provided at the 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 cycle for the California High School Exit Examination (CAHSEE). The specifications maintained by Educational Testing Service (ETS) to conduct each of those processes are described. This chapter is organized to provide a brief description of each process followed by a summary of the associated specifications. More details about the specifications and the analyses associated with each process are described in 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. Item writing emphasis 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 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 multiple-choice (MC) items. The ELA test also includes one constructed-response (CR) item, which is a writing prompt that is polytomously scored.

Model for Generating Item Statistics

Item response theory (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 English-language arts (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 (1PL) model 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 trait value 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 trait value 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 the analysis of the 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, as required by the CAHSEE program. 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 (i.e. February and March) administrations to obtain information about item performance and to obtain statistics that can be used to assemble operational forms. ETS prepares the 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. However,

only 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. Each year there are a number of items that are 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. There were no RTQs for 2008–09.

Criteria for Selecting RTQs

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' achievement of the content standards can be assessed. These released test questions, however, do not reflect all of 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 the account the factors described below.

Test Blueprint

ETS selects all CAHSEE test items to conform to the State Board of Education-approved California content standards and test blueprints. The ELA examination measures reading and writing skills as defined by the SBE standards through grade 10. 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 is comprised of 73 operational items and 7 field-test items for a total of 80 items. The mathematics test form is comprised of 80 operational items and 12 field-test items for a total of 92 items.

Number of Test Forms

In order to maintain the item pool for construction of future forms, multiple forms are administered, each containing a set of operational items and 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 item performance. 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 2008–09 Administrations

Administration	Operational Items	Field-Test Forms	Number of Field-Test Items[1]
July 2008			
English-Language Arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
October 2008			
English-Language Arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
November 2008			
English-Language Arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
December 2008			
English-Language Arts	72 MC, 1 CR	1	7
Mathematics	80 MC	1	12
February 2009			
English-Language Arts	72 MC, 1 CR	96	646
Mathematics	80 MC	37	444
March 2009			
English-Language Arts	72 MC, 1 CR	144	970
Mathematics	80 MC	56	672
May 2009			
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: ELA 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 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 administration 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 block 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 2008–09 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's 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 district CAHSEE coordinators, test examiners, proctors, and scribes. The responsibilities for each staff member are included in the *CAHSEE District and Test Site Coordinator's Manual* (CDE, 2008a), 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 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 Directions for Administration—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 *District and Test Site Coordinator's Manual* provides specific information and forms appropriate to the respective roles carried out by the Local Education Agency (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 – Other Asian Pacific Islander – Native Hawaiian – Guamanian – Samoan – Tahitian – Other Pacific Islander Filipino Hispanic or Latino African American White (not Hispanic)
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

Test Variations, Accommodations and Modifications

All grade 10, 11, and 12 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, including students with disabilities and English learners (EL). Many students with disabilities and 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. Students in these categories 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 may interfere with 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 "Far Below Basic" for API calculations and "Not Proficient" for AYP calculations.

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 grid 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 gridded for any portion of the CAHSEE are not applicable to that portion of the CAHSEE. For example, code "Q" (Calculators on mathematics test, i.e. CA) is considered a modification for mathematics and therefore cannot be gridded for the ELA portion.

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, scribes, spell checker, grammar checker or an English dictionary.

Table 2.4: CAHSEE Modification/Accommodation Table: 2008–09

A. Student marks in test booklet (other than responses) -*Cannot be gridded 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 gridded 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 gridded on the math portion of box 13, ELA side only.* ELA Accommodation	EO	Q. Calculators on mathematics test. - *Cannot be gridded 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 gridded on the math portion of box number 13, ELA side only.* ELA Accommodation	AN	R. Arithmetic table on mathematics test. - *Cannot be gridded 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 gridded 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 gridded for the math portion of box 13, ELA side only.* ELA Modification	SC
I. Extra time on test within a testing day. -*Cannot be gridded 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 gridded 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 gridded 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 2008–09 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 disabilities across the seven administrations ranged from 712 to 30,939 for ELA and 651 to 31,207 for mathematics. For students who tested with accommodations or modifications, the rates of achieving a score of 350 or higher ranged from 20 to 25 percent for ELA and from 19 to 25 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 districts whether each student with a passing score 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. While the accommodation group with the largest percentage of students passing varied across the seven administrations, students receiving the following accommodations had the highest passing percentage most often: test administered at home or hospital, responses dictated to a scribe for multiple-choice items, and responses in test booklet transferred to scannable answer document. 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, English learner) were directions read aloud or signed, additional breaks, and the use of a dictionary (ELA) or calculator (mathematics). English-only speakers 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. Also reported is passing status. 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 summing 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 summed 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 item response theory.

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 2008–09 items were calibrated and placed on the reference scale using a set of linking items selected from the 2008 forms and re-administered in 2009. The number correct or raw scores are converted to scale scores via true score equating. The raw-score to scale-score conversion reflects the relationship between the difficulty of individual test items comprising 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 2008

Table 2.A.1: Summary Statistics by Testing Variations and Disability—ELA, July 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	201	325	29	21
Modifications	1208	333	24	23
All	1409	332	25	23
Disability				
Mental Retardation	111	310	23	5
Hard of Hearing	38	328	27	21
Deaf	28	326	23	18
Speech or Language Impairment	120	334	23	24
Visual Impairment	-	-	-	-
Emotional Disturbance	210	331	30	28
Orthopedic Impairment	17	325	21	12
Other Health Impairment	149	329	24	18
Specific Learning Disability	2882	329	24	20
Deaf-Blindness	-	-	-	-
Multiple Disabilities	27	328	23	11
Autism	59	333	23	29
Traumatic Brain Injury	16	318	21	13

[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 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	221	330	25	18
Modifications	1687	333	21	20
All	1908	333	21	20
Disability				
Mental Retardation	112	319	21	8
Hard of Hearing	32	329	22	19
Deaf	-	-	-	-
Speech or Language Impairment	103	338	20	32
Visual Impairment	-	-	-	-
Emotional Disturbance	269	335	26	23
Orthopedic Impairment	27	330	19	11
Other Health Impairment	175	333	20	20
Specific Learning Disability	2876	331	21	17
Multiple Disabilities	22	331	31	18
Autism	44	335	27	30
Traumatic Brain Injury	19	322	16	5

[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 2008

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	2441	495	20	1946	80	330	56	49	51	37	48	1.9
Transfer of T/B Responses to A/D (TS)	19	7	37	12	63	338	53	50	60	46	57	1.9
Oral Responses Dictated to a Scribe (OR)	14	4	29	10	71	333	54	53	51	43	53	1.7
Spell Checker/Grammar Checker Off (SO)	35	12	34	23	66	338	64	53	58	38	53	1.9
Essay Responses (EO)	16	4	25	12	75	332	59	53	53	34	51	1.9
Assistive Device No Interference (AN)	11	3	27	8	73	336	62	56	60	40	52	1.8
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	113	29	26	84	74	328	54	48	50	38	44	1.9
Supervised Breaks (SB)	298	75	25	223	75	331	56	50	53	38	47	1.8
Beneficial Time (BT)	47	13	28	34	72	326	56	51	51	34	43	1.5
Tested At Home or Hospital (HH)	-	-		-	-	-	-	-	-	-	-	-
Dictionary (DI)	380	85	22	295	78	331	58	49	53	38	47	2.0
Sign Language (SL)	28	6	21	22	79	328	51	46	50	36	45	2.1
Oral Presentation (OP)	960	227	24	733	76	334	60	52	55	40	49	1.9
Spell Checker or Grammar Checker (SC)	156	42	27	114	73	337	60	53	55	41	49	2.1
Essay Responses (ER)	40	9	23	31	78	334	58	47	56	35	45	2.3
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-

[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 2.A.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, July 2008

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	2504	463	18	2041	82	332	44	46	41	40	33
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	12	5	42	7	58	338	55	65	40	39	29
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	90	31	34	59	66	337	48	52	43	42	38
Supervised Breaks (SB)	231	56	24	175	76	335	46	48	43	43	34
Beneficial Time (BT)	55	13	24	42	76	334	42	47	42	43	33
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	28	8	29	20	71	340	48	54	47	44	34
Sign Language (SL)	14	1	7	13	93	324	36	42	36	34	30
Oral Presentation (OP)	504	116	23	388	77	333	46	48	42	40	33
Calculator (CA)	1674	335	20	1339	80	333	45	48	41	41	33
Arithmetic Table (AT)	98	29	30	69	70	338	50	49	44	42	39
Math Manipulatives (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	292	311	329	344	372	399	330	24	2441
Transfer of T/B Responses to A/D (TS)	294	294	311	338	365	381	381	338	28	19
Oral Responses Dictated to a Scribe (OR)	292	292	298	325	363	390	390	333	34	14
Spell Checker/Grammar Checker Off (SO)	275	294	319	338	363	381	381	338	27	35
Essay Responses (EO)	282	282	313	338	349	384	384	332	29	16
Assistive Device No Interference (AN)	313	313	329	336	350	352	352	336	13	11
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	275	282	304	325	350	376	399	328	30	113
Supervised Breaks (SB)	275	286	308	329	350	387	426	331	31	298
Beneficial Time (BT)	275	275	296	321	356	387	417	326	37	47
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	286	293	313	330	346	375	399	331	24	380
Sign Language (SL)	275	292	314	324	345	367	393	328	25	28
Oral Presentation (OP)	286	298	317	332	348	376	409	334	24	960
Spell Checker or Grammar Checker (SC)	298	304	321	333	350	384	409	337	24	156
Essay Responses (ER)	275	297	319	329	344	383	409	334	25	40
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	291	302	317	329	345	369	399	332	21	2504
Transfer of T/B Responses to A/D (TS)	-	-	-	-	-	-	-	-	-	-
Oral Responses Dictated to a Scribe (OR)	296	296	315	337	364	381	381	338	28	12
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	275	291	317	334	358	396	438	337	31	90
Supervised Breaks (SB)	277	299	317	331	349	388	409	335	26	231
Beneficial Time (BT)	275	284	311	325	347	402	438	334	33	55
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	309	311	317	335	351	388	438	340	28	28
Sign Language (SL)	299	299	313	324	329	358	358	324	17	14
Oral Presentation (OP)	294	302	317	331	347	371	399	333	22	504
Calculator (CA)	294	304	319	331	345	371	396	333	21	1674
Arithmetic Table (AT)	302	306	317	330	354	391	412	338	26	98
Math Manipulatives (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	Emotional Disturbance	21	325	38	24
	Specific Learning Disability	19	336	36	37
Dictionary	Mental Retardation	20	315	20	0
	Other Health Impairment	14	326	20	21
	Specific Learning Disability	288	333	24	24
Essay Responses (EO)	Specific Learning Disability	11	334	25	27
Essay Responses (ER)	Specific Learning Disability	18	336	24	22
IEP or Section 504 plan	Mental Retardation	75	309	23	3
	Hard of Hearing	28	330	29	25
	Deaf	19	329	26	26
	Speech or Language Impairment	69	333	23	17
	Emotional Disturbance	131	330	29	25
	Orthopedic Impairment	13	325	23	15
	Other Health Impairment	92	330	24	20
	Specific Learning Disability	1698	330	24	20
	Autism	32	329	24	28
Oral Presentation	Mental Retardation	36	315	24	8
	Speech or Language Impairment	41	339	23	24
	Emotional Disturbance	25	336	24	12
	Other Health Impairment	36	332	22	19
	Specific Learning Disability	743	335	24	25
Sign Language	Hard of Hearing	12	323	25	17
	Deaf	14	331	27	29
Spell Checker or Grammar Checker	Specific Learning Disability	115	336	24	25
Spell Checker or Grammar Checker Off	Specific Learning Disability	24	338	27	33
Supervised Breaks	Mental Retardation	11	298	19	0
	Emotional Disturbance	49	331	37	27
	Other Health Impairment	14	326	25	21
	Specific Learning Disability	195	333	30	27
Test Over More Than One Day	Specific Learning Disability	78	331	30	29
Unlisted Accommodation	Specific Learning Disability	21	344	47	33

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	70	336	24	30
Beneficial Time	Emotional Disturbance	25	340	37	28
	Specific Learning Disability	20	328	24	15
Calculator	Mental Retardation	54	323	21	13
	Hard of Hearing	18	334	21	22
	Speech or Language Impairment	51	336	20	22
	Emotional Disturbance	89	336	24	26
	Orthopedic Impairment	15	330	19	7
	Other Health Impairment	80	335	21	24
	Specific Learning Disability	1194	334	21	21
	Autism	14	333	19	29
Dictionary for Math	Specific Learning Disability	19	346	31	32
IEP or Section 504 plan	Mental Retardation	75	319	21	9
	Hard of Hearing	25	330	23	20
	Speech or Language Impairment	63	335	20	22
	Emotional Disturbance	182	336	27	24
	Orthopedic Impairment	22	331	20	14
	Other Health Impairment	112	333	22	22
	Specific Learning Disability	1743	332	21	19
	Autism	26	328	22	23
Oral Presentation	Mental Retardation	25	321	17	4
	Speech or Language Impairment	16	341	21	44
	Emotional Disturbance	24	334	33	29
	Other Health Impairment	23	337	25	26
	Specific Learning Disability	386	334	21	24
Supervised Breaks	Emotional Disturbance	49	340	35	33
	Other Health Impairment	12	327	27	8
	Specific Learning Disability	140	335	22	22
Test Over More Than One Day	Emotional Disturbance	15	346	48	47
	Specific Learning Disability	58	335	25	29
Unlisted Accommodation	Specific Learning Disability	16	341	29	38

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	English Only	44	326	39	30
Dictionary	English Only	206	333	24	26
	Initially Fluent English Proficient	13	326	21	8
	Reclassified Fluent English Proficient	18	332	30	22
	English Learner	141	328	23	18
Essay Responses (EO)	English Only	11	338	21	27
Essay Responses (ER)	English Only	23	334	26	17
	English Learner	13	332	22	23
IEP or Section 504 plan	English Only	1229	333	26	25
	Initially Fluent English Proficient	72	332	21	19
	Reclassified Fluent English Proficient	81	334	28	27
	English Learner	1012	325	22	13
	Unknown	47	343	23	47
Oral Presentation	English Only	493	338	24	30
	Initially Fluent English Proficient	30	335	21	17
	Reclassified Fluent English Proficient	36	335	26	25
	English Learner	398	329	23	16
Sign Language	English Only	15	328	25	20
Spell Checker or Grammar Checker	English Only	88	340	26	34
	Reclassified Fluent English Proficient	12	345	24	33
	English Learner	52	328	18	12
Spell Checker or Grammar Checker Off	English Only	23	339	25	35
Supervised Breaks	English Only	187	333	33	28
	Reclassified Fluent English Proficient	15	330	32	27
	English Learner	83	324	25	16
Test Over More Than One Day	English Only	67	329	32	28
	English Learner	39	326	27	21
Transfer of Student T/B Responses to A/D	English Only	14	338	27	36
Unlisted Accommodation	English Only	18	340	45	28
	English Learner	14	325	41	14

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	60	336	24	25
	English Learner	29	333	24	24
Beneficial Time	English Only	47	336	35	28
Calculator	English Only	1000	335	21	22
	Initially Fluent English Proficient	52	338	21	19
	Reclassified Fluent English Proficient	54	335	19	20
	English Learner	537	330	20	16
	Unknown	31	335	18	16
Dictionary for Math	English Only	21	338	30	24
IEP or Section 504 plan	English Only	1428	334	22	21
	Initially Fluent English Proficient	76	336	22	17
	Reclassified Fluent English Proficient	78	336	18	23
	English Learner	877	328	20	14
	Unknown	45	338	19	18
Oral Presentation	English Only	302	334	23	24
	Initially Fluent English Proficient	12	338	30	25
	Reclassified Fluent English Proficient	20	333	20	30
	English Learner	165	332	20	20
Oral Responses Dictated to a Scribe	English Only	11	342	26	45
Supervised Breaks	English Only	155	336	28	26
	Reclassified Fluent English Proficient	14	337	13	21
	English Learner	56	330	21	13
Test Over More Than One Day	English Only	51	345	34	47
	English Learner	34	323	19	12
Unlisted Accommodation	English Only	16	341	23	38

[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 2008

Table 2.B.1: Summary Statistics by Testing Variations and Disability—ELA, October 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1172	327	25	19
Modifications	3022	332	24	24
All	4194	331	25	22
Disability				
Unknown	-	-	-	-
Mental Retardation	367	307	21	3
Hard of Hearing	125	326	26	17
Deaf	116	320	23	11
Speech or Language Impairment	473	330	22	20
Visual Impairment	33	330	25	30
Emotional Disturbance	814	329	31	26
Orthopedic Impairment	92	332	28	27
Other Health Impairment	608	336	27	31
Specific Learning Disability	9202	329	23	19
Deaf-Blindness	-	-	-	-
Multiple Disabilities	71	334	22	25
Autism	207	323	27	16
Traumatic Brain Injury	45	320	26	11

[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 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1022	330	23	18
Modifications	4984	336	23	24
All	6006	335	23	23
Disability				
Unknown	-	-	-	-
Mental Retardation	368	316	17	4
Hard of Hearing	107	332	21	16
Deaf	79	334	22	18
Speech or Language Impairment	458	334	23	23
Visual Impairment	34	338	32	29
Emotional Disturbance	911	330	25	20
Orthopedic Impairment	95	333	22	22
Other Health Impairment	684	337	25	26
Specific Learning Disability	9497	332	21	19
Multiple Disabilities	66	333	23	21
Autism	216	332	22	20
Traumatic Brain Injury	48	325	17	8

[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 2008

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	8034	1713	21	6321	79	330	56	53	52	37	47	1.8
Transfer of T/B Responses to A/D (TS)	54	14	26	40	74	336	62	58	55	42	50	1.9
Oral Responses Dictated to a Scribe (OR)	39	10	26	29	74	332	54	57	55	41	50	1.6
Spell Checker/Grammar Checker Off (SO)	76	28	37	48	63	342	67	62	61	42	50	2.1
Essay Responses (EO)	68	27	40	41	60	338	65	60	57	46	52	1.7
Assistive Device No Interference (AN)	34	13	38	21	62	341	63	65	60	40	55	1.9
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	30	7	23	23	77	331	61	55	53	39	47	1.7
Test Over More Than One Day (TD)	480	107	22	373	78	330	56	52	52	38	46	2.0
Supervised Breaks (SB)	1383	306	22	1077	78	330	56	53	52	38	47	1.9
Beneficial Time (BT)	252	37	15	215	85	325	54	50	49	34	42	1.8
Tested At Home or Hospital (HH)	18	6	33	12	67	333	64	55	56	40	46	1.8
Dictionary (DI)	875	214	24	661	76	333	59	56	54	38	49	1.9
Sign Language (SL)	47	12	26	35	74	330	54	48	55	39	49	1.8
Oral Presentation (OP)	2570	622	24	1948	76	333	58	55	55	39	48	1.9
Spell Checker or Grammar Checker (SC)	237	72	30	165	70	337	61	58	57	39	49	2.1
Essay Responses (ER)	68	25	37	43	63	339	59	58	56	39	51	2.3
Assistive Device (AD)	35	8	23	27	77	331	57	55	55	38	45	1.9
Unlisted Modification (UM)	15	2	13	13	87	316	52	44	39	34	43	1.5

[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

Table 2.B.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, October 2008

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	8567	1835	21	6732	79	333	49	51	42	38	31
Transfer of T/B Responses to A/D (TS)	43	13	30	30	70	336	53	54	43	41	30
Oral Responses Dictated to a Scribe (OR)	29	13	45	16	55	343	58	57	48	43	33
Braille Version (BV)	16	5	31	11	69	342	56	58	42	43	35
Large Print Version (LV)	28	10	36	18	64	345	57	57	49	48	35
Test Over More Than One Day (TD)	287	73	25	214	75	333	50	52	41	38	29
Supervised Breaks (SB)	1195	268	22	927	78	333	49	51	41	39	31
Beneficial Time (BT)	208	48	23	160	77	334	49	51	42	39	32
Tested At Home or Hospital (HH)	19	4	21	15	79	328	47	47	38	37	28
Dictionary for Math (DM)	99	32	32	67	68	341	54	58	47	44	31
Sign Language (SL)	33	8	24	25	76	337	48	57	46	40	31
Oral Presentation (OP)	1686	426	25	1260	75	336	51	53	42	41	32
Calculator (CA)	4927	1204	24	3723	76	336	50	55	42	40	31
Arithmetic Table (AT)	424	126	30	298	70	339	53	57	44	43	31
Math Manipulatives (MM)	33	10	30	23	70	342	54	55	45	45	36
Assistive Device (AD)	17	6	35	11	65	340	52	56	41	49	33
Unlisted Modification (UM)	30	11	37	19	63	338	50	51	44	44	38

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	290	313	329	346	370	393	330	25	8034
Transfer of T/B Responses to A/D (TS)	275	294	318	333	356	410	410	336	29	54
Oral Responses Dictated to a Scribe (OR)	275	275	309	329	350	399	414	332	33	39
Spell Checker/Grammar Checker Off (SO)	294	302	323	338	361	387	410	342	25	76
Essay Responses (EO)	275	275	315	337	362	393	414	338	33	68
Assistive Device No Interference (AN)	275	297	331	345	354	370	387	341	22	34
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	275	275	318	326	348	372	410	331	30	30
Test Over More Than One Day (TD)	275	287	313	329	348	374	403	330	26	480
Supervised Breaks (SB)	275	288	311	331	348	372	396	330	26	1383
Beneficial Time (BT)	275	284	307	323	338	367	387	325	24	252
Tested At Home or Hospital (HH)	290	290	306	330	356	410	410	333	32	18
Dictionary (DI)	282	295	316	331	348	382	403	333	25	875
Sign Language (SL)	275	290	313	327	350	370	374	330	25	47
Oral Presentation (OP)	280	295	316	333	348	377	393	333	24	2570
Spell Checker or Grammar Checker (SC)	290	302	322	335	352	377	390	337	23	237
Essay Responses (ER)	275	295	322	340	356	374	393	339	25	68
Assistive Device (AD)	286	288	313	331	346	379	379	331	24	35
Unlisted Modification (UM)	286	286	297	301	346	365	365	316	27	15

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	289	300	318	332	347	374	400	333	23	8567
Transfer of T/B Responses to A/D (TS)	295	300	312	334	358	389	395	336	29	43
Oral Responses Dictated to a Scribe (OR)	283	295	314	341	364	387	450	343	34	29
Braille Version (BV)	305	305	314	333	357	450	450	342	37	16
Large Print Version (LV)	300	303	325	347	367	389	410	345	29	28
Test Over More Than One Day (TD)	275	298	316	330	350	374	403	333	25	287
Supervised Breaks (SB)	289	300	316	330	347	380	410	333	24	1195
Beneficial Time (BT)	289	298	312	330	348	392	410	334	28	208
Tested At Home or Hospital (HH)	275	275	310	326	347	395	395	328	29	19
Dictionary for Math (DM)	298	308	323	338	356	387	410	341	24	99
Sign Language (SL)	295	298	326	340	349	369	374	337	19	33
Oral Presentation (OP)	295	303	318	332	350	385	414	336	25	1686
Calculator (CA)	292	303	320	334	349	378	400	336	23	4927
Arithmetic Table (AT)	298	308	320	336	352	385	400	339	24	424
Math Manipulatives (MM)	295	310	316	336	354	410	450	342	33	33
Assistive Device (AD)	308	308	324	334	356	382	382	340	22	17
Unlisted Modification (UM)	298	300	314	324	367	397	400	338	32	30

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	26	333	23	27
Assistive Device No Interference	Specific Learning Disability	23	345	22	43
Beneficial Time	Mental Retardation	13	301	16	0
	Emotional Disturbance	54	320	25	13
	Other Health Impairment	11	337	23	27
	Specific Learning Disability	136	325	23	13
Dictionary	Mental Retardation	36	318	26	14
	Hard of Hearing	12	337	30	33
	Speech or Language Impairment	19	332	16	16
	Emotional Disturbance	33	346	34	39
	Other Health Impairment	48	336	28	33
	Specific Learning Disability	604	333	24	24
	Autism	15	325	20	13
Essay Responses	Deaf	13	347	18	54
	Specific Learning Disability	41	344	31	41
	Specific Learning Disability	38	339	24	34
IEP or Section 504 plan	Mental Retardation	239	309	20	4
	Hard of Hearing	83	330	24	22
	Deaf	65	321	23	15
	Speech or Language Impairment	263	329	21	17
	Visual Impairment	27	332	26	37
	Emotional Disturbance	492	330	31	27
	Orthopedic Impairment	60	327	25	18
	Other Health Impairment	359	337	26	31
	Specific Learning Disability	5518	329	24	20
	Multiple Disabilities	23	329	26	22
	Autism	126	323	25	16
	Traumatic Brain Injury	24	322	31	17
Oral Presentation	Mental Retardation	103	311	23	7
	Hard of Hearing	25	332	25	20
	Speech or Language Impairment	103	333	22	23
	Emotional Disturbance	82	342	28	33
	Orthopedic Impairment	15	326	24	13
	Other Health Impairment	124	340	25	35
	Specific Learning Disability	1907	333	23	24
	Multiple Disabilities	13	323	31	15
	Autism	39	327	26	21
Oral Responses Dictated to a Scribe	Specific Learning Disability	19	339	31	32
Sign Language	Hard of Hearing	11	340	15	27
	Deaf	34	326	25	24
Spell Checker or Grammar Checker	Other Health Impairment	16	344	26	38
	Specific Learning Disability	158	336	23	30
Spell Checker or Grammar Checker Off	Specific Learning Disability	37	335	22	24

Table 2.B.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Supervised Breaks	Mental Retardation	64	309	21	5
	Hard of Hearing	12	337	33	42
	Speech or Language Impairment	43	329	21	21
	Emotional Disturbance	135	331	33	27
	Other Health Impairment	57	333	25	32
	Specific Learning Disability	847	331	24	21
	Autism	27	316	26	7
Test Over More Than One Day	Mental Retardation	23	308	23	9
	Speech or Language Impairment	16	333	19	31
	Emotional Disturbance	42	338	31	29
	Other Health Impairment	18	330	28	17
	Specific Learning Disability	319	330	25	22
	Autism	12	321	24	8
Transfer of Student T/B Responses to A/D	Visual Impairment	11	328	32	36
	Specific Learning Disability	14	333	16	7
Unlisted Accommodation	Emotional Disturbance	14	352	47	50
	Other Health Impairment	15	343	15	40
	Specific Learning Disability	156	331	25	18

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Mental Retardation	13	329	21	15
	Speech or Language Impairment	11	337	18	36
	Emotional Disturbance	24	334	19	25
	Other Health Impairment	37	345	25	32
	Specific Learning Disability	310	338	23	30
Assistive Device	Specific Learning Disability	14	344	21	43
Beneficial Time	Mental Retardation	12	323	26	17
	Emotional Disturbance	55	328	27	16
	Specific Learning Disability	95	331	22	18
Braille Version	Visual Impairment	12	348	39	33
Calculator	Mental Retardation	165	318	17	4
	Hard of Hearing	36	336	21	22
	Deaf	16	338	17	25
	Speech or Language Impairment	188	335	24	25
	Visual Impairment	13	354	35	46
	Emotional Disturbance	257	336	24	26
	Orthopedic Impairment	34	334	22	21
	Other Health Impairment	291	341	25	30
	Specific Learning Disability	3489	336	22	24
	Multiple Disabilities	19	336	25	21
	Autism	74	335	21	18
	Traumatic Brain Injury	14	328	16	7
Dictionary for Math	Other Health Impairment	15	357	24	60
	Specific Learning Disability	64	342	24	34
IEP or Section 504 plan	Mental Retardation	238	316	17	3
	Hard of Hearing	64	334	22	19
	Deaf	44	333	17	16
	Speech or Language Impairment	269	334	24	24
	Visual Impairment	28	343	32	36
	Emotional Disturbance	573	332	25	21
	Orthopedic Impairment	62	334	23	26
	Other Health Impairment	452	339	26	27
	Specific Learning Disability	5862	333	22	21
	Multiple Disabilities	25	337	23	20
	Autism	143	334	21	20
	Traumatic Brain Injury	25	325	18	8
Math Manipulatives	Specific Learning Disability	16	341	31	31
Oral Presentation	Mental Retardation	69	319	19	6
	Hard of Hearing	17	336	19	18
	Speech or Language Impairment	52	338	24	37
	Emotional Disturbance	60	340	26	35
	Other Health Impairment	99	348	27	37
	Specific Learning Disability	1190	335	24	23

Table 2.B.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Autism	31	339	23	26
Sign Language	Deaf	14	332	16	7
Supervised Breaks	Mental Retardation	56	321	20	7
	Speech or Language Impairment	37	330	20	19
	Emotional Disturbance	128	333	28	24
	Other Health Impairment	46	334	24	22
	Specific Learning Disability	727	333	22	21
	Autism	26	334	21	27
Test Over More Than One Day	Mental Retardation	15	326	25	20
	Emotional Disturbance	45	331	31	31
	Other Health Impairment	14	337	22	29
	Specific Learning Disability	174	332	23	22
Unlisted Accommodation	Emotional Disturbance	15	336	26	20
	Other Health Impairment	19	340	17	16
	Specific Learning Disability	161	331	22	17
Unlisted Modification	Specific Learning Disability	18	333	32	33

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	31	332	24	23
Assistive Device No Interference	English Only	27	344	20	44
Beneficial Time	English Only	194	325	24	15
	English Learner	33	329	21	15
	Unknown	17	305	23	6
Dictionary	English Only	522	335	27	29
	Initially Fluent English Proficient	14	342	18	36
	Reclassified Fluent English Proficient	27	346	24	33
	English Learner	300	328	22	17
	Unknown	12	322	22	0
Essay Responses	English Only	46	340	26	39
	English Only	57	338	34	39
	English Learner	19	336	21	26
IEP or Section 504 plan	English Only	4676	331	26	24
	Initially Fluent English Proficient	270	333	25	27
	Reclassified Fluent English Proficient	218	337	25	31
	English Learner	2758	326	22	15
	Unknown	112	319	24	11
Large Print Version	English Only	18	327	27	17
Oral Presentation	English Only	1484	335	25	28
	Initially Fluent English Proficient	78	337	24	29
	Reclassified Fluent English Proficient	72	340	22	31
	English Learner	899	328	22	18
	Unknown	37	313	18	0
Oral Responses Dictated to a Scribe	English Only	26	334	35	27
Sign Language	English Only	16	337	27	38
	English Learner	28	326	23	18
Spell Checker or Grammar Checker	English Only	155	340	22	37
	English Learner	69	330	21	14
Spell Checker or Grammar Checker Off	English Only	57	344	24	39
	English Learner	17	331	22	24
Supervised Breaks	English Only	951	332	27	26
	Initially Fluent English Proficient	48	335	21	27
	Reclassified Fluent English Proficient	29	329	23	17
	English Learner	328	325	21	12
	Unknown	27	315	25	7
Test Over More Than One Day	English Only	328	332	27	24
	Initially Fluent English Proficient	11	330	27	18
	English Learner	116	327	22	19

Table 2.B.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Unknown	15	303	22	7
Tested At Home or Hospital	English Only	12	337	36	42
Transfer of Student T/B Responses to A/D	English Only	33	338	28	27
	English Learner	17	332	33	18
Unlisted Accommodation	English Only	191	335	28	27
	English Learner	58	321	19	5

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Unknown	15	303	22	7
Tested At Home or Hospital	English Only	12	337	36	42
Transfer of Student T/B Responses to A/D	English Only	33	338	28	27
	English Learner	17	332	33	18
Unlisted Accommodation	English Only	191	335	28	27
	English Learner	58	321	19	5
Arithmetic Table	English Only	295	339	24	32
	Reclassified Fluent English Proficient	11	341	15	18
	English Learner	110	336	22	24
Assistive Device	English Only	15	338	22	33
Beneficial Time	English Only	173	336	28	25
	English Learner	17	336	28	24
	Unknown	15	316	17	7
Braille Version	English Only	11	352	40	45
Calculator	English Only	3120	337	24	27
	Initially Fluent English Proficient	154	337	22	25
	Reclassified Fluent English Proficient	141	343	22	31
	English Learner	1468	333	21	19
	Unknown	44	333	19	16
Dictionary for Math	English Only	81	342	25	36
	English Learner	17	332	17	12
IEP or Section 504 plan	English Only	5314	334	24	23
	Initially Fluent English Proficient	285	335	22	22
	Reclassified Fluent English Proficient	255	341	22	30
	English Learner	2599	331	21	17
	Unknown	114	322	19	7
Large Print Version	English Only	23	346	30	35
Math Manipulatives	English Only	23	349	37	43
Oral Presentation	English Only	1062	338	26	29
	Initially Fluent English Proficient	63	334	24	21
	Reclassified Fluent English Proficient	48	346	27	38
	English Learner	483	332	22	17
	Unknown	30	319	24	13
Oral Responses Dictated to a Scribe	English Only	19	343	41	53
Sign Language	English Only	16	337	21	25
	English Learner	15	337	15	20
Supervised Breaks	English Only	855	335	26	24
	Initially Fluent English Proficient	45	337	21	31
	Reclassified Fluent English Proficient	19	330	19	16
	English Learner	252	330	20	16
	Unknown	24	321	18	8
Test Over More Than One Day	English Only	200	334	26	28

Table 2.B.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	66	332	22	23
	Unknown	14	314	17	7
Tested At Home or Hospital	English Only	12	333	33	25
Transfer of Student T/B Responses to A/D	English Only	39	338	30	31
Unlisted Accommodation	English Only	192	337	26	24
	English Learner	59	328	16	10
Unlisted Modification	English Only	23	340	34	39

[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 2008

Table 2.C.1: Summary Statistics by Testing Variations and Disability—ELA, November 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	2929	327	30	22
Modifications	6880	329	27	24
All	9809	329	28	23
Disability				
Unknown	-	-	-	-
Mental Retardation	684	303	23	4
Hard of Hearing	299	324	28	19
Deaf	266	312	25	9
Speech or Language Impairment	1354	330	26	25
Visual Impairment	96	333	35	33
Emotional Disturbance	1867	329	36	29
Orthopedic Impairment	227	329	30	26
Other Health Impairment	1645	334	31	32
Specific Learning Disability	21462	327	27	21
Deaf-Blindness	-	-	-	-
Multiple Disabilities	66	319	32	15
Autism	685	334	31	30
Traumatic Brain Injury	128	323	31	20

[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 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	2723	331	24	20
Modifications	11307	335	22	24
All	14030	334	22	23
Disability				
Unknown	-	-	-	-
Mental Retardation	683	316	18	6
Hard of Hearing	249	335	23	25
Deaf	189	328	19	14
Speech or Language Impairment	1227	334	21	22
Visual Impairment	114	337	30	32
Emotional Disturbance	2087	332	26	22
Orthopedic Impairment	261	332	24	22
Other Health Impairment	1857	335	23	25
Specific Learning Disability	21904	332	21	20
Deaf-Blindness	-	-	-	-
Multiple Disabilities	61	327	23	10
Autism	640	336	25	25
Traumatic Brain Injury	134	332	25	20

[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 2008

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	20862	4791	23	16071	77	328	51	52	57	42	50	1.9
Transfer of T/B Responses to A/D (TS)	124	35	28	89	72	331	51	52	58	43	53	2.0
Oral Responses Dictated to a Scribe (OR)	83	35	42	48	58	342	58	61	62	52	55	2.2
Spell Checker/Grammar Checker Off (SO)	219	62	28	157	72	333	55	55	57	44	53	2.1
Essay Responses (EO)	123	43	35	80	65	334	54	56	57	47	52	2.0
Assistive Device No Interference (AN)	98	17	17	81	83	329	50	52	59	43	49	2.0
Braille Version (BV)	12	2	17	10	83	335	43	45	54	38	50	2.3
Large Print Version (LV)	76	23	30	53	70	330	51	54	56	42	50	2.0
Test Over More Than One Day (TD)	1171	307	26	864	74	331	52	54	58	44	51	1.9
Supervised Breaks (SB)	3366	765	23	2601	77	328	50	52	56	42	50	1.9
Beneficial Time (BT)	419	104	25	315	75	329	52	53	57	43	50	1.8
Tested At Home or Hospital (HH)	40	22	55	18	45	349	64	66	69	56	58	1.8
Dictionary (DI)	1961	445	23	1516	77	329	51	53	56	43	51	1.9
Sign Language (SL)	110	7	6	103	94	316	45	42	48	33	45	1.8
Oral Presentation (OP)	5879	1422	24	4457	76	330	51	54	58	44	51	1.9
Spell Checker or Grammar Checker (SC)	785	201	26	584	74	332	52	54	58	43	52	2.1
Essay Responses (ER)	130	40	31	90	69	338	55	55	60	46	52	2.4
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	68	25	37	43	63	335	56	55	60	45	55	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

Table 2.C.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, November 2008

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	21558	4616	21	16942	79	333	46	46	42	38	32
Transfer of T/B Responses to A/D (TS)	103	37	36	66	64	341	52	50	48	43	36
Oral Responses Dictated to a Scribe (OR)	55	19	35	36	65	339	49	52	46	39	36
Braille Version (BV)	25	8	32	17	68	344	52	56	48	42	39
Large Print Version (LV)	75	21	28	54	72	337	51	47	44	41	33
Test Over More Than One Day (TD)	702	153	22	549	78	333	47	46	43	38	33
Supervised Breaks (SB)	3014	694	23	2320	77	334	46	46	43	39	33
Beneficial Time (BT)	370	85	23	285	77	332	46	44	42	38	32
Tested At Home or Hospital (HH)	40	26	65	14	35	357	66	65	58	51	40
Dictionary for Math (DM)	338	96	28	242	72	338	50	52	44	41	34
Sign Language (SL)	107	23	21	84	79	331	43	44	40	35	36
Oral Presentation (OP)	3816	885	23	2931	77	334	47	48	42	39	33
Calculator (CA)	11228	2637	23	8591	77	335	47	49	43	39	32
Arithmetic Table (AT)	640	177	28	463	72	336	49	50	43	40	32
Math Manipulatives (MM)	132	58	44	74	56	343	53	56	48	42	37
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	52	11	21	41	79	332	46	47	42	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.C.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, November 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	284	307	328	348	374	398	328	28	20862
Transfer of T/B Responses to A/D (TS)	275	282	307	329	351	395	423	331	35	124
Oral Responses Dictated to a Scribe (OR)	275	284	312	340	374	398	414	342	35	83
Spell Checker/Grammar Checker Off (SO)	275	286	312	332	355	382	398	333	30	219
Essay Responses (EO)	278	284	307	334	357	395	423	334	35	123
Assistive Device No Interference (AN)	275	292	310	330	344	371	398	329	24	98
Braille Version (BV)	275	275	308	318	335	450	450	335	53	12
Large Print Version (LV)	282	288	304	332	352	385	423	330	31	76
Test Over More Than One Day (TD)	275	284	310	330	350	377	398	331	28	1171
Supervised Breaks (SB)	275	282	307	328	348	377	402	328	29	3366
Beneficial Time (BT)	275	275	305	330	348	382	410	329	31	419
Tested At Home or Hospital (HH)	275	275	319	359	380	414	423	349	44	40
Dictionary (DI)	275	286	307	330	348	374	395	329	27	1961
Sign Language (SL)	275	282	297	313	330	355	406	316	25	110
Oral Presentation (OP)	275	288	310	330	348	377	398	330	27	5879
Spell Checker or Grammar Checker (SC)	275	288	312	332	350	374	398	332	26	785
Essay Responses (ER)	278	295	314	334	352	395	429	338	33	130
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	275	276	311	338	359	385	398	335	32	68

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	291	302	316	330	346	372	398	333	22	21558
Transfer of T/B Responses to A/D (TS)	288	302	318	338	357	392	433	341	30	103
Oral Responses Dictated to a Scribe (OR)	291	294	314	330	357	407	427	339	32	55
Braille Version (BV)	291	297	314	338	369	428	448	344	42	25
Large Print Version (LV)	275	302	316	332	352	390	433	337	29	75
Test Over More Than One Day (TD)	282	302	318	332	346	372	398	333	22	702
Supervised Breaks (SB)	288	302	316	332	348	374	401	334	23	3014
Beneficial Time (BT)	282	300	316	330	346	374	401	332	24	370
Tested At Home or Hospital (HH)	275	305	336	359	377	403	407	357	31	40
Dictionary for Math (DM)	300	305	322	338	352	376	392	338	21	338
Sign Language (SL)	294	297	314	328	346	368	374	331	23	107
Oral Presentation (OP)	291	302	318	332	348	376	404	334	23	3816
Calculator (CA)	294	305	318	332	348	374	398	335	22	11228
Arithmetic Table (AT)	294	306	320	334	352	374	407	336	23	640
Math Manipulatives (MM)	288	305	324	342	358	385	433	343	26	132
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	291	297	317	331	348	372	378	332	20	52

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	Specific Learning Disability	71	332	24	20
Beneficial Time	Mental Retardation	15	307	19	0
	Emotional Disturbance	82	327	35	27
	Other Health Impairment	17	328	36	29
	Specific Learning Disability	255	331	28	25
	Autism	11	309	26	0
Dictionary	Mental Retardation	57	310	26	7
	Hard of Hearing	25	314	28	16
	Deaf	19	317	25	11
	Speech or Language Impairment	70	330	24	24
	Emotional Disturbance	47	332	34	28
	Other Health Impairment	87	334	27	29
	Specific Learning Disability	1427	329	27	23
	Autism	27	331	25	26
Essay Responses (EO)	Orthopedic Impairment	12	324	38	33
	Specific Learning Disability	54	332	27	33
Essay Responses (ER)	Deaf	21	323	27	5
	Specific Learning Disability	65	341	29	42
IEP or Section 504 plan	Mental Retardation	489	304	23	4
	Hard of Hearing	201	321	27	17
	Deaf	171	311	26	9
	Speech or Language Impairment	782	329	26	24
	Visual Impairment	74	330	36	27
	Emotional Disturbance	1171	329	35	28
	Orthopedic Impairment	155	329	30	28
	Other Health Impairment	1084	333	30	30
	Specific Learning Disability	13859	327	27	22
	Multiple Disabilities	47	318	34	15
	Autism	379	330	30	27
	Traumatic Brain Injury	77	322	31	18
Large Print Version	Visual Impairment	40	330	35	33
	Specific Learning Disability	14	328	22	21
Oral Presentation	Mental Retardation	199	310	24	6
	Hard of Hearing	48	318	25	8
	Speech or Language Impairment	273	331	27	25
	Visual Impairment	22	335	35	32
	Emotional Disturbance	175	331	31	26
	Orthopedic Impairment	45	323	27	20
	Other Health Impairment	280	336	27	31
	Specific Learning Disability	4246	331	27	25
	Multiple Disabilities	23	324	34	17
	Autism	94	332	28	24
	Traumatic Brain Injury	24	321	31	21

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Responses Dictated to a Scribe	Orthopedic Impairment	13	323	30	15
	Specific Learning Disability	31	350	29	55
Sign Language	Hard of Hearing	36	318	26	8
	Deaf	53	314	25	6
	Specific Learning Disability	12	316	24	8
Spell Checker or Grammar Checker	Mental Retardation	41	320	23	7
	Speech or Language Impairment	39	341	23	44
	Emotional Disturbance	21	330	29	38
	Other Health Impairment	43	340	25	37
	Specific Learning Disability	559	331	26	24
	Autism	18	336	27	33
Spell Checker or Grammar Checker Off	Emotional Disturbance	19	335	25	26
	Orthopedic Impairment	13	333	26	23
	Other Health Impairment	25	344	28	52
	Specific Learning Disability	105	332	28	29
Supervised Breaks	Mental Retardation	110	307	24	5
	Hard of Hearing	22	317	35	23
	Speech or Language Impairment	104	323	23	13
	Visual Impairment	14	327	31	36
	Emotional Disturbance	315	325	35	23
	Orthopedic Impairment	30	329	31	27
	Other Health Impairment	159	334	32	29
	Specific Learning Disability	2257	329	27	23
	Multiple Disabilities	15	311	32	7
	Autism	62	329	32	24
	Traumatic Brain Injury	19	329	33	26
Test Over More Than One Day	Mental Retardation	18	304	21	6
	Deaf	11	334	31	18
	Speech or Language Impairment	34	332	23	26
	Emotional Disturbance	90	330	32	31
	Orthopedic Impairment	12	327	36	25
	Other Health Impairment	61	342	26	46
	Specific Learning Disability	802	331	27	24
	Autism	24	331	39	33
Tested At Home or Hospital	Specific Learning Disability	13	345	37	54
Transfer of Student T/B Responses to A/D	Visual Impairment	28	328	39	29
	Specific Learning Disability	41	326	30	20
Unlisted Accommodation	Mental Retardation	19	303	19	0
	Speech or Language Impairment	15	331	24	20
	Emotional Disturbance	31	330	28	23
	Other Health Impairment	34	336	37	38
	Specific Learning Disability	339	322	26	12

Table 2.C.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Autism	13	313	28	15
Unlisted Modification	Specific Learning Disability	48	338	28	38

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Mental Retardation	30	326	20	13
	Speech or Language Impairment	13	342	19	38
	Emotional Disturbance	28	341	26	32
	Other Health Impairment	36	336	20	31
	Specific Learning Disability	450	337	23	29
Beneficial Time	Mental Retardation	12	315	8	0
	Emotional Disturbance	75	328	23	21
	Other Health Impairment	23	333	24	26
	Specific Learning Disability	207	333	23	24
	Autism	12	331	18	8
Braille Version	Visual Impairment	20	338	41	30
Calculator	Mental Retardation	267	321	19	9
	Hard of Hearing	81	336	23	27
	Deaf	30	334	22	23
	Speech or Language Impairment	443	336	21	23
	Visual Impairment	30	337	33	40
	Emotional Disturbance	489	335	25	26
	Orthopedic Impairment	105	335	24	19
	Other Health Impairment	654	337	21	26
	Specific Learning Disability	7992	335	22	23
	Multiple Disabilities	28	324	20	11
	Autism	208	337	21	28
	Traumatic Brain Injury	50	331	23	18
Dictionary for Math	Other Health Impairment	17	344	16	35
	Specific Learning Disability	223	338	21	28
IEP or Section 504 plan	Mental Retardation	481	317	18	7
	Hard of Hearing	170	333	22	22
	Deaf	118	328	20	14
	Speech or Language Impairment	740	334	21	21
	Visual Impairment	92	336	30	32
	Emotional Disturbance	1314	332	25	22
	Orthopedic Impairment	171	333	24	20
	Other Health Impairment	1239	335	22	23
	Specific Learning Disability	14541	333	22	21
	Multiple Disabilities	43	324	22	9
	Autism	385	335	22	24
	Traumatic Brain Injury	83	334	26	23
Large Print Version	Visual Impairment	35	337	30	34
	Specific Learning Disability	15	325	14	7
Math Manipulatives	Specific Learning Disability	91	345	25	47
Oral Presentation	Mental Retardation	149	322	21	12
	Hard of Hearing	30	334	26	20

Table 2.C.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Speech or Language Impairment	152	338	23	25
	Visual Impairment	21	349	35	57
	Emotional Disturbance	116	332	26	21
	Orthopedic Impairment	33	338	27	24
	Other Health Impairment	207	337	21	28
	Specific Learning Disability	2706	335	23	23
	Multiple Disabilities	11	334	26	27
	Autism	60	334	24	23
	Traumatic Brain Injury	18	333	22	17
Oral Responses Dictated to a Scribe	Orthopedic Impairment	13	344	38	38
Sign Language	Hard of Hearing	19	331	22	21
	Deaf	59	328	21	20
	Specific Learning Disability	18	335	20	22
Supervised Breaks	Mental Retardation	98	322	20	11
	Hard of Hearing	17	333	21	24
	Speech or Language Impairment	87	334	21	21
	Visual Impairment	21	341	38	38
	Emotional Disturbance	285	329	25	20
	Orthopedic Impairment	22	341	26	32
	Other Health Impairment	154	335	23	25
	Specific Learning Disability	1998	334	23	24
	Autism	57	331	18	12
	Traumatic Brain Injury	20	344	29	35
Test Over More Than One Day	Mental Retardation	15	320	21	7
	Speech or Language Impairment	20	334	20	20
	Emotional Disturbance	68	330	25	24
	Orthopedic Impairment	12	355	36	50
	Other Health Impairment	36	332	21	14
	Specific Learning Disability	461	333	22	21
	Autism	16	343	26	44
Tested At Home or Hospital	Specific Learning Disability	13	341	26	38
Transfer of Student T/B Responses to A/D	Visual Impairment	30	339	34	40
	Specific Learning Disability	30	341	27	33
Unlisted Accommodation	Mental Retardation	12	315	12	0
	Speech or Language Impairment	14	333	17	14
	Emotional Disturbance	30	331	24	17
	Other Health Impairment	34	335	24	26
	Specific Learning Disability	304	331	23	17
	Autism	16	329	21	13
Unlisted Modification	Specific Learning Disability	34	335	19	24

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device No Interference	English Only	58	330	26	21
	English Learner	33	324	23	12
Beneficial Time	English Only	279	331	32	27
	Reclassified Fluent English Proficient	16	331	29	31
	English Learner	116	322	27	16
Dictionary	English Only	1098	333	28	29
	Initially Fluent English Proficient	46	333	26	22
	Reclassified Fluent English Proficient	67	326	24	13
	English Learner	716	322	25	14
	Unknown	34	328	26	18
Essay Responses (EO)	English Only	85	339	36	38
	English Learner	30	324	35	30
Essay Responses (ER)	English Only	79	341	36	35
	English Learner	33	327	24	21
IEP or Section 504 plan	English Only	11674	331	30	27
	Initially Fluent English Proficient	729	329	28	23
	Reclassified Fluent English Proficient	845	335	26	31
	English Learner	7443	322	25	15
	Unknown	171	333	30	27
Large Print Version	English Only	36	336	33	39
	English Learner	29	319	29	21
Oral Presentation	English Only	3281	333	28	29
	Initially Fluent English Proficient	173	331	25	20
	Reclassified Fluent English Proficient	225	334	23	26
	English Learner	2159	325	26	17
	Unknown	41	325	29	15
Oral Responses Dictated to a Scribe	English Only	58	347	34	45
	English Learner	22	330	35	36
Sign Language	English Only	54	311	25	2
	English Learner	46	316	20	9
Spell Checker or Grammar Checker	English Only	406	335	27	31
	Initially Fluent English Proficient	17	332	27	24
	Reclassified Fluent English Proficient	33	336	24	27
	English Learner	309	326	25	19
	Unknown	20	331	25	15

Table 2.C.9 (continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Spell Checker or Grammar Checker Off	English Only	152	336	32	34
	English Learner	51	326	24	18
Supervised Breaks	English Only	1995	331	31	26
	Initially Fluent English Proficient	98	329	27	23
	Reclassified Fluent English Proficient	150	335	28	29
	English Learner	1091	323	25	16
	Unknown	32	322	30	19
Test Over More Than One Day	English Only	614	334	31	34
	Initially Fluent English Proficient	34	335	25	24
	Reclassified Fluent English Proficient	34	332	24	29
	English Learner	481	325	24	16
Tested At Home or Hospital	English Only	31	354	42	55
Transfer of Student T/B Responses to A/D	English Only	81	335	35	33
	English Learner	30	320	32	17
Unlisted Accommodation	English Only	306	328	30	21
	Initially Fluent English Proficient	14	328	25	21
	Reclassified Fluent English Proficient	11	331	28	27
	English Learner	174	313	22	4
Unlisted Modification	English Only	39	339	31	38
	English Learner	20	319	30	20

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	402	338	24	30
	Initially Fluent English Proficient	18	345	24	39
	Reclassified Fluent English Proficient	23	339	21	35
	English Learner	178	331	18	20
	Unknown	19	342	22	42
Beneficial Time	English Only	258	334	26	27
	English Learner	95	326	19	13
Braille Version	English Only	12	356	49	42
Calculator	English Only	6916	336	22	25
	Initially Fluent English Proficient	407	338	21	27
	Reclassified Fluent English Proficient	464	340	20	30
	English Learner	3337	332	21	18
	Unknown	104	335	22	23
Dictionary for Math	English Only	225	340	21	31
	English Learner	92	335	21	21
IEP or Section 504 plan	English Only	12872	334	23	23
	Initially Fluent English Proficient	784	336	22	24
	Reclassified Fluent English Proficient	904	339	22	30
	English Learner	6817	330	21	17
	Unknown	181	335	23	24
Large Print Version	English Only	44	343	33	36
	English Learner	19	326	24	11
Math Manipulatives	English Only	95	344	28	44
	English Learner	27	337	23	37
Oral Presentation	English Only	2244	336	24	26
	Initially Fluent English Proficient	102	332	21	14
	Reclassified Fluent English Proficient	152	336	22	24
	English Learner	1287	332	21	19
	Unknown	31	338	27	26
Oral Responses Dictated to a Scribe	English Only	40	338	29	38
Sign Language	English Only	64	331	25	23
	Initially Fluent English Proficient	13	323	20	8
	English Learner	28	331	18	18
Supervised Breaks	English Only	1829	335	24	25
	Initially Fluent English Proficient	92	337	25	27
	Reclassified Fluent English Proficient	134	338	21	32
	English Learner	935	331	21	17

Table 2.C.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Unknown	24	330	21	21
Test Over More Than One Day	English Only	353	335	25	24
	Initially Fluent English Proficient	32	338	24	34
	Reclassified Fluent English Proficient	20	339	25	40
	English Learner	292	331	19	17
Tested At Home or Hospital	English Only	29	360	31	69
Transfer of Student T/B Responses to A/D	English Only	76	341	31	36
	English Learner	18	339	24	33
Unlisted Accommodation	English Only	277	331	23	19
	Initially Fluent English Proficient	15	328	19	13
	English Learner	156	328	23	14
Unlisted Modification	English Only	38	333	21	29

[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 2008

Table 2.D.1: Summary Statistics by Testing Variations and Disability—ELA, December 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	99	323	26	13
Modifications	334	330	28	22
All	433	328	27	20
Disability				
Mental Retardation	36	307	21	0
Hard of Hearing	21	326	19	10
Deaf	20	329	18	10
Speech or Language Impairment	35	330	28	23
Visual Impairment	-	-	-	-
Emotional Disturbance	23	346	33	52
Orthopedic Impairment	-	-	-	-
Other Health Impairment	36	338	24	36
Specific Learning Disability	712	327	25	19
Multiple Disabilities	-	-	-	-
Autism	13	339	40	38
Traumatic Brain Injury	-	-	-	-

[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 2008

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	58	338	18	21
Modifications	383	336	18	21
All	441	336	18	21
Disability				
Mental Retardation	30	321	13	0
Hard of Hearing	19	336	18	11
Deaf	17	337	19	12
Speech or Language Impairment	33	335	17	27
Visual Impairment	-	-	-	-
Emotional Disturbance	29	338	12	17
Orthopedic Impairment	-	-	-	-
Other Health Impairment	36	338	21	25
Specific Learning Disability	651	336	19	20
Multiple Disabilities	-	-	-	-
Autism	19	335	19	21
Traumatic Brain Injury	-	-	-	-

[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 2008

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	564	111	20	453	80	327	58	50	53	41	49	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 Responses (EO)	-	-	-	-	-	-	-	-	-	-	-	-
Assistive Device No Interference (AN)	-	-	-	-	-	-	-	-	-	-	-	-
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	-	-	-	-	-	-	-	-	-	-	-	-
Test Over More Than One Day (TD)	40	4	10	36	90	316	50	44	47	36	43	1.9
Supervised Breaks (SB)	125	15	12	110	88	322	54	47	51	38	48	1.9
Beneficial Time (BT)	23	2	9	21	91	307	49	41	43	32	37	1.4
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	118	32	27	86	73	334	61	54	56	45	52	2.2
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	284	64	23	220	77	331	59	52	57	45	50	2.0
Spell Checker or Grammar Checker (SC)	89	10	11	79	89	322	51	45	52	40	46	2.0
Essay Responses (ER)	16	4	25	12	75	332	68	54	51	43	46	2.5
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-	-

[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

Table 2.D.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, December 2008

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	511	107	21	404	79	336	45	48	43	38	31
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)	15	5	33	10	67	338	48	52	42	41	31
Supervised Breaks (SB)	76	14	18	62	82	336	44	46	42	40	34
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-	-
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	129	28	22	101	78	337	46	49	42	39	31
Calculator (CA)	379	81	21	298	79	336	44	50	42	38	31
Arithmetic Table (AT)	37	5	14	32	86	333	40	50	41	34	32
Math Manipulatives (MM)	-	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	284	306	326	344	373	399	327	27	564
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	277	289	316	340	368	378	316	29	40
Supervised Breaks (SB)	275	280	304	324	340	359	378	322	25	125
Beneficial Time (BT)	275	275	280	302	336	354	366	307	28	23
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary (DI)	280	288	308	330	354	399	407	334	33	118
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	275	288	308	330	347	389	407	331	29	284
Spell Checker or Grammar Checker (SC)	275	282	304	322	342	352	407	322	24	89
Essay Responses (ER)	298	298	317	323	352	375	375	332	23	16
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	305	312	323	334	347	369	397	336	18	511
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)	312	312	325	336	352	379	379	338	18	15
Supervised Breaks (SB)	308	312	325	334	346	369	379	336	16	76
Beneficial Time (BT)	-	-	-	-	-	-	-	-	-	-
Tested At Home or Hospital (HH)	-	-	-	-	-	-	-	-	-	-
Dictionary for Math (DM)	-	-	-	-	-	-	-	-	-	-
Sign Language (SL)	-	-	-	-	-	-	-	-	-	-
Oral Presentation (OP)	308	312	323	336	349	363	397	337	18	129
Calculator (CA)	300	310	323	334	347	367	383	336	18	379
Arithmetic Table (AT)	308	312	321	333	340	363	371	333	15	37
Math Manipulatives (MM)	-	-	-	-	-	-	-	-	-	-
Assistive Device (AD)	-	-	-	-	-	-	-	-	-	-
Unlisted Modification (UM)	-	-	-	-	-	-	-	-	-	-

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	Specific Learning Disability	14	298	24	7
Dictionary	Mental Retardation	11	306	14	0
	Specific Learning Disability	85	336	33	29
IEP or Section 504 plan	Mental Retardation	30	309	21	0
	Hard of Hearing	14	326	22	14
	Deaf	15	328	19	7
	Speech or Language Impairment	17	326	32	24
	Emotional Disturbance	14	344	37	43
	Other Health Impairment	25	340	23	36
	Specific Learning Disability	413	327	27	20
Oral Presentation	Mental Retardation	19	312	19	0
	Other Health Impairment	18	343	23	39
	Specific Learning Disability	211	331	28	23
Spell Checker or Grammar Checker	Specific Learning Disability	60	321	22	10
Supervised Breaks	Specific Learning Disability	95	321	25	12
Test Over More Than One Day	Specific Learning Disability	32	314	30	13
Unlisted Accommodation	Specific Learning Disability	12	317	24	8

[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 2008

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Specific Learning Disability	30	336	15	17
Calculator	Mental Retardation	24	322	13	0
	Hard of Hearing	11	335	18	9
	Speech or Language Impairment	14	335	17	21
	Other Health Impairment	19	341	25	32
	Specific Learning Disability	273	337	17	23
IEP or Section 504 plan	Mental Retardation	24	323	12	0
	Hard of Hearing	13	337	19	15
	Deaf	16	336	19	13
	Speech or Language Impairment	17	333	16	18
	Emotional Disturbance	14	336	14	21
	Other Health Impairment	26	339	22	23
	Specific Learning Disability	360	338	18	23
Oral Presentation	Specific Learning Disability	94	338	18	24
Supervised Breaks	Specific Learning Disability	57	337	17	19
Unlisted Accommodation	Specific Learning Disability	15	336	19	27

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Beneficial Time	English Only	18	308	30	11
Dictionary	English Only	60	328	30	25
	English Learner	52	339	37	31
IEP or Section 504 plan	English Only	288	327	26	21
	Initially Fluent English Proficient	16	326	30	19
	Reclassified Fluent English Proficient	18	328	26	22
	English Learner	227	327	28	19
	Unknown	15	330	24	13
Oral Presentation	English Only	130	331	26	22
	English Learner	140	330	31	22
Spell Checker or Grammar Checker	English Only	41	327	24	15
	English Learner	45	316	24	9
Supervised Breaks	English Only	69	320	25	13
	English Learner	37	320	22	5
Test Over More Than One Day	English Only	27	316	31	11

[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 2008

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	26	332	15	8
	English Learner	11	338	14	27
Calculator	English Only	229	337	19	24
	English Learner	122	335	16	16
	Unknown	12	332	20	25
IEP or Section 504 plan	English Only	304	337	19	22
	Initially Fluent English Proficient	19	342	13	26
	Reclassified Fluent English Proficient	13	339	19	31
	English Learner	160	336	17	18
	Unknown	15	332	19	20
Oral Presentation	English Only	75	336	18	21
	English Learner	47	339	16	23
Supervised Breaks	English Only	42	336	15	14
	English Learner	20	332	17	10
Test Over More Than One Day	English Only	11	338	18	27
Unlisted Accommodation	English Only	14	345	29	36

[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 2009

Table 2.E.1: Summary Statistics by Testing Variations and Disability—ELA, February 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3721	332	31	26
Modifications	5476	332	26	24
All	9197	332	28	24
Disability				
Unknown	-	-	-	-
Mental Retardation	752	306	22	4
Hard of Hearing	304	334	32	25
Deaf	220	326	34	19
Speech or Language Impairment	1473	343	35	36
Visual Impairment	115	346	40	47
Emotional Disturbance	2230	335	37	31
Orthopedic Impairment	223	340	37	32
Other Health Impairment	1801	345	33	41
Specific Learning Disability	19879	332	27	24
Deaf-Blindness	-	-	-	-
Multiple Disabilities	88	327	28	18
Autism	707	348	41	44
Traumatic Brain Injury	142	329	32	21

[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 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	3292	336	28	26
Modifications	8724	337	23	25
All	12016	337	24	25
Disability				
Unknown	-	-	-	-
Mental Retardation	745	317	20	5
Hard of Hearing	265	344	30	32
Deaf	165	338	29	25
Speech or Language Impairment	1436	350	35	40
Visual Impairment	121	347	38	36
Emotional Disturbance	2494	338	31	28
Orthopedic Impairment	244	341	32	31
Other Health Impairment	1976	345	29	37
Specific Learning Disability	20230	337	24	25
Deaf-Blindness	-	-	-	-
Multiple Disabilities	80	332	26	20
Autism	689	352	39	42
Traumatic Brain Injury	135	337	28	25

[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 2009

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	19109	5147	27	13962	73	334	58	56	53	45	50	2.0
Transfer of T/B Responses to A/D (TS)	142	58	41	84	59	343	62	59	60	51	53	2.1
Oral Responses Dictated to a Scribe (OR)	100	46	46	54	54	349	70	64	63	54	54	2.1
Spell Checker/Grammar Checker Off (SO)	274	93	34	181	66	340	60	60	57	48	53	2.0
Essay Responses (EO)	145	67	46	78	54	347	67	63	60	56	55	2.1
Assistive Device No Interference (AN)	63	24	38	39	62	339	62	57	57	49	54	2.0
Braille Version (BV)	41	11	27	30	73	329	56	61	55	33	35	2.1
Large Print Version (LV)	70	28	40	42	60	342	61	58	57	50	53	2.0
Test Over More Than One Day (TD)	1334	344	26	990	74	333	56	55	54	45	51	2.0
Supervised Breaks (SB)	3540	922	26	2618	74	332	57	55	52	44	50	1.9
Beneficial Time (BT)	623	149	24	474	76	331	55	54	52	43	48	1.9
Tested At Home or Hospital (HH)	62	33	53	29	47	350	68	66	62	54	61	2.0
Dictionary (DI)	1715	411	24	1304	76	333	59	55	54	44	50	2.0
Sign Language (SL)	109	22	20	87	80	327	51	50	52	40	46	1.9
Oral Presentation (OP)	4494	1082	24	3412	76	333	58	56	53	45	49	1.9
Spell Checker or Grammar Checker (SC)	543	132	24	411	76	333	58	55	52	44	48	2.1
Essay Responses (ER)	133	47	35	86	65	341	60	58	56	49	53	2.3
Assistive Device (AD)	30	13	43	17	57	342	68	60	55	49	58	2.2
Unlisted Modification(UM)	68	18	26	50	74	331	60	55	50	43	50	1.9

[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

Table 2.E.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, February 2009

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	19732	5388	27	14344	73	338	49	52	46	43	33
Transfer of T/B Responses to A/D (TS)	104	37	36	67	64	346	55	56	50	47	37
Oral Responses Dictated to a Scribe (OR)	55	23	42	32	58	353	58	59	55	49	42
Braille Version (BV)	19	0	0	19	100	320	35	55	29	29	16
Large Print Version (LV)	70	25	36	45	64	345	52	53	51	47	35
Test Over More Than One Day (TD)	969	254	26	715	74	337	49	51	45	42	32
Supervised Breaks (SB)	3057	797	26	2260	74	337	49	51	45	41	33
Beneficial Time (BT)	533	136	26	397	74	335	49	50	43	39	31
Tested At Home or Hospital (HH)	53	30	57	23	43	360	63	62	59	58	41
Dictionary for Math (DM)	237	59	25	178	75	339	50	55	45	42	32
Sign Language (SL)	71	18	25	53	75	336	42	51	44	41	35
Oral Presentation (OP)	3209	853	27	2356	73	338	49	52	46	43	33
Calculator (CA)	7791	1877	24	5914	76	337	48	52	45	43	32
Arithmetic Table (AT)	681	208	31	473	69	341	52	55	46	45	34
Math Manipulatives (MM)	1101	313	28	788	72	340	50	54	47	45	33
Assistive Device (AD)	26	15	58	11	42	357	65	69	56	50	40
Unlisted Modification (UM)	56	16	29	40	71	345	48	58	49	49	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 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	289	312	331	351	390	416	334	30	19109
Transfer of T/B Responses to A/D (TS)	275	289	312	343	369	404	432	343	37	142
Oral Responses Dictated to a Scribe (OR)	275	289	323	346	375	416	444	349	37	100
Spell Checker/Grammar Checker Off (SO)	275	295	316	335	360	404	445	340	34	274
Essay Responses (EO)	275	291	323	345	369	416	445	347	36	145
Assistive Device No Interference (AN)	275	295	316	341	355	393	432	339	33	63
Braille Version (BV)	277	291	314	321	351	364	450	329	30	41
Large Print Version (LV)	275	287	308	334	369	426	445	342	42	70
Test Over More Than One Day (TD)	275	289	314	333	351	381	412	333	28	1334
Supervised Breaks (SB)	275	289	310	331	351	384	412	332	29	3540
Beneficial Time (BT)	275	283	308	329	349	387	421	331	31	623
Tested At Home or Hospital (HH)	275	275	325	354	374	407	445	350	37	62
Dictionary (DI)	279	295	316	331	349	379	400	333	25	1715
Sign Language (SL)	275	289	308	321	341	374	390	327	26	109
Oral Presentation (OP)	275	291	314	331	349	379	400	333	27	4494
Spell Checker or Grammar Checker (SC)	281	293	316	331	349	379	407	333	27	543
Essay Responses (ER)	279	299	319	341	360	393	412	341	29	133
Assistive Device (AD)	281	297	319	341	360	393	393	342	30	30
Unlisted Accommodation (UA)	275	295	308	321	351	379	421	331	29	68

[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 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	290	304	319	334	352	389	422	338	27	19732
Transfer of T/B Responses to A/D (TS)	301	306	319	340	365	439	450	346	36	104
Oral Responses Dictated to a Scribe (OR)	276	299	321	343	389	433	450	353	43	55
Braille Version (BV)	287	287	308	319	334	347	347	320	17	19
Large Print Version (LV)	287	308	319	333	364	427	450	345	36	70
Test Over More Than One Day (TD)	287	301	319	334	351	382	414	337	26	969
Supervised Breaks (SB)	287	304	317	333	351	387	418	337	26	3057
Beneficial Time (BT)	276	299	317	331	351	384	427	335	28	533
Tested At Home or Hospital (HH)	290	304	325	356	392	439	450	360	40	53
Dictionary for Math (DM)	293	308	323	336	349	382	414	339	23	237
Sign Language (SL)	293	301	315	331	351	397	418	336	27	71
Oral Presentation (OP)	293	304	321	334	351	384	410	338	24	3209
Calculator (CA)	296	306	321	334	349	378	403	337	22	7791
Arithmetic Table (AT)	296	308	325	338	354	382	406	341	23	681
Math Manipulatives (MM)	296	306	323	338	352	384	414	340	24	1101
Assistive Device (AD)	311	319	334	354	367	406	450	357	29	26
Unlisted Modification (UM)	290	296	323	335	365	410	450	345	35	56

[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 2009

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	19	335	21	32
Assistive Device No Interference	Specific Learning Disability	31	336	34	39
Beneficial Time	Mental Retardation	19	303	20	5
	Speech or Language Impairment	16	327	22	13
	Emotional Disturbance	184	332	36	29
	Other Health Impairment	42	335	27	33
	Specific Learning Disability	282	328	25	17
Braille Version	Visual Impairment	21	332	36	33
	Specific Learning Disability	13	329	22	23
Dictionary	Mental Retardation	54	320	26	11
	Hard of Hearing	18	323	27	22
	Deaf	12	321	17	0
	Speech or Language Impairment	50	331	25	20
	Emotional Disturbance	49	331	28	24
	Other Health Impairment	78	337	25	26
	Specific Learning Disability	1270	335	25	26
	Autism	29	335	32	31
Essay Responses	Hard of Hearing	14	331	20	21
	Deaf	26	333	30	23
	Emotional Disturbance	15	355	32	60
	Specific Learning Disability	58	343	28	34
	Specific Learning Disability	66	345	27	44
	Autism	11	344	57	45
IEP or Section 504 plan	Mental Retardation	523	308	22	5
	Hard of Hearing	175	331	31	23
	Deaf	111	325	33	18
	Speech or Language Impairment	675	334	27	24
	Visual Impairment	85	343	40	45
	Emotional Disturbance	1360	332	36	28
	Orthopedic Impairment	132	336	33	26
	Other Health Impairment	1007	342	32	37
	Specific Learning Disability	12020	331	26	23
	Multiple Disabilities	41	323	27	17
	Autism	406	343	39	41
	Traumatic Brain Injury	75	321	28	15
Large Print Version	Visual Impairment	39	348	45	44
	Specific Learning Disability	14	324	27	21
Oral Presentation	Mental Retardation	193	313	25	8
	Hard of Hearing	41	329	26	22
	Speech or Language Impairment	180	334	26	22
	Visual Impairment	13	334	38	38
	Emotional Disturbance	157	326	27	16
	Orthopedic Impairment	36	333	29	22
	Other Health Impairment	186	338	26	31

Table 2.E.7 (continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Specific Learning Disability	3314	334	26	25
	Multiple Disabilities	16	321	21	6
	Autism	96	337	34	34
	Traumatic Brain Injury	19	326	27	26
Oral Responses Dictated to a Scribe	Specific Learning Disability	49	342	29	41
Sign Language	Hard of Hearing	24	328	25	25
	Deaf	65	324	28	17
Spell Checker or Grammar Checker	Mental Retardation	41	315	20	5
	Speech or Language Impairment	17	339	31	29
	Emotional Disturbance	18	325	20	6
	Other Health Impairment	25	350	23	48
	Specific Learning Disability	372	333	25	25
	Autism	12	357	46	50
Spell Checker or Grammar Checker Off	Emotional Disturbance	21	337	26	33
	Other Health Impairment	15	362	28	67
	Specific Learning Disability	120	341	30	35
	Autism	12	338	46	33
Supervised Breaks	Mental Retardation	122	314	25	7
	Hard of Hearing	29	326	32	17
	Deaf	19	333	43	26
	Speech or Language Impairment	113	331	25	21
	Visual Impairment	23	345	42	52
	Emotional Disturbance	450	332	36	29
	Orthopedic Impairment	22	346	32	36
	Other Health Impairment	203	339	29	35
	Specific Learning Disability	2226	332	27	24
	Autism	82	345	41	45
	Traumatic Brain Injury	17	323	27	18
Test Over More Than One Day	Mental Retardation	39	312	25	8
	Deaf	14	326	38	14
	Speech or Language Impairment	52	331	32	25
	Emotional Disturbance	97	329	34	28
	Other Health Impairment	75	334	24	25
	Specific Learning Disability	928	334	26	25
	Autism	30	345	34	47
Tested At Home or Hospital	Specific Learning Disability	27	351	30	44
Transfer of Student T/B Responses to A/D	Visual Impairment	36	350	45	53
	Orthopedic Impairment	12	341	32	33
	Specific Learning Disability	44	339	29	30

Table 2.E.7 (continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Unlisted Accommodation	Mental Retardation	30	310	16	0
	Hard of Hearing	12	317	30	17
	Speech or Language Impairment	26	324	20	15
	Emotional Disturbance	32	327	27	19
	Orthopedic Impairment	15	324	23	13
	Other Health Impairment	41	342	33	39
	Specific Learning Disability	434	328	24	18
	Autism	17	335	33	35
Unlisted Modification	Specific Learning Disability	32	326	19	16

[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 2009

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Mental Retardation	16	328	17	19
	Speech or Language Impairment	19	345	25	37
	Emotional Disturbance	17	343	19	29
	Other Health Impairment	32	348	31	34
	Specific Learning Disability	522	341	22	30
Assistive Device	Specific Learning Disability	17	356	18	65
Beneficial Time	Mental Retardation	16	320	12	0
	Speech or Language Impairment	17	343	22	35
	Emotional Disturbance	150	332	33	24
	Other Health Impairment	37	339	21	27
	Specific Learning Disability	235	333	24	23
Braille Version	Visual Impairment	18	319	17	0
Calculator	Mental Retardation	246	321	19	6
	Hard of Hearing	58	335	19	10
	Deaf	39	332	19	18
	Speech or Language Impairment	305	338	22	26
	Visual Impairment	18	336	27	28
	Emotional Disturbance	447	337	24	26
	Orthopedic Impairment	55	339	22	25
	Other Health Impairment	428	341	23	31
	Specific Learning Disability	5519	337	22	24
	Multiple Disabilities	25	337	27	24
	Autism	139	340	26	27
	Traumatic Brain Injury	28	338	22	21
Dictionary for Math	Other Health Impairment	13	336	19	23
	Specific Learning Disability	179	338	22	23
IEP or Section 504 plan	Mental Retardation	518	318	20	6
	Hard of Hearing	149	338	27	21
	Deaf	82	333	24	24
	Speech or Language Impairment	664	340	26	27
	Visual Impairment	90	344	37	31
	Emotional Disturbance	1522	336	30	25
	Orthopedic Impairment	154	340	31	29
	Other Health Impairment	1145	343	27	34
	Specific Learning Disability	12477	336	23	24
	Multiple Disabilities	43	331	28	16
	Autism	401	347	35	36
	Traumatic Brain Injury	72	334	22	19
Large Print Version	Visual Impairment	37	347	40	35
	Specific Learning Disability	13	332	24	23
Math Manipulatives	Mental Retardation	33	325	18	12
	Speech or Language Impairment	38	337	16	18
	Emotional Disturbance	63	334	26	19

Table 2.E.8 (continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Other Health Impairment	71	342	21	30
	Specific Learning Disability	727	340	23	29
	Autism	19	337	20	21
Oral Presentation	Mental Retardation	144	324	20	9
	Hard of Hearing	27	334	19	15
	Speech or Language Impairment	104	336	22	20
	Visual Impairment	14	340	38	21
	Emotional Disturbance	129	337	25	27
	Orthopedic Impairment	39	343	22	31
	Other Health Impairment	170	339	23	27
	Specific Learning Disability	2274	339	24	28
	Autism	64	337	29	23
	Traumatic Brain Injury	16	340	24	25
Oral Responses Dictated to a Scribe	Orthopedic Impairment	13	365	50	46
	Specific Learning Disability	12	327	23	17
Sign Language	Hard of Hearing	15	330	17	7
	Deaf	38	330	21	21
Supervised Breaks	Mental Retardation	103	321	18	7
	Hard of Hearing	25	331	20	12
	Deaf	14	338	19	36
	Speech or Language Impairment	106	336	20	25
	Visual Impairment	22	349	39	36
	Emotional Disturbance	396	334	31	24
	Orthopedic Impairment	25	351	31	40
	Other Health Impairment	180	339	24	28
	Specific Learning Disability	1910	337	25	27
	Autism	71	346	37	31
	Traumatic Brain Injury	12	332	14	8
Test Over More Than One Day	Mental Retardation	24	330	22	13
	Speech or Language Impairment	35	336	25	20
	Emotional Disturbance	76	341	31	33
	Other Health Impairment	62	335	22	16
	Specific Learning Disability	655	336	24	27
	Autism	17	352	35	41
Tested At Home or Hospital	Specific Learning Disability	23	349	46	39
Transfer of Student T/B Responses to A/D	Visual Impairment	34	348	43	35
	Orthopedic Impairment	12	351	35	33
	Specific Learning Disability	23	340	19	30

Table 2.E.8 (continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Unlisted Accommodation	Mental Retardation	28	315	15	4
	Hard of Hearing	11	318	16	0
	Speech or Language Impairment	28	334	20	21
	Emotional Disturbance	37	330	25	14
	Orthopedic Impairment	13	335	27	15
	Other Health Impairment	47	343	28	32
	Specific Learning Disability	427	332	23	20
	Autism	13	347	32	38
Unlisted Modification	Specific Learning Disability	36	340	25	25

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	13	346	25	54
Assistive Device No Interference	English Only	43	346	35	44
	English Learner	16	322	22	13
Beneficial Time	English Only	467	333	33	27
	Initially Fluent English Proficient	14	334	39	21
	Reclassified Fluent English Proficient	12	338	28	25
	English Learner	126	322	23	13
Braille Version	English Only	27	329	33	22
Dictionary	English Only	883	335	26	27
	Initially Fluent English Proficient	58	339	30	33
	Reclassified Fluent English Proficient	57	339	27	32
	English Learner	699	330	24	19
	Unknown	18	336	28	33
Essay Responses	English Only	74	341	30	32
	English Only	101	353	38	54
	English Learner	35	334	27	29
	English Learner	47	340	27	36
IEP or Section 504 plan	English Only	11487	337	32	32
	Initially Fluent English Proficient	697	339	31	34
	Reclassified Fluent English Proficient	711	347	31	42
	English Learner	6044	326	24	15
	Unknown	170	328	29	23
Large Print Version	English Only	44	345	43	45
	English Learner	18	322	25	17
Oral Presentation	English Only	2423	334	27	27
	Initially Fluent English Proficient	132	339	24	32
	Reclassified Fluent English Proficient	125	341	27	31
	English Learner	1771	329	25	19
	Unknown	43	324	29	19
Oral Responses Dictated to a Scribe	English Only	77	354	38	53
	English Learner	21	330	29	24
Sign Language	English Only	60	330	29	23
	English Learner	44	321	22	14
Spell Checker or Grammar Checker	English Only	303	336	27	27
	Initially Fluent English Proficient	12	322	22	17
	Reclassified Fluent English Proficient	23	338	24	26
	English Learner	192	329	25	19
	Unknown	13	338	25	38
Spell Checker or Grammar Checker Off	English Only	168	349	36	46

Table 2.E.9 (continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	96	324	22	13
Supervised Breaks	English Only	2228	335	31	30
	Initially Fluent English Proficient	114	335	31	32
	Reclassified Fluent English Proficient	108	339	31	34
	English Learner	1070	326	25	16
	Unknown	20	330	34	25
Test Over More Than One Day	English Only	731	335	30	29
	Initially Fluent English Proficient	38	341	23	29
	Reclassified Fluent English Proficient	46	343	27	39
	English Learner	502	329	24	20
	Unknown	17	329	19	18
Tested At Home or Hospital	English Only	46	355	35	57
	English Learner	12	330	41	33
Transfer of Student T/B Responses to A/D	English Only	99	346	37	42
	English Learner	30	330	32	30
Unlisted Accommodation	English Only	445	334	29	27
	Initially Fluent English Proficient	22	334	35	23
	English Learner	182	319	20	8
Unlisted Modification	English Only	34	337	32	38
	English Learner	30	318	15	7

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	394	342	23	34
	Initially Fluent English Proficient	26	346	21	38
	Reclassified Fluent English Proficient	19	345	22	37
	English Learner	238	338	22	24
Assistive Device	English Only	14	360	35	57
Beneficial Time	English Only	398	335	28	26
	Initially Fluent English Proficient	14	339	24	43
	English Learner	109	332	22	19
Calculator	English Only	4705	338	23	27
	Initially Fluent English Proficient	248	339	23	27
	Reclassified Fluent English Proficient	251	343	22	31
	English Learner	2481	334	21	18
	Unknown	106	335	21	20
Dictionary for Math	English Only	132	344	24	31
	English Learner	92	331	19	14
IEP or Section 504 plan	English Only	12428	340	28	30
	Initially Fluent English Proficient	751	341	27	30
	Reclassified Fluent English Proficient	770	352	30	45
	English Learner	5615	333	22	19
	Unknown	168	335	23	23
Large Print Version	English Only	47	349	40	43
	English Learner	15	330	18	20
Math Manipulatives	English Only	731	341	25	31
	Initially Fluent English Proficient	36	338	24	22
	Reclassified Fluent English Proficient	35	357	29	57
	English Learner	291	336	21	20
Oral Presentation	English Only	1837	339	25	29
	Initially Fluent English Proficient	105	335	25	18
	Reclassified Fluent English Proficient	98	347	23	41
	English Learner	1137	335	24	23
	Unknown	32	337	27	28
Oral Responses Dictated to a Scribe	English Only	40	352	45	45
Sign Language	English Only	48	337	30	27
	English Learner	21	331	20	19
Supervised Breaks	English Only	2004	337	27	27
	Initially Fluent English Proficient	97	340	26	30
	Reclassified Fluent English Proficient	91	352	34	42
	English Learner	849	333	23	20
	Unknown	16	341	32	38
Test Over More Than One Day	English Only	516	339	29	29
	Initially Fluent English Proficient	34	338	15	24
	Reclassified Fluent English Proficient	35	345	27	46
	English Learner	372	334	21	22

Table 2.E.10 (continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	Unknown	12	331	26	8
Tested At Home or Hospital	English Only	38	363	42	61
Transfer of Student T/B Responses to A/D	English Only	72	347	36	38
	English Learner	19	332	25	21
Unlisted Accommodation	English Only	456	335	25	24
	Initially Fluent English Proficient	22	341	35	32
	English Learner	181	327	21	12
Unlisted Modification	English Only	32	342	28	25
	English Learner	22	346	44	27

[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 2009

Table 2.F.1: Summary Statistics by Testing Variations and Disability—ELA, March 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	4891	335	32	30
Modifications	6879	331	25	22
All	11770	332	28	25
Disability				
Mental Retardation	895	307	24	6
Hard of Hearing	483	341	36	34
Deaf	383	315	29	9
Speech or Language Impairment	2271	345	35	39
Visual Impairment	197	352	41	50
Emotional Disturbance	3030	336	39	36
Orthopedic Impairment	386	347	39	42
Other Health Impairment	3329	348	34	49
Specific Learning Disability	30939	332	27	26
Deaf-Blindness	-	-	-	-
Multiple Disabilities	136	327	31	21
Autism	1487	350	39	48
Traumatic Brain Injury	185	330	31	25

[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 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	4408	339	30	31
Modifications	11746	335	23	23
All	16154	336	25	25
Disability				
Mental Retardation	871	315	21	5
Hard of Hearing	446	350	36	41
Deaf	313	329	30	16
Speech or Language Impairment	2250	352	37	43
Visual Impairment	210	353	38	46
Emotional Disturbance	3221	338	32	30
Orthopedic Impairment	400	347	37	39
Other Health Impairment	3522	349	32	44
Specific Learning Disability	31207	338	26	28
Deaf-Blindness	-	-	-	-
Multiple Disabilities	133	333	31	22
Autism	1446	356	41	48
Traumatic Brain Injury	183	338	30	29

[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 2009

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	29206	8794	30	20412	70	335	58	53	59	46	52	2.0
Transfer of T/B Responses to A/D (TS)	211	93	44	118	56	347	62	58	64	52	57	2.1
Oral Responses Dictated to a Scribe (OR)	162	53	33	109	67	338	61	55	61	50	53	1.9
Spell Checker/Grammar Checker Off (SO)	325	148	46	177	54	349	63	60	65	53	59	2.2
Essay Responses (EO)	182	77	42	105	58	347	64	59	64	52	56	2.2
Assistive Device No Interference (AN)	157	59	38	98	62	344	63	58	65	49	56	2.1
Braille Version (BV)	29	13	45	16	55	357	68	62	62	56	65	2.4
Large Print Version (LV)	133	63	47	70	53	347	64	59	63	54	56	2.1
Test Over More Than One Day (TD)	1659	511	31	1148	69	336	57	53	58	47	52	2.0
Supervised Breaks (SB)	4426	1074	24	3352	76	331	56	50	56	44	49	1.9
Beneficial Time (BT)	603	132	22	471	78	327	55	48	54	41	45	1.8
Tested At Home or Hospital (HH)	64	37	58	27	42	359	72	66	72	56	65	2.2
Dictionary (DI)	2407	486	20	1921	80	330	58	49	55	43	48	2.0
Sign Language (SL)	138	20	14	118	86	322	50	40	49	39	48	1.8
Oral Presentation (OP)	5405	1226	23	4179	77	332	57	51	58	45	50	1.9
Spell Checker or Grammar Checker (SC)	657	153	23	504	77	333	57	51	57	44	50	2.1
Essay Responses (ER)	178	66	37	112	63	341	59	55	60	48	54	2.4
Assistive Device (AD)	32	10	31	22	69	338	57	55	59	43	51	2.2
Unlisted Modification (UM)	82	20	24	62	76	332	56	50	56	46	52	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

Table 2.F.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, March 2009

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	29684	9231	31	20453	69	340	51	52	47	44	37
Transfer of T/B Responses to A/D (TS)	180	68	38	112	62	347	53	55	51	47	39
Oral Responses Dictated to a Scribe (OR)	99	48	48	51	52	354	60	59	57	48	46
Braille Version (BV)	30	12	40	18	60	347	57	55	50	46	42
Large Print Version (LV)	133	59	44	74	56	353	58	58	56	51	42
Test Over More Than One Day (TD)	1105	353	32	752	68	339	51	52	46	43	36
Supervised Breaks (SB)	3949	1070	27	2879	73	336	49	50	44	41	34
Beneficial Time (BT)	570	133	23	437	77	334	49	48	41	41	33
Tested At Home or Hospital (HH)	62	31	50	31	50	356	63	61	56	53	45
Dictionary for Math (DM)	307	95	31	212	69	340	52	56	44	43	35
Sign Language (SL)	147	28	19	119	81	330	43	42	42	36	36
Oral Presentation (OP)	4033	1011	25	3022	75	336	49	51	43	41	34
Calculator (CA)	10680	2452	23	8228	77	335	48	52	43	40	33
Arithmetic Table (AT)	712	166	23	546	77	335	49	52	42	39	33
Math Manipulatives (MM)	1274	323	25	951	75	337	50	53	44	41	34
Assistive Device (AD)	18	9	50	9	50	362	62	67	60	55	46
Unlisted Modification (UM)	40	14	35	26	65	343	48	55	50	48	38

[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 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	289	314	334	355	389	419	335	31	29206
Transfer of T/B Responses to A/D (TS)	275	293	315	344	373	424	450	347	40	211
Oral Responses Dictated to a Scribe (OR)	275	289	314	337	359	399	432	338	33	162
Spell Checker/Grammar Checker Off (SO)	281	297	323	346	368	419	450	349	37	325
Essay Responses (EO)	285	295	317	342	370	410	450	347	37	182
Assistive Device No Interference (AN)	287	302	323	342	361	395	436	344	28	157
Braille Version (BV)	281	300	316	345	393	438	450	357	47	29
Large Print Version (LV)	275	279	317	344	370	424	450	347	42	133
Test Over More Than One Day (TD)	275	287	315	334	355	386	410	336	30	1659
Supervised Breaks (SB)	275	287	310	331	348	380	406	331	29	4426
Beneficial Time (BT)	275	277	302	325	346	375	395	327	30	603
Tested At Home or Hospital (HH)	275	299	325	361	388	436	450	359	42	64
Dictionary (DI)	275	291	312	331	346	370	395	330	24	2407
Sign Language (SL)	279	285	302	316	334	375	414	322	28	138
Oral Presentation (OP)	275	293	315	333	348	373	395	332	25	5405
Spell Checker or Grammar Checker (SC)	275	295	315	333	348	378	399	333	26	657
Essay Responses (ER)	275	302	323	338	359	392	419	341	27	178
Assistive Device (AD)	275	277	316	340	357	406	410	338	31	32
Unlisted Modification (UM)	275	289	314	334	348	373	424	332	28	82

[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 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	290	303	319	336	356	395	433	340	29	29684
Transfer of T/B Responses to A/D (TS)	275	303	317	336	369	430	450	347	38	180
Oral Responses Dictated to a Scribe (OR)	275	308	327	349	369	433	450	354	36	99
Braille Version (BV)	296	313	319	342	365	427	433	347	33	30
Large Print Version (LV)	290	301	323	341	378	448	450	353	40	133
Test Over More Than One Day (TD)	287	301	321	336	356	387	428	339	28	1105
Supervised Breaks (SB)	287	301	317	332	352	385	418	336	27	3949
Beneficial Time (BT)	280	298	317	328	349	387	423	334	28	570
Tested At Home or Hospital (HH)	293	301	328	352	387	414	450	356	38	62
Dictionary for Math (DM)	293	303	323	338	352	389	414	340	26	307
Sign Language (SL)	290	298	308	321	343	387	450	330	32	147
Oral Presentation (OP)	290	303	319	332	351	382	407	336	25	4033
Calculator (CA)	293	303	319	332	349	376	404	335	23	10680
Arithmetic Table (AT)	290	301	319	332	349	378	397	335	23	712
Math Manipulatives (MM)	293	303	321	334	351	389	414	337	26	1274
Assistive Device (AD)	314	314	334	349	374	440	440	362	36	18
Unlisted Modification (UM)	293	307	316	338	367	397	428	343	33	40

[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 2009

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	16	333	22	13
Assistive Device No Interference	Visual Impairment	11	358	31	55
	Other Health Impairment	12	368	23	83
	Specific Learning Disability	92	339	24	30
Beneficial Time	Mental Retardation	16	306	16	0
	Speech or Language Impairment	14	329	23	21
	Emotional Disturbance	136	323	33	26
	Other Health Impairment	24	333	38	25
	Specific Learning Disability	345	326	27	19
	Autism	23	339	39	30
Braille Version	Visual Impairment	28	355	45	43
Dictionary	Mental Retardation	97	318	22	11
	Hard of Hearing	22	323	22	5
	Deaf	25	323	20	8
	Speech or Language Impairment	96	331	25	21
	Emotional Disturbance	66	330	27	23
	Orthopedic Impairment	14	325	25	14
	Other Health Impairment	99	332	26	22
	Specific Learning Disability	1785	330	24	21
	Multiple Disabilities	11	323	16	9
	Autism	43	333	29	26
Essay Responses	Mental Retardation	11	334	46	18
	Deaf	17	328	19	18
	Emotional Disturbance	11	351	33	55
	Orthopedic Impairment	19	360	39	58
	Other Health Impairment	15	349	42	40
	Specific Learning Disability	104	346	27	46
	Specific Learning Disability	82	341	31	38
IEP or Section 504 plan	Mental Retardation	593	308	22	5
	Hard of Hearing	249	332	32	26
	Deaf	224	313	29	9
	Speech or Language Impairment	1008	333	27	25
	Visual Impairment	145	347	39	45
	Emotional Disturbance	1738	333	37	33
	Orthopedic Impairment	236	344	36	37
	Other Health Impairment	1873	345	33	46
	Specific Learning Disability	17924	331	26	24
	Multiple Disabilities	76	323	28	17
	Autism	739	345	38	42
	Traumatic Brain Injury	100	330	27	22
Large Print Version	Visual Impairment	60	344	42	40
	Specific Learning Disability	25	339	29	44

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Oral Presentation	Mental Retardation	231	316	22	8
	Hard of Hearing	40	333	25	28
	Speech or Language Impairment	214	331	22	20
	Visual Impairment	28	348	30	39
	Emotional Disturbance	172	332	27	22
	Orthopedic Impairment	50	334	29	24
	Other Health Impairment	250	334	26	26
	Specific Learning Disability	3990	333	24	23
	Multiple Disabilities	17	319	12	0
	Autism	108	333	29	27
	Traumatic Brain Injury	37	326	24	22
Oral Responses Dictated to a Scribe	Visual Impairment	13	345	45	31
	Orthopedic Impairment	25	347	37	44
	Specific Learning Disability	56	339	27	39
Sign Language	Hard of Hearing	30	321	22	13
	Deaf	90	316	21	7
Spell Checker or Grammar Checker	Mental Retardation	36	320	19	8
	Speech or Language Impairment	28	337	32	32
	Emotional Disturbance	27	331	26	22
	Orthopedic Impairment	11	338	26	27
	Other Health Impairment	37	341	34	35
	Specific Learning Disability	442	333	24	21
	Autism	18	340	39	39
Spell Checker or Grammar Checker Off	Speech or Language Impairment	12	354	33	58
	Emotional Disturbance	22	355	43	59
	Orthopedic Impairment	27	360	41	56
	Other Health Impairment	24	358	45	54
	Specific Learning Disability	123	339	29	37
	Autism	35	374	38	71
Supervised Breaks	Mental Retardation	154	309	21	5
	Hard of Hearing	34	334	31	35
	Speech or Language Impairment	167	333	24	23
	Visual Impairment	18	343	31	33
	Emotional Disturbance	413	330	36	31
	Orthopedic Impairment	38	344	34	45
	Other Health Impairment	253	345	34	43
	Specific Learning Disability	2958	330	26	22
	Multiple Disabilities	19	327	23	21
	Autism	106	335	36	33
	Traumatic Brain Injury	24	324	28	17

Table 2.F.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Test Over More Than One Day	Mental Retardation	44	317	31	7
	Hard of Hearing	18	329	33	33
	Deaf	17	312	30	6
	Speech or Language Impairment	70	338	29	31
	Visual Impairment	12	334	41	25
	Emotional Disturbance	93	333	37	33
	Orthopedic Impairment	22	341	37	41
	Other Health Impairment	95	356	32	60
	Specific Learning Disability	1163	334	27	28
	Autism	47	351	37	55
Tested At Home or Hospital	Specific Learning Disability	21	342	37	38
Transfer of Student T/B Responses to A/D	Speech or Language Impairment	13	324	26	15
	Visual Impairment	52	355	42	60
	Orthopedic Impairment	19	344	26	37
	Other Health Impairment	18	365	47	61
	Specific Learning Disability	48	332	30	31
	Autism	16	339	33	31
Unlisted Accommodation	Mental Retardation	27	312	21	4
	Speech or Language Impairment	19	329	26	26
	Emotional Disturbance	26	313	27	8
	Other Health Impairment	41	355	28	61
	Specific Learning Disability	365	331	28	24
	Autism	18	342	33	56
Unlisted Modification	Specific Learning Disability	42	332	22	19

[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 2009

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	Mental Retardation	33	322	20	9
	Speech or Language Impairment	30	331	23	17
	Emotional Disturbance	34	332	20	18
	Other Health Impairment	41	338	21	24
	Specific Learning Disability	485	335	23	25
	Autism	11	326	19	9
Beneficial Time	Mental Retardation	16	321	20	6
	Speech or Language Impairment	15	334	25	27
	Emotional Disturbance	134	324	25	12
	Other Health Impairment	27	338	24	26
	Specific Learning Disability	313	336	24	26
	Autism	21	352	46	48
Braille Version	Visual Impairment	29	345	32	38
Calculator	Mental Retardation	312	319	18	6
	Hard of Hearing	76	339	27	25
	Deaf	57	328	23	16
	Speech or Language Impairment	409	335	22	23
	Visual Impairment	26	342	24	35
	Emotional Disturbance	535	335	25	25
	Orthopedic Impairment	93	335	25	24
	Other Health Impairment	664	339	24	30
	Specific Learning Disability	7615	336	22	23
	Multiple Disabilities	26	327	15	8
	Autism	236	337	27	28
	Traumatic Brain Injury	47	328	23	15
Dictionary for Math	Speech or Language Impairment	14	325	16	7
	Specific Learning Disability	232	341	25	31
IEP or Section 504 plan	Mental Retardation	584	316	18	4
	Hard of Hearing	228	342	32	33
	Deaf	185	329	30	17
	Speech or Language Impairment	1000	339	28	29
	Visual Impairment	154	348	34	41
	Emotional Disturbance	1874	335	29	27
	Orthopedic Impairment	246	343	34	35
	Other Health Impairment	2019	347	31	41
	Specific Learning Disability	18286	337	25	26
	Multiple Disabilities	76	331	28	20
	Autism	751	350	38	43
	Traumatic Brain Injury	105	333	28	21
Large Print Version	Visual Impairment	58	349	35	36
	Specific Learning Disability	20	336	28	25
Math Manipulatives	Mental Retardation	37	316	18	5
	Speech or Language Impairment	31	333	25	16

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Emotional Disturbance	59	325	25	12
	Orthopedic Impairment	13	330	23	8
	Other Health Impairment	87	350	32	40
	Specific Learning Disability	859	337	24	25
	Autism	33	337	31	30
	Traumatic Brain Injury	11	321	26	9
Oral Presentation	Mental Retardation	159	322	21	9
	Hard of Hearing	35	342	27	34
	Speech or Language Impairment	161	338	25	27
	Visual Impairment	25	355	39	36
	Emotional Disturbance	147	333	25	21
	Orthopedic Impairment	40	341	30	30
	Other Health Impairment	184	337	24	26
	Specific Learning Disability	2924	336	24	26
	Autism	84	339	31	32
	Traumatic Brain Injury	35	332	31	26
Oral Responses Dictated to a Scribe	Visual Impairment	14	343	17	29
	Orthopedic Impairment	21	359	39	48
	Specific Learning Disability	27	337	25	33
Sign Language	Hard of Hearing	23	327	23	13
	Deaf	95	328	31	16
	Specific Learning Disability	15	327	21	20
Supervised Breaks	Mental Retardation	110	317	17	4
	Hard of Hearing	29	341	35	38
	Speech or Language Impairment	148	337	30	29
	Visual Impairment	19	346	28	47
	Emotional Disturbance	388	331	29	23
	Orthopedic Impairment	29	350	39	48
	Other Health Impairment	252	343	32	38
	Specific Learning Disability	2632	336	25	26
	Multiple Disabilities	15	325	19	13
	Autism	90	345	34	41
	Traumatic Brain Injury	19	338	30	37
Test Over More Than One Day	Mental Retardation	17	315	20	6
	Hard of Hearing	12	350	28	58
	Speech or Language Impairment	35	342	27	43
	Visual Impairment	11	332	30	18
	Emotional Disturbance	70	332	29	21
	Orthopedic Impairment	16	354	31	56
	Other Health Impairment	64	353	35	48
	Specific Learning Disability	776	338	25	29
	Autism	31	347	39	45

Table 2.F.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Tested At Home or Hospital	Specific Learning Disability	27	350	33	44
Transfer of Student T/B Responses to A/D	Visual Impairment	47	354	39	49
	Orthopedic Impairment	12	345	34	42
	Other Health Impairment	16	356	47	56
	Specific Learning Disability	44	337	23	25
	Autism	12	343	37	33
Unlisted Accommodation	Mental Retardation	22	316	15	5
	Speech or Language Impairment	16	341	24	38
	Emotional Disturbance	30	328	27	20
	Other Health Impairment	33	351	23	52
	Specific Learning Disability	322	334	23	22
	Autism	18	345	38	33
Unlisted Modification	Specific Learning Disability	17	346	31	41

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	18	339	34	28
	English Learner	11	330	26	18
Assistive Device No Interference	English Only	94	347	30	45
	Reclassified Fluent English Proficient	11	353	24	64
	English Learner	41	328	16	10
Beneficial Time	English Only	409	328	31	24
	Initially Fluent English Proficient	13	322	28	23
	Reclassified Fluent English Proficient	24	345	33	50
	English Learner	151	321	24	13
Braille Version	English Only	14	367	44	57
Dictionary	English Only	1266	331	26	23
	Initially Fluent English Proficient	65	335	25	28
	Reclassified Fluent English Proficient	77	334	23	31
	English Learner	972	327	21	14
	Unknown	27	336	29	44
Essay Responses	English Only	120	350	38	46
	English Only	100	345	29	41
	English Learner	59	336	23	29
	English Learner	48	330	27	27
IEP or Section 504 plan	English Only	16703	339	33	35
	Initially Fluent English Proficient	1025	342	32	41
	Reclassified Fluent English Proficient	1589	353	28	58
	English Learner	9725	326	23	15
	Unknown	164	334	27	29
Large Print Version	English Only	88	351	44	53
	English Learner	30	327	26	17
Oral Presentation	English Only	2940	333	26	25
	Initially Fluent English Proficient	151	337	24	31
	Reclassified Fluent English Proficient	186	338	23	32
	English Learner	2075	329	23	17
	Unknown	53	342	21	38
Oral Responses Dictated to a Scribe	English Only	102	340	35	32
	English Learner	51	334	26	27
Sign Language	English Only	80	326	30	20
	English Learner	49	313	15	2

Table 2.F.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Spell Checker or Grammar Checker	English Only	380	336	28	28
	Initially Fluent English Proficient	26	337	23	23
	Reclassified Fluent English Proficient	20	332	25	25
	English Learner	216	327	21	13
	Unknown	15	333	30	40
Spell Checker or Grammar Checker Off	English Only	221	354	39	52
	Reclassified Fluent English Proficient	11	373	37	64
	English Learner	77	330	20	16
Supervised Breaks	English Only	2611	333	31	29
	Initially Fluent English Proficient	121	339	26	32
	Reclassified Fluent English Proficient	197	339	26	35
	English Learner	1456	324	24	14
	Unknown	41	331	24	24
Test Over More Than One Day	English Only	970	338	32	34
	Initially Fluent English Proficient	73	339	29	40
	Reclassified Fluent English Proficient	57	350	27	54
	English Learner	531	330	24	21
	Unknown	28	329	26	18
Tested At Home or Hospital	English Only	46	368	42	65
Transfer of Student T/B Responses to A/D	English Only	157	350	41	48
	English Learner	37	326	24	16
Unlisted Accommodation	English Only	370	336	32	32
	Initially Fluent English Proficient	18	323	24	11
	Reclassified Fluent English Proficient	13	345	18	46
	English Learner	175	323	24	17
Unlisted Modification	English Only	44	333	31	27
	English Learner	32	326	23	13

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	424	335	23	25
	Initially Fluent English Proficient	15	347	28	33
	Reclassified Fluent English Proficient	19	347	35	32
	English Learner	246	332	21	19
Beneficial Time	English Only	394	334	28	23
	Initially Fluent English Proficient	13	348	43	31
	Reclassified Fluent English Proficient	22	357	34	55
	English Learner	135	330	22	20
Braille Version	English Only	17	350	35	41
Calculator	English Only	6486	336	24	25
	Initially Fluent English Proficient	349	337	24	25
	Reclassified Fluent English Proficient	397	342	23	32
	English Learner	3359	332	21	18
	Unknown	89	335	20	27
Dictionary for Math	English Only	161	343	28	39
	English Learner	135	335	22	22
IEP or Section 504 plan	English Only	17594	342	30	33
	Initially Fluent English Proficient	1054	345	32	37
	Reclassified Fluent English Proficient	1658	357	31	56
	English Learner	9214	334	24	22
	Unknown	164	335	22	24
Large Print Version	English Only	87	356	44	48
	Reclassified Fluent English Proficient	11	361	38	55
	English Learner	29	338	22	28
Math Manipulatives	English Only	812	338	27	25
	Initially Fluent English Proficient	48	343	28	29
	Reclassified Fluent English Proficient	49	354	29	55
	English Learner	343	333	22	22
	Unknown	22	332	16	18
Oral Presentation	English Only	2307	336	25	25
	Initially Fluent English Proficient	105	335	25	26
	Reclassified Fluent English Proficient	135	344	27	37
	English Learner	1458	335	23	24
	Unknown	28	341	19	39
Oral Responses Dictated to a Scribe	English Only	66	356	35	52
	English Learner	26	339	30	38
Sign Language	English Only	74	331	32	23
	Initially Fluent English Proficient	20	326	36	15

Table 2.F.10 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	49	325	19	10
Supervised Breaks	English Only	2399	337	28	28
	Initially Fluent English Proficient	107	343	30	37
	Reclassified Fluent English Proficient	168	349	30	44
	English Learner	1232	332	23	22
	Unknown	43	331	22	19
Test Over More Than One Day	English Only	639	341	29	35
	Initially Fluent English Proficient	34	342	35	41
	Reclassified Fluent English Proficient	54	354	30	56
	English Learner	359	333	22	22
	Unknown	19	340	37	21
Tested At Home or Hospital	English Only	43	360	40	56
Transfer of Student T/B Responses to A/D	English Only	137	349	40	39
	English Learner	27	328	26	22
Unlisted Accommodation	English Only	334	338	28	28
	Initially Fluent English Proficient	23	329	17	13
	Reclassified Fluent English Proficient	15	362	26	67
	English Learner	153	329	22	19
Unlisted Modification	English Only	26	342	28	31

[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 2009

Table 2.G.1: Summary Statistics by Testing Variations and Disability—ELA, May 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1387	325	28	17
Modifications	3664	330	26	21
All	5051	329	27	20
Disability				
Unknown	39	374	40	74
Mental Retardation	397	306	22	4
Hard of Hearing	129	320	23	8
Deaf	164	311	22	4
Speech or Language Impairment	589	329	26	18
Visual Impairment	46	334	36	26
Emotional Disturbance	942	326	34	22
Orthopedic Impairment	106	329	32	18
Other Health Impairment	713	332	29	25
Specific Learning Disability	9017	327	26	17
Deaf-Blindness	-	-	-	-
Multiple Disabilities	252	349	34	42
Autism	361	338	34	32
Traumatic Brain Injury	121	337	30	33

[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 2009

Testing Variations	N[1]	Mean	SD[2]	Percent (≥350)
Accommodations	1181	330	23	17
Modifications	5432	334	22	19
All	6613	333	22	19
Disability				
Unknown	25	365	45	64
Mental Retardation	404	319	20	6
Hard of Hearing	110	332	19	15
Deaf	108	331	21	16
Speech or Language Impairment	518	333	22	15
Visual Impairment	47	336	30	23
Emotional Disturbance	1141	331	26	21
Orthopedic Impairment	119	334	23	18
Other Health Impairment	826	334	24	20
Specific Learning Disability	9131	331	21	17
Multiple Disabilities	165	347	33	36
Autism	314	338	29	26
Traumatic Brain Injury	103	337	27	28

[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 2009

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
							RW	RC	RL	WS	WC	Essay
IEP or Section 504 plan	8170	1461	18	6709	82	327	60	50	52	38	48	1.9
Transfer of T/B Responses to A/D (TS)	69	13	19	56	81	326	56	51	48	38	49	2.0
Oral Responses Dictated to a Scribe (OR)	50	14	28	36	72	335	68	59	58	43	47	1.9
Spell Checker/Grammar Checker Off (SO)	171	41	24	130	76	333	67	52	54	40	51	2.1
Essay Responses (EO)	76	24	32	52	68	335	63	54	57	43	51	2.1
Assistive Device No Interference (AN)	87	11	13	76	87	329	60	50	53	38	52	2.1
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	36	9	25	27	75	333	62	55	55	43	49	2.0
Test Over More Than One Day (TD)	641	156	24	485	76	332	63	52	56	42	50	2.0
Supervised Breaks (SB)	1551	317	20	1234	80	328	61	50	52	40	48	2.0
Beneficial Time (BT)	235	49	21	186	79	328	61	51	53	39	46	2.0
Tested At Home or Hospital (HH)	34	11	32	23	68	340	66	57	58	46	55	2.0
Dictionary (DI)	1153	253	22	900	78	332	67	53	55	41	52	2.0
Sign Language (SL)	60	5	8	55	92	317	49	42	49	30	46	1.8
Oral Presentation (OP)	3113	687	22	2426	78	331	62	53	55	41	51	2.0
Spell Checker or Grammar Checker (SC)	352	75	21	277	79	332	63	51	55	41	50	2.2
Essay Responses (ER)	95	28	29	67	71	339	62	54	58	44	54	2.4
Assistive Device (AD)	39	11	28	28	72	336	65	57	59	42	53	2.1
Unlisted Modification (UM)	38	6	16	32	84	326	64	47	51	36	50	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

Table 2.G.4: Demographic Summary for All Examinees by Testing Variations—Mathematics, May 2009

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	8623	1501	17	7122	83	332	46	44	41	40	33
Transfer of T/B Responses to A/D (TS)	63	14	22	49	78	332	46	44	41	41	35
Oral Responses Dictated to a Scribe (OR)	37	9	24	28	76	334	47	47	43	41	33
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	36	6	17	30	83	335	45	46	45	40	35
Test Over More Than One Day (TD)	383	85	22	298	78	336	50	48	45	42	34
Supervised Breaks (SB)	1363	250	18	1113	82	332	45	44	41	40	33
Beneficial Time (BT)	184	33	18	151	82	326	42	41	37	37	29
Tested At Home or Hospital (HH)	33	8	24	25	76	341	46	51	50	45	37
Dictionary for Math (DM)	192	32	17	160	83	333	48	47	42	39	31
Sign Language (SL)	50	11	22	39	78	332	47	42	43	41	33
Oral Presentation (OP)	1849	415	22	1434	78	335	48	46	44	43	35
Calculator (CA)	4854	907	19	3947	81	333	47	46	42	41	33
Arithmetic Table (AT)	487	105	22	382	78	336	48	48	44	43	35
Math Manipulatives (MM)	699	153	22	546	78	336	48	47	44	43	35
Assistive Device (AD)	23	4	17	19	83	333	47	46	42	42	31
Unlisted Modification (UM)	20	5	25	15	75	339	47	48	48	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 2.G.5: Percentiles of Scale Scores for Students with Testing Variations—ELA, May 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	275	285	307	326	343	374	400	327	27	8170
Transfer of T/B Responses to A/D (TS)	275	277	311	326	345	376	382	326	28	69
Oral Responses Dictated to a Scribe (OR)	275	289	312	334	355	393	417	335	31	50
Spell Checker/Grammar Checker Off (SO)	275	289	314	332	349	382	404	333	28	171
Essay Responses (EO)	275	277	316	335	358	384	422	335	31	76
Assistive Device No Interference (AN)	287	295	314	328	341	360	400	329	20	87
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	275	277	302	334	363	397	404	333	38	36
Test Over More Than One Day (TD)	275	291	314	332	349	379	408	332	27	641
Supervised Breaks (SB)	275	285	309	326	345	376	400	328	28	1551
Beneficial Time (BT)	275	283	307	328	347	379	422	328	30	235
Tested At Home or Hospital (HH)	275	275	318	335	355	422	450	340	41	34
Dictionary (DI)	279	293	316	332	347	376	400	332	25	1153
Sign Language (SL)	277	280	302	317	332	359	379	317	22	60
Oral Presentation (OP)	275	291	312	330	347	379	404	331	27	3113
Spell Checker or Grammar Checker (SC)	283	293	312	332	347	376	400	332	25	352
Essay Responses (ER)	289	297	320	332	355	390	417	339	27	95
Assistive Device (AD)	279	293	318	337	353	379	417	336	26	39
Unlisted Modification (UM)	291	293	309	321	345	362	393	326	23	38

[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 2009

Testing Variations	Percentiles							Mean Scale Score	SD[1]	N Tested[2]
	1	5	25	50	75	95	99			
IEP or Section 504 plan	291	302	317	329	343	372	404	332	22	8623
Transfer of T/B Responses to A/D (TS)	275	300	315	331	343	378	425	332	26	63
Oral Responses Dictated to a Scribe (OR)	275	297	315	329	348	378	415	334	28	37
Braille Version (BV)	-	-	-	-	-	-	-	-	-	-
Large Print Version (LV)	302	307	319	327	338	401	414	335	27	36
Test Over More Than One Day (TD)	285	304	319	333	347	390	425	336	26	383
Supervised Breaks (SB)	285	302	315	329	343	376	404	332	23	1363
Beneficial Time (BT)	275	291	311	321	342	366	408	326	25	184
Tested At Home or Hospital (HH)	288	307	319	334	345	408	450	341	33	33
Dictionary for Math (DM)	300	307	319	331	343	366	411	333	20	192
Sign Language (SL)	297	304	315	331	347	370	385	332	21	50
Oral Presentation (OP)	294	304	317	331	347	383	404	335	24	1849
Calculator (CA)	291	304	319	331	345	374	401	333	22	4854
Arithmetic Table (AT)	297	304	319	333	347	388	430	336	25	487
Math Manipulatives (MM)	288	304	319	333	347	383	411	336	24	699
Assistive Device (AD)	294	313	323	331	345	357	361	333	16	23
Unlisted Modification (UM)	302	303	321	333	346	403	404	339	30	20

[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 2009

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	Specific Learning Disability	29	339	26	31
Assistive Device No Interference	Specific Learning Disability	63	328	19	11
Beneficial Time	Emotional Disturbance	67	332	35	33
	Specific Learning Disability	114	328	23	18
Dictionary	Mental Retardation	32	315	25	9
	Hard of Hearing	11	331	29	9
	Deaf	16	318	21	6
	Speech or Language Impairment	47	328	25	13
	Emotional Disturbance	27	333	25	19
	Other Health Impairment	42	339	23	36
	Specific Learning Disability	888	334	25	23
	Autism	15	324	23	13
Essay Responses	Deaf	12	324	18	8
	Specific Learning Disability	55	342	28	35
	Specific Learning Disability	39	333	28	28
IEP or Section 504 plan	Mental Retardation	272	307	22	4
	Hard of Hearing	77	323	25	9
	Deaf	95	311	24	5
	Speech or Language Impairment	350	329	25	19
	Visual Impairment	38	336	37	29
	Emotional Disturbance	537	325	34	22
	Orthopedic Impairment	77	327	32	16
	Other Health Impairment	422	331	27	23
	Specific Learning Disability	5616	327	26	18
	Multiple Disabilities	24	317	24	13
	Autism	144	324	31	17
	Traumatic Brain Injury	43	321	29	19
Large Print Version	Visual Impairment	17	341	44	41
Oral Presentation	Mental Retardation	143	314	23	8
	Hard of Hearing	25	322	22	8
	Speech or Language Impairment	167	333	25	21
	Visual Impairment	13	328	32	15
	Emotional Disturbance	104	330	32	23
	Orthopedic Impairment	34	329	27	18
	Other Health Impairment	126	337	26	28
	Specific Learning Disability	2285	332	26	23
	Multiple Disabilities	11	324	25	18
	Autism	52	323	27	15
	Traumatic Brain Injury	15	319	26	7
Oral Responses Dictated to a Scribe	Specific Learning Disability	16	348	33	44
Sign Language	Hard of Hearing	11	322	29	9
	Deaf	44	312	18	2

Table 2.G.7 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
Spell Checker or Grammar Checker	Mental Retardation	20	315	17	0
	Speech or Language Impairment	24	325	17	4
	Other Health Impairment	13	328	24	8
	Specific Learning Disability	250	334	25	25
Spell Checker or Grammar Checker Off	Emotional Disturbance	24	326	26	21
	Specific Learning Disability	90	331	26	20
Supervised Breaks	Mental Retardation	48	312	24	6
	Hard of Hearing	16	329	28	13
	Deaf	12	315	21	0
	Speech or Language Impairment	53	326	24	15
	Visual Impairment	11	314	31	9
	Emotional Disturbance	156	322	29	19
	Orthopedic Impairment	13	328	31	15
	Other Health Impairment	77	331	28	25
	Specific Learning Disability	1033	329	28	22
	Autism	29	321	29	14
Test Over More Than One Day	Mental Retardation	19	316	26	5
	Speech or Language Impairment	25	331	20	20
	Emotional Disturbance	55	334	31	31
	Other Health Impairment	34	326	29	24
	Specific Learning Disability	439	333	26	25
	Autism	13	323	28	23
Tested At Home or Hospital	Specific Learning Disability	12	323	21	8
Transfer of Student T/B Responses to A/D	Specific Learning Disability	16	321	27	13
Unlisted Accommodation	Emotional Disturbance	13	340	29	46
	Other Health Impairment	13	310	19	0
	Specific Learning Disability	209	325	24	12
Unlisted Modification	Specific Learning Disability	30	325	20	17

[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 2009

Testing Variations	Disability	N[1]	Mean	SD [2]	Percent (≥350)
Arithmetic Table	Mental Retardation	17	331	21	12
	Speech or Language Impairment	24	339	32	25
	Emotional Disturbance	17	330	27	24
	Other Health Impairment	23	330	22	13
	Specific Learning Disability	358	337	26	22
Assistive Device	Specific Learning Disability	16	331	15	6
Beneficial Time	Emotional Disturbance	60	329	25	27
	Specific Learning Disability	79	328	25	18
Calculator	Mental Retardation	159	322	18	6
	Hard of Hearing	43	332	17	19
	Deaf	31	332	20	16
	Speech or Language Impairment	210	331	23	14
	Visual Impairment	17	332	15	12
	Emotional Disturbance	259	333	24	22
	Orthopedic Impairment	52	331	20	13
	Other Health Impairment	284	336	23	20
	Specific Learning Disability	3417	334	21	19
	Multiple Disabilities	12	343	23	42
	Autism	87	331	21	16
	Traumatic Brain Injury	25	322	19	0
Dictionary for Math	Specific Learning Disability	151	333	21	15
IEP or Section 504 plan	Mental Retardation	270	320	20	6
	Hard of Hearing	70	332	18	17
	Deaf	63	331	25	21
	Speech or Language Impairment	321	331	21	13
	Visual Impairment	38	337	32	26
	Emotional Disturbance	651	330	26	20
	Orthopedic Impairment	86	332	22	16
	Other Health Impairment	506	334	23	20
	Specific Learning Disability	5862	332	22	17
	Multiple Disabilities	23	336	23	26
	Autism	163	332	26	19
	Traumatic Brain Injury	44	325	22	11
Large Print Version	Visual Impairment	15	342	37	27
Math Manipulatives	Mental Retardation	26	325	20	8
	Speech or Language Impairment	26	336	17	15
	Emotional Disturbance	42	328	23	12
	Other Health Impairment	45	338	26	24
	Specific Learning Disability	479	337	24	23
Oral Presentation	Mental Retardation	83	321	18	7
	Hard of Hearing	13	331	18	15

Table 2.G.8 (Continued)

Testing Variations	Disability	N[1]	Mean	SD[2]	Percent (≥350)
	Speech or Language Impairment	73	329	24	8
	Emotional Disturbance	73	338	26	32
	Orthopedic Impairment	20	336	23	15
	Other Health Impairment	84	344	30	35
	Specific Learning Disability	1341	336	24	23
	Autism	36	331	22	17
Sign Language	Deaf	26	333	24	27
Supervised Breaks	Mental Retardation	45	319	17	7
	Speech or Language Impairment	47	331	22	15
	Emotional Disturbance	146	330	23	21
	Orthopedic Impairment	11	341	34	36
	Other Health Impairment	65	332	25	14
	Specific Learning Disability	895	333	24	20
	Autism	31	328	21	13
Test Over More Than One Day	Mental Retardation	11	326	23	9
	Speech or Language Impairment	15	330	24	13
	Emotional Disturbance	39	337	23	23
	Other Health Impairment	33	336	31	24
	Specific Learning Disability	234	339	26	25
Tested At Home or Hospital	Specific Learning Disability	17	334	22	12
Transfer of Student T/B Responses to A/D	Visual Impairment	12	330	22	17
	Specific Learning Disability	21	329	28	19
Unlisted Accommodation	Emotional Disturbance	14	328	23	14
	Other Health Impairment	12	333	15	8
	Specific Learning Disability	190	330	19	17
Unlisted Modification	Specific Learning Disability	15	340	28	27

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Assistive Device	English Only	28	343	21	32
Assistive Device No Interference	English Only	54	330	22	15
	English Learner	25	325	17	4
Beneficial Time	English Only	153	331	33	25
	English Learner	65	323	21	11
Dictionary	English Only	585	334	26	24
	Initially Fluent English Proficient	18	332	22	22
	Reclassified Fluent English Proficient	27	334	24	37
	English Learner	510	331	24	20
	Unknown	13	327	18	8
Essay Responses	English Only	50	341	29	26
	English Only	64	338	32	36
	English Learner	40	337	26	33
IEP or Section 504 plan	English Only	4520	328	29	20
	Initially Fluent English Proficient	210	329	28	23
	Reclassified Fluent English Proficient	240	331	25	24
	English Learner	3082	324	24	14
	Unknown	118	327	23	11
Large Print Version	English Only	18	330	44	28
	English Learner	13	330	30	15
Oral Presentation	English Only	1623	333	28	24
	Initially Fluent English Proficient	67	329	29	19
	Reclassified Fluent English Proficient	81	337	23	32
	English Learner	1299	329	25	19
	Unknown	43	330	26	14
Oral Responses Dictated to a Scribe	English Only	38	341	31	34
	English Learner	11	318	25	9
Sign Language	English Only	20	324	19	5
	English Learner	38	312	23	8
Spell Checker or Grammar Checker	English Only	204	331	24	19
	Reclassified Fluent English Proficient	19	328	21	16
	English Learner	122	333	26	25
Spell Checker or Grammar Checker Off	English Only	90	338	32	31
	English Learner	72	327	20	17
Supervised Breaks	English Only	847	329	29	23
	Initially Fluent English Proficient	43	328	28	19
	Reclassified Fluent English Proficient	57	332	25	28

Table 2.G.9 (Continued)

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
	English Learner	594	326	26	17
Test Over More Than One Day	English Only	310	334	28	27
	Initially Fluent English Proficient	19	336	37	21
	Reclassified Fluent English Proficient	27	333	29	33
	English Learner	283	328	25	21
Tested At Home or Hospital	English Only	17	356	48	53
	English Learner	12	324	31	17
Transfer of Student T/B Responses to A/D	English Only	43	323	30	19
	English Learner	21	327	22	10
Unlisted Accommodation	English Only	141	326	27	16
	English Learner	141	320	22	6
Unlisted Modification	English Only	17	328	29	18
	English Learner	17	324	18	12

[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 2009

Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Arithmetic Table	English Only	313	337	25	24
	Initially Fluent English Proficient	16	347	37	25
	Reclassified Fluent English Proficient	12	326	15	0
	English Learner	143	333	24	17
Assistive Device	English Only	21	333	17	19
Beneficial Time	English Only	133	328	26	21
	English Learner	35	319	19	9
Calculator	English Only	2896	334	22	19
	Initially Fluent English Proficient	150	337	25	23
	Reclassified Fluent English Proficient	161	337	22	26
	English Learner	1553	332	20	16
	Unknown	94	337	20	19
Dictionary for Math	English Only	124	335	20	19
	English Learner	60	329	17	12
IEP or Section 504 plan	English Only	5233	332	23	18
	Initially Fluent English Proficient	247	336	25	23
	Reclassified Fluent English Proficient	261	335	21	23
	English Learner	2737	330	20	15
	Unknown	145	332	19	14
Large Print Version	English Only	20	334	27	15
Math Manipulatives	English Only	478	336	24	22
	Initially Fluent English Proficient	20	346	27	35
	Reclassified Fluent English Proficient	20	338	16	20
	English Learner	168	334	25	22
	Unknown	13	329	11	0
Oral Presentation	English Only	1084	335	25	23
	Initially Fluent English Proficient	49	338	23	22
	Reclassified Fluent English Proficient	47	340	27	23
	English Learner	643	334	22	20
	Unknown	26	339	22	35
Oral Responses Dictated to a Scribe	English Only	26	336	30	27
Sign Language	English Only	27	328	22	15
	English Learner	21	337	20	33
Supervised Breaks	English Only	822	332	24	21
	Initially Fluent English Proficient	43	335	19	23
	Reclassified Fluent English Proficient	39	333	26	15
	English Learner	441	330	22	14
	Unknown	18	335	22	22
Test Over More Than One Day	English Only	231	338	27	25
	Initially Fluent English Proficient	12	350	39	42
	English Learner	128	334	25	16
Tested At Home or Hospital	English Only	22	347	36	27

Table 2.G.10 (Continued)

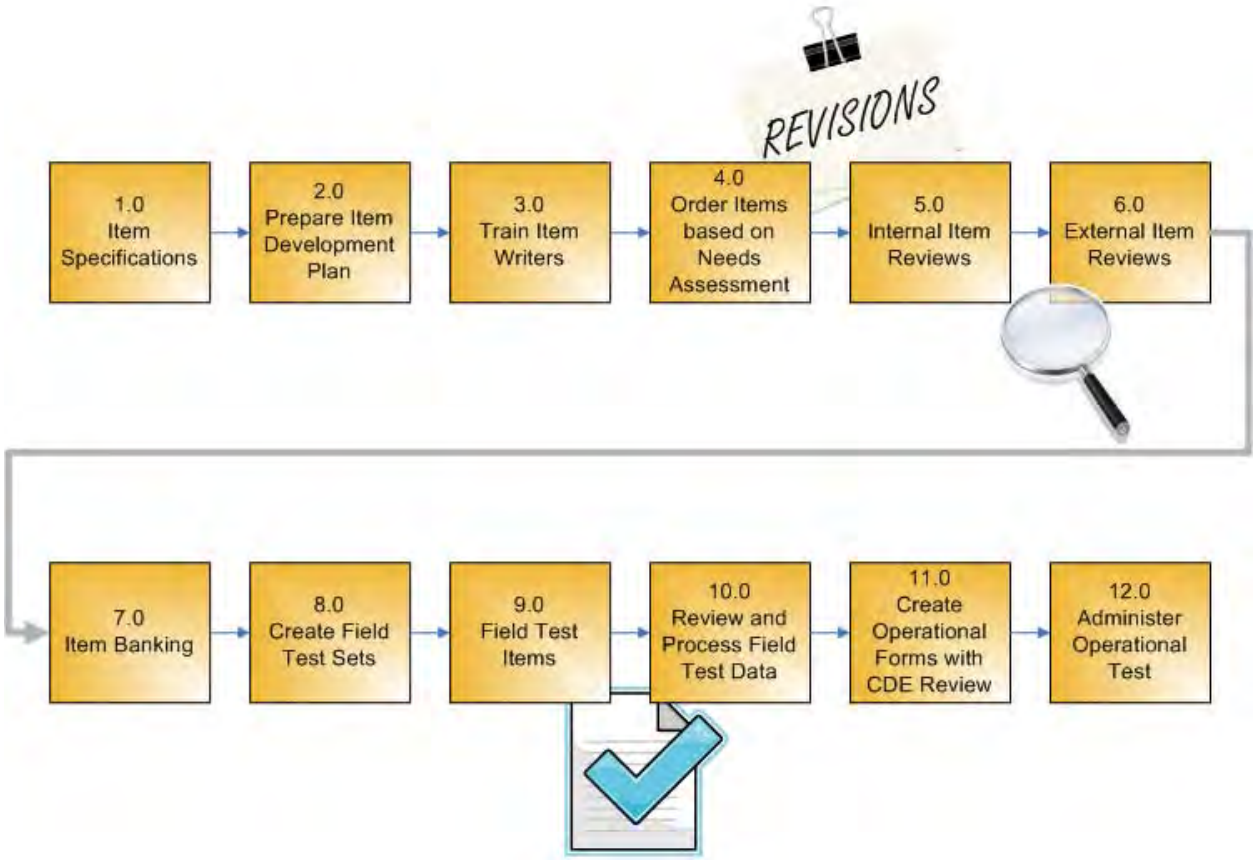
Testing Variations	Language Fluency	N[1]	Mean	SD[2]	Percent (≥350)
Transfer of Student T/B Responses to A/D	English Only	43	334	28	23
	English Learner	15	327	21	13
Unlisted Accommodation	English Only	134	329	23	15
	English Learner	119	330	19	16
Unlisted Modification	English Only	18	339	31	22

[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 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 in the CAHSEE measure the content standards in the same way. To accomplish this, the item specifications provide detailed information to item writers that 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, 2009).
- 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., items with a negative connotation in the stem, such as, “Which of the following is NOT…”).

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 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 blueprint is used to determine item needs and helps 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 test. 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 and mathematics RTQs were released.

The field-test percentages and item counts are shown in Table 3.1. The percentages are based on the ratio of the number of items to be field tested to the number of items on an operational test form.

Table 3.1 Field-test Percentages for the CAHSEE

Content Area	Percentage of Operational Form to Be Field Tested	Number of Items to Be Field Tested
English-–Language Arts	≈ 500%	1956 items
Mathematics	≈ 350%	1560 items

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 were screened by senior ETS content staff and approved by CDE staff. Only those with strong content and teaching backgrounds were approved for inclusion in the training. Thus, the participants were particularly experienced in writing test questions to the standards measured by the CAHSEE. All item writers met 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 and 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 CAHSEE ELA, passages must be written for both reading and language. Applicants for passage writing were screened by senior ETS content staff, and only those with strong writing skills were 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. Then, they 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's Guidelines for Bias and Sensitivity.
- A copy of the test item specifications document for the CAHSEE, which includes sample stems, sample items, and a checklist for item writers.
- An explanation of the depth-of-knowledge ratings.
- The Internet Web link to the previous Released Test Questions.

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

5.0 Internal Item Reviews (Educational Testing Service)

Purpose

The items selected for the CAHSEE undergo an extensive item review process that is designed to provide CDE 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 included the following stages:

1. Internal Content Review
2. Internal Editorial Review
3. Internal Bias and Sensitivity 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's 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 depth of knowledge
- 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/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/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

For the 2008-2009 development cycle, CAHSEE Content and Bias and Sensitivity Review committees were assembled to review newly developed items for accuracy of item content, clarity of phrasing, item quality, and fairness. The review panels were provided with the opportunity to review newly developed items and to make recommendations for the use of items in embedded field tests.

Composition

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 were recruited through an application process. Recommendations were solicited from districts and county offices of education in addition to CDE staff. Applications were received and reviewed throughout the year. Committee applications were reviewed by ETS Assessment Directors, who confirmed that the applicant's qualifications met the specified criteria. Applications that met the criteria were forwarded to the CDE for review and final approval. Upon approval, the applicant was notified that he or she had been selected to serve on the Content Review or the Bias and Sensitivity Review committee.

Bias and Sensitivity Review Meetings for Review of CAHSEE Items

The ETS test development manager facilitates the Bias and Sensitivity Review meetings. 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 principles of universal design.
- An overview of guidelines for reviewing CAHSEE items for potential issues of bias and sensitivity.

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

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. 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

The committee was also trained on how to make recommendations for revising items. 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?
- Does the stem or question include administrative directions 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 depth of knowledge 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 booklet.

ETS staff maintains the master binders showing committee recommendations. Any discrepancies are reconciled with CDE consultants.

Statewide Pupil Assessment Review

The SPAR is responsible for reviewing and approving a single achievement test to be used statewide for the testing of students in California public schools. At the SPAR panel meetings, all new items are presented in binders for review. The SPAR panel representatives ensure that the test items conform to the requirements of *Education Code* Section 60614. The CR writing tasks are also presented 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 in a nearby location or near a telephone to be available 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 content 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 the 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 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 three 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’s 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 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 December 2008 ELA field test was to try out a large number of writing prompts to augment the CAHSEE item bank. The results of the field-test analyses are not used to provide scores to students.

Samples of grade 10 students were selected to ensure that all geographic areas of the state were represented and that districts of various sizes, including charter schools, were included in the sample. Thirty-five forms were spiraled at the student level to ensure that each form was administered to randomly equivalent groups. Every

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

Embedded Field-Test Items

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

Allocation of Students to Field-Test Items

The operational test forms for 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 2009 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.

Table 3.2 Summary of Items and Forms Presented in 2009 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	199,809	195,091	96	646	1,994-4,647	195,091
	March	73	378,401	372,707	144	970	2,407-6,666	372,707
Mathematics	**February**	80	204,325	196,387	37	444	5,867-10,715	196,387
	March	80	382,066	371,704	56	672	5,389-7,965	371,704

[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.

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 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 student 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 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.

11.0 Create Operational Forms with California Department of Education Approval

Forms are created using the California Item Bank by ETS based on the CAHSEE blueprints and meeting statistical requirements given by ETS psychometricians. These are reviewed internally by ETS for content, by ETS psychometricians for statistical considerations, and by editorial staff before being posted to CDE for their review. These forms are posted to and reviewed by 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 include July, October, November, December, February, March, and May. The census administrations, consisting primarily of grade 10 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 12 and adult education students.

Chapter 4: Test Development

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 examinees' 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 number 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 includes 72 operational MC items, one operational CR item, and seven embedded field-test MC items. Each mathematics form consists of 92 items, which includes 80 operational MC items and 12 field-test MC items. For more details on the distribution of items across seven 2008–09 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 call for 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 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, and that some items 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 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 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 for approval to the psychometricians.

Figure 4.1: Example of Test Characteristic Curves

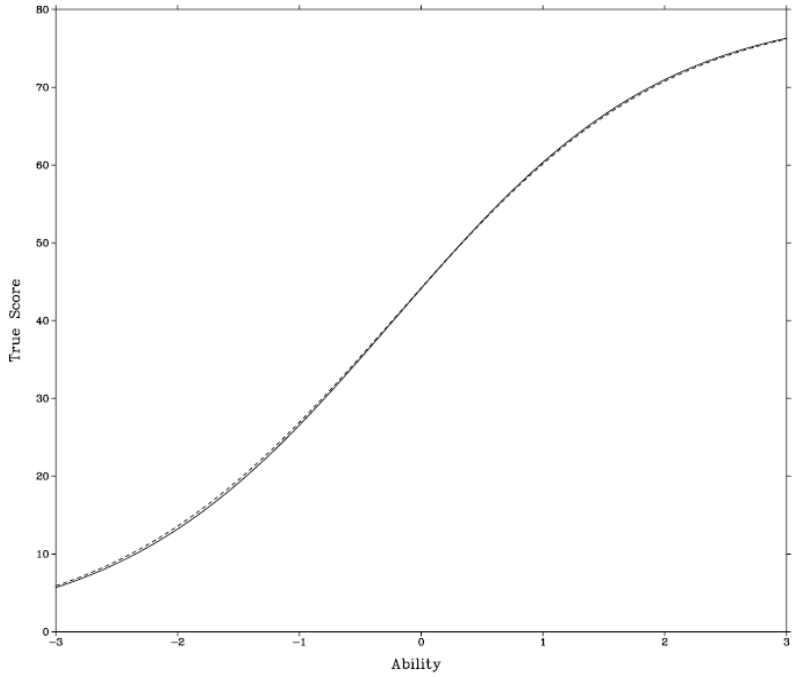


Table 4.1: Difficulty (B) and Discrimination (R-bis) Specifications for ELA MC Items

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	3 - 7
0.3	0.4	10 - 12
0.4	0.5	12 - 16
0.5	0.6	24 - 28
0.6	0.7	19 - 23
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 English-language arts 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 standalone 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 were organized in the following order: Number Sense, Probability and Statistics, Algebra and Functions, Measurement and Geometry, and Algebra I. Mathematical reasoning items were 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 administration, every person having access to test materials maintains the security and confidentiality of the tests. ETS's 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 an Office of Testing Integrity (OTI), which is described in the next section.

Educational Testing Service's Office of Testing Integrity

The Office of Testing Integrity is a division of ETS that provides quality assurance services for all testing programs administered by ETS and resides in the ETS legal department. The Quality Assurance division of ETS 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

OTI helps prevent misconduct on the part of test takers and administrators, detect potential misconduct through empirically established indicators, and resolve situations in a fair and balanced way that reflects the laws and professional standards governing the integrity of testing. In its pursuit of enforcing secure practices, ETS and 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.

- Use of electronic devices is strictly prohibited in the meeting rooms.

Item Bank

Once the item review panel completes its review the items are placed in the item bank along with their statistics and reviewers' evaluations of their quality. 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 that ETS takes 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 authorize 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 and other authorized locations. Pearson distributes the booklets to school districts in securely packaged boxes.

Test Administration

Pearson receives testing materials from printers, packages them, and sends them to school districts. Pearson ships secure and non-secure test materials in separate shipments, both via secure, expedited delivery. When materials arrive at the district they must be signed for and inventoried. 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 district CAHSEE 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 secure box contain the bar code range of books 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 district CAHSEE coordinators are required to keep all test materials in central locked storage except during actual test administration times. After inventorying all test booklets, at the district and local sites, the boxes in which they were shipped are resealed and not opened until the day of the examination. ETS audits district offices and schools several days prior to the exam, checking to see if they properly handle materials according to the published required security procedures described in the *District and Test Site Coordinator's Manual* (CDE, 2008a). Test site coordinators are responsible for accounting for and returning all secure materials to the district CAHSEE coordinator, who is responsible for returning them to the CAHSEE Scoring and Processing Center at Pearson Educational Measurement. All test booklets, upon return receipt at Pearson's receiving facility, are scanned and accounted for. If books are missing, district test coordinators are contacted and asked to recover and return the booklets. A final report of missing books is forwarded to the CAHSEE Office at the CDE.

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

- District CAHSEE 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 district CAHSEE 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 CDE of their planned testing schedule.

- Test security during testing is managed by the district CAHSEE coordinator as well as the CAHSEE coordinator at each test site. All students are to be 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 gathers all pertinent information on the *Test Security Breach Report* supplied in the *CAHSEE District and Test Site Coordinator's*

Manual (CDE, 2008a) 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 *District and Test Site Coordinator's Manual* (CDE, 2008a) and are reviewed annually with the CDE CAHSEE Office.

Processing Security

Pearson has established the following security safeguards of their sites:

- Access to the facility is controlled.

- No test materials may leave the facility during the project without the permission of a person or persons designated by the CDE.

- All staff must wear Pearson identification badges at all times in Pearson facilities.

- No recording or photographic equipment is allowed in the processing area without the consent of the CDE.

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, after which 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 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 secure FTP 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 SKM 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 at ETS performs quality control procedures before passing files to the ETS Statistical Analysis group. The Statistical Analysis group then keeps 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, including student results, are posted in secure district folders on the CAHSEE Web site. Individual student score reports are printed on security paper and shipped by secure delivery to the LEA's. 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 at ETS 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 National School Lunch Program (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 ETS and 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.

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, ETS's Code of Ethics further prohibits ETS employees from financial misuse, conflicts of interest, and unauthorized appropriation of ETS's 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 ETS Office of Testing Integrity 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 District and Test Site Coordinator's Manual* (CDE, 2008a). This manual is described in the next section.

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

District CAHSEE Coordinator

Each local education agency (LEA)[3] designates a district CAHSEE 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) who has been trained to administer the tests and has signed a *CAHSEE Test Security Affidavit*. Test examiners must follow the directions in the *CAHSEE Directions for Administration* (DFA) (CDE, 2008b) 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 *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 *CAHSEE Test Security Affidavits*.

CAHSEE Directions for Administration

CAHSEE Directions for Administration are manuals used by test examiners to administer the CAHSEE to students. They must follow all directions and guidelines and read, word-for-word, the instructions to students in "SAY" boxes to ensure test standardization.

[3] Local education agencies include 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.

CAHSEE District and Test Site Coordinator's Manual

Test administration procedures are to be followed exactly so that all students have an equal opportunity to demonstrate their academic achievement. The *CAHSEE District and Test Site Coordinator's Manual* (CDE, 2008a) 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, Education Code,* or to detail all of the coordinator's responsibilities.

CAHSEE Online

CAHSEE Online is a secure, Web-based application that allows district CAHSEE 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 they contain. 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 school, and may be sorted by group(s) within the school, depending on whether the district CAHSEE coordinator provided a sorting order on the Pre-ID file. All materials are sent to the district CAHSEE 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.

Accommodations for Students with Disabilities

All students participate in the CAHSEE Program, including students with disabilities and English learners. For eligible students with disabilities, the CAHSEE requirement can be satisfied by passing the examination, meeting the exemption requirement pursuant to California *Education Code (EC)* Section 60852.3, or receiving a local waiver pursuant to *EC* Section 60851(c). CAHSEE practice is that reasonable testing accommodations be provided to candidates with documented disabilities that are identified in the Americans with Disabilities Act (ADA). The ADA mandates that test accommodations be individualized, meaning that no single type of test accommodation may be adequate or appropriate for all individuals with any given type of disability. CAHSEE test takers with disabilities may be tested under standard conditions if only minor adjustments to the testing environment are required (e.g.,

wheelchair access, large print versions, sign language interpreter for spoken directions).

Identification of Students with Test Variations, Accommodations or Modifications

Most students with disabilities and English learners take the CAHSEE under standard conditions. 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 to provide test examiners with any information about students who require accommodations or modifications as specified in their individualized education program (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 has 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 therefore the meaning of the obtained score. 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 examination. In addition, scores for students tested with modifications are counted as Far Below Basic for API calculations and Not Proficient for AYP calculations. Test administration variations and accommodations do not result in changes to students' scores for API or AYP calculations.

Demographic Data Corrections

After reviewing student data, some school districts may discover demographic data that are incorrect. The Demographics Data Corrections function module of the CAHSEE Online gives school districts the means 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 and, if more than five percent of the students tested are involved, could affect a school's API and AYP.

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 District and Test Site Coordinator's Manual* (CDE, 2008a).

Test Administration Incidents

A test administration incident is any event that occurs before, during, or after test administrations that does not conform to the instructions stated in the *DFA*s and the *CAHSEE District and Test Site Coordinator's Manual* (CDE, 2008a). These events include test administration errors, disruptions, and student 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 district CAHSEE 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 2008–09. The statistics presented in this chapter are divided into five sections and are presented in the following order:

1. Classical Item Analyses
2. Differential Item Functioning Analyses
3. Item Response Theory Analyses (Calibration, Scaling, and Equating)
4. Reliability Analyses
5. Validity Evidence

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

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

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

    - Results of IRT analyses that include the distribution of items based on their fit to the Rasch model.

    - DIF analyses which include the list of items flagged for significant DIF and the distributions of items across DIF categories.

    - 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 error 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. Finally, 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 as a result of 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 final item analyses (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 key information about the quality of each item empirically. 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 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 what is observed, which is a binary classification — the correct answer or the incorrect answer.

Point-biserial correlations are computed using the following formula:

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

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. It can be seen that the r_{bis} values are always greater than those of r_{ptbis}.

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 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 using the formula:

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

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 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.

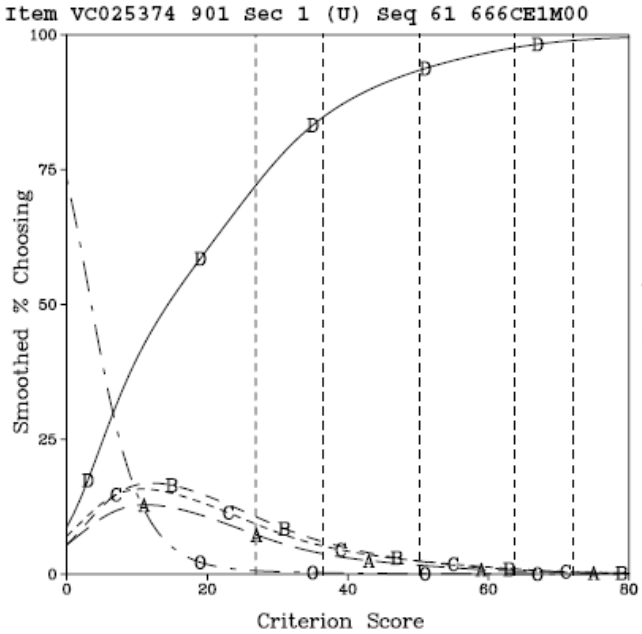
Percent of students omitting an item:

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 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 files loaded 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 related 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

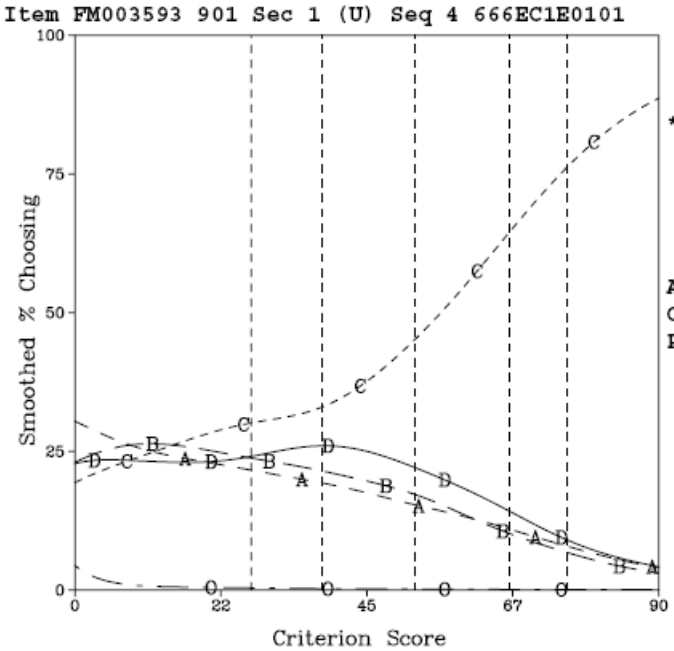


Multiple Choice, 4 choice

	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 some 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 2008 through May 2009 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 Characteristics

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 characteristic 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. The purposes of flagging items for review are to verify that the item is correctly keyed, there is one clear and correct answer, and the 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, differential item functioning studies were conducted following the classical item analyses. Differential item functioning 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 differences that are seen.

ETS used 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, was used.[4] 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 ETS 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.

---

[4] 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[5], Hispanic, and African American), and English language proficiency (English proficient compared to English learner[6]). Operational items flagged for C DIF are reviewed by an expert committee consisting of CAHSEE item development staff, 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.

[5] Combined Asian group includes students from Asian, Filipino, and Pacific Islander groups.
[6] According to "CAHSEE DIF Procedures," DIF analyses were performed for English Learners on mathematics items only. From September 2005, DIF analyses were also performed for English Learners 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 2009 items were placed on the reference scale through a set of linking items that appeared in the 2008 operational forms and were re-administered in 2009. 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 details on both models.

ETS uses a computer system called the 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 result in 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 parameters.

The resulting calibrations are then transformed to the existing scale, using the Stocking and Lord (1983) 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 item bank values. The linking process is

iterative and involves an inspection of differences between the new estimates and the bank estimates for the linking items. Items with large weighted or unweighted 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:

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

where abilities are grouped in the intervals of .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 at ETS.

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 bank 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 2008–09 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 2008–09 forms were linked back to scale established in February 2004.

Complete raw-to-scale score conversion tables for the 2008–09 CAHSEE are presented in Appendix 6.O. The raw scores and corresponding unrounded converted scale scores are listed in those tables. For all the 2008–09 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 May 2008 operational form was used as the Braille, large-print, and audio CD forms for the July 2008 administration for both ELA and mathematics. The October 2008 operational form was used for the October, November, and December 2008 administrations, and the February operational form was used for the February, March, and May 2009 administrations of the Braille, large-print, and 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 final item analyses. 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 2009	ELA	570 (88)
	Mathematics	383 (86)
March 2009	ELA	852 (88)
	Mathematics	578 (86)

Summaries of Scaled Item Response Theory B-Values

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

Reliability Analyses

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

Test Reliabilities and Standard Errors of Measurement

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

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

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

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

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

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

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

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

Reliability estimates for the ELA section across the seven administrations ranged from 0.87 to 0.94. Reliability estimates for the mathematics section across the seven administrations ranged from 0.85 to 0.95. Reliabilities for the MC strands for ELA ranged from 0.42 to 0.83, while the strand reliabilities for mathematics ranged from 0.45 to 0.85. The reliabilities for the ELA CR (unweighted) item score ranged from 0.24 to 0.44.

The standard error of measurement (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 standard errors of measurement for the ELA total raw scores and the mathematics raw scores across the 2008–09 administrations ranged from 3.98 to 4.63

and 3.63 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 2008–09 administrations, the correlations between ELA strands (including essay) ranged from 0.30 to 0.80, while the correlations between mathematics strands ranged from 0.26 to 0.80. The reliabilities of the ELA strands ranged from 0.24 to 0.83, and the reliability of the mathematics strands ranged from 0.45 to 0.85. The SEM for the MC strands for ELA ranged from 1.03 to 2.03, while the SEM for mathematics ranged from 1.45 to 2.09. The SEM for the ELA CR items ranged from 0.46 to 0.59.

Subgroup Reliabilities and Standard Errors of Measurement

Reliabilities (REL) and SEM estimates are reported for subgroups from the February and March 2009 census administrations, where larger case counts are available. Table 6.K.8 shows REL and SEM for the MC total and the composite score for gender, ethnic, accommodation, non-accommodation, English learner, and English-proficient groups for the ELA February 2009 administration. Table 6.K.9 provides the same information for the mathematics test of the February 2009 administration. Tables 6.K.10 and 6.K.11 provide REL and SEM for the subgroups for ELA and mathematics for the March 2009 administration. The sample sizes for these REL also appear in these tables.

Reliability estimates for the ELA composite (MC+CR) 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, REL are similar for the ethnic groups; the SEM are similar for most groups with a range of 2.95 (White) to 3.60 (Hispanic) for the MC total and a range of 3.71 (White, February admin; Asian, March admin) to 4.18 (African American) for the composite. All REL are greater than or equal to 0.88.

The REL for the accommodation groups are slightly lower than the nonaccommodation groups and the SEM are higher. REL for the English Learner group are lower than that for the English-proficient group. SEM is higher for the English learner group than the English-proficient group.

REL and SEM for the mathematics total score are similar for gender groups. REL for ethnic groups are similar; the SEM are similar for most groups with a range of 2.86 (Asian) to 3.96 (African American). The REL for the accommodation groups are slightly lower than the nonaccommodation groups and the SEM are slightly higher. REL for the English-proficient group is slightly higher than that for the English learner group, whereas the SEM is higher for the English learner group than the English-proficient group. All REL are greater than or equal to 0.90.

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. Nonvalid responses are scored as zero. 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, it is considered discrepant, and a scoring leader or content specialist provides a third score, which becomes the score of record. The ratings obtained from the first two readers are used to carry out inter-rater agreement and generalizability analyses to assess the reliability of the writing scores.

Rater Agreement

An important part of the analysis of any multiple-rated CR item is the degree to which the individual rater scores agree. Tables 1, 3, 5, 7, 9, 11, and 13 in Appendix 6.L present the possible score combinations, the distribution of the two ratings, and the distribution of differences between the first and second ratings of each CR item. As the diagonals of the tables show, the majority of raters assigned the same score. When ratings differed, most were by only one score point. Only the February and March 2009 administrations had about 1 percent of CR item scores assigned by the two raters differing by two score points or more.

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 0.18 to 0.32. The correlation between first and second ratings across the seven administrations ranged from 0.67 to 0.78.

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 7 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 2009 administrations and the largest percentage was in the October 2009 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 0.80 to 0.88 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 ETS-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 0.93 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 0.91 to 0.99, while the consistency for mathematics ranged from 0.94 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 0.90 to 0.96, while the consistency for mathematics ranged from 0.91 to 0.97.

The reliability of classification results for the cuts 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 level ranged from 0.90 to 0.94, while the decision accuracy for mathematics at the Pass/Not Pass level ranged from 0.88 to 0.95. The decision consistency for ELA at the Pass/Not Pass level ranged from 0.87 to 0.92, while the decision consistency for mathematics at the Pass/Not Pass ranged from 0.84 to 0.92. The magnitudes 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 grade-level competency in reading, writing, 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 10. The mathematics examination measures grade 7 level mathematics standards and Algebra I.

CAHSEE test blueprints provide the number of items per standard that will appear on an operational form. 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 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 grade 10, 11, or 12 or are enrolled in adult schools operated by a school district. Beginning in the 2005–06 school year, with the

exception of eligible students with disabilities, 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 score 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. The total score also is used to classify students in terms of their level of knowledge and skill in the subject area. These three levels are called performance levels and are as follows: advanced, proficient, and pass/fail.

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' achievement 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:

Description 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 2009 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. One of the reviews is conducted by external reviewers, that is, the content review. The content review committees are responsible for reviewing all newly developed items for alignment to the California content standards. Additionally, the Statewide Pupil Assessment Review (SPAR) is responsible for reviewing and approving the achievement tests to be used statewide for the testing of students in California public schools. 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. Alignment results should demonstrate that the assessments represent the full range of the content standards, and that these assessments measure student knowledge in the same manner and at the same level of complexity as expected in the content standards. The Human Resource Research Organization (HumRRO) performed an alignment study for the February 2008 ELA and mathematics forms. The report, titled *2008 Review of the CAHSEE Test Quality* (Taylor, Johnstone, Wise, Thacker, & Hardoin-Mandeville, 2008), may be found in the Independent Evaluation of the California High School Exit Examination (CAHSEE): 2008 Evaluation Report, Chapter 2.

HumRRO utilized the Webb alignment method to evaluate the alignment of the 2008 CAHSEE 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 depth of knowledge 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. Overall, the alignment was judged to be good. The 2008 independent evaluation report is 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

Student performance on the CAHSEE and CSTs 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 students with disabilities. Results suggest the CST results were good indicators of how students performed on the CAHSEE, demonstrating consistency of test results across testing programs.

To investigate the performance of students who were repeat test-takers, HumRRO examined the average CST scores for all students and for grade 12 students taking the CAHSEE. The latter group had mean CST scores between 0.75 and 1.00 standard deviations below the means scores for all students on the grade 11 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 students with disabilities, the 2007 CAHSEE results were compared with the 2006 CST results. Most of the grade 10 students in special education programs in 2007 had taken the ELA end-of-course test in 2006. Results indicate that the majority of students who took the Grade 9 or 10 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 10 students with disabilities 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

Differential item functioning 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 English learner).

Across all 2008–09 administrations few operational items were flagged for significant DIF per administration. Of the 31 items that were flagged, 24 were ELA items and 7 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 of 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., 0.81 and 0.78 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 a different construct. 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 0.69.

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 0.80 to 0.88 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 69 to 82 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 corrrelations 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 0.41 to 0.56 for ELA and from 0.37 to 0.57 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 standard error of measurement estimates (SEM) are reported for demographic groups for the February and March 2009 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 English learner 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 0.72 (SD = 0.06) and 0.67 (SD = 0.08) for the February and March administrations, respectively. The mathematics strands exhibit mean correlations of 0.74 (SD = 0.04) for both administrations. The mean intercorrelations between the ELA and mathematics strand scores are 0.63 (SD = 0.05) and 0.59 (SD = 0.07) 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 10 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, empirical information in support of the CAHSEE follows, as 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 California 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 item characteristic curves 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)

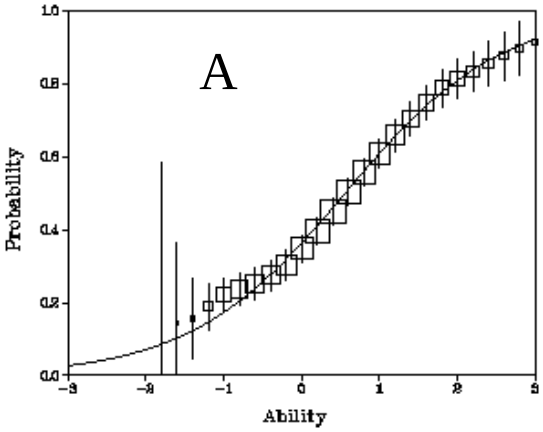
- Fit at the passing score may be slightly out of error range.
- Theoretical curve outside of error range at some regions across ability range.
- Empirical curve may have a zero slope at and around the passing score.
- Large Chi-square value relative to the other items in the calibration with similar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

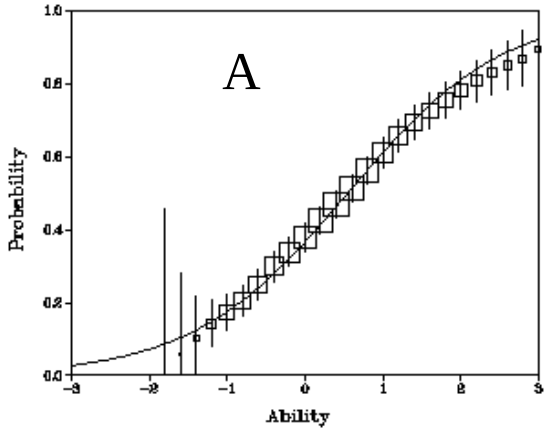
Flag F (Item 98, FM002768)

- 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.

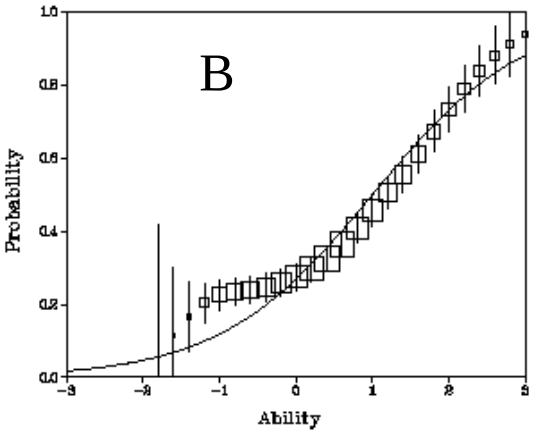
Figure 6.3: Samples of CAHSEE Item-Fit Rating Categories



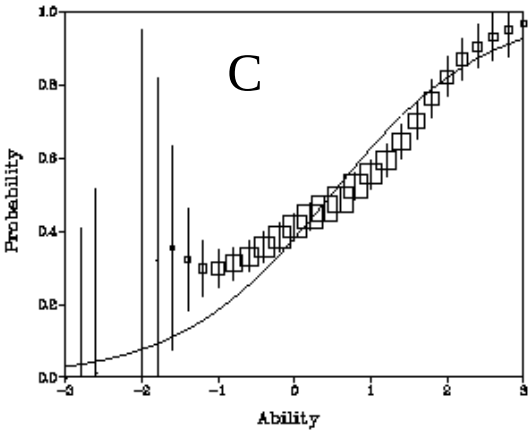
#93 FM002619901 4 Choice P+=0.494
a=0.588 F, b=0.561, c=0.000 F, CHI=11.85



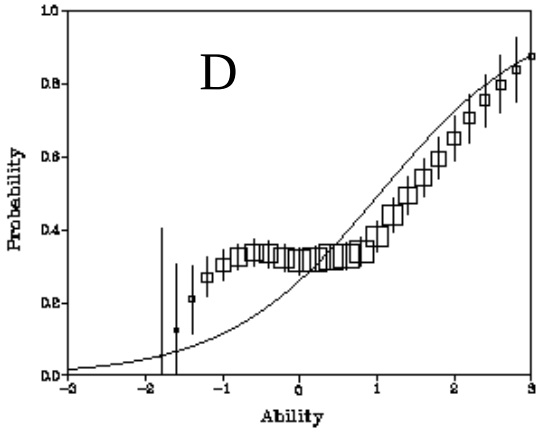
#95 FM002640901 4 Choice P+=0.496
a=0.588 F, b=0.550, c=0.000 F, CHI=10.50



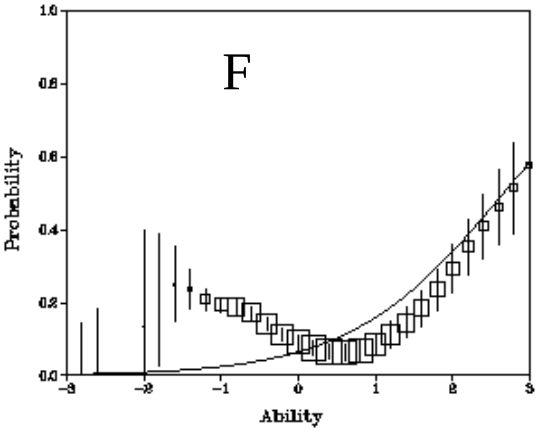
#96 FM002641901 4 Choice P+=0.402
a=0.588 F, b=1.003, c=0.000 F, CHI=116.63



#97 FM002767901 4 Choice P+=0.506
a=0.588 F, b=0.486, c=0.000 F, CHI=151.58



#94 FM002620901 4 Choice P+=0.395
a=0.588 F, b=1.040, c=0.000 F, CHI=480.45



#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 2008

Table 6.B.1: Summary of Operational Item Statistics—ELA

Content area[1]	Number of items	IRT b value				Point-biserial/Pearson correlation			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	0.00	0.67	-2.08	1.18	0.32	0.07	0.16	0.46
RC	18	-0.11	0.79	-2.08	0.72	0.33	0.06	0.23	0.45
RL	20	-0.09	0.66	-1.45	0.89	0.31	0.07	0.17	0.46
RW	7	-0.29	0.46	-1.06	0.34	0.33	0.11	0.18	0.45
WA	1	0.88	-	0.88	0.88	0.61	-	0.61	0.61
WC	15	-0.06	0.54	-0.86	0.80	0.33	0.05	0.24	0.39
WS	12	0.51	0.49	-0.52	1.18	0.29	0.07	0.16	0.43

[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	-0.23	0.65	-2.06	1.21	0.29	0.08	0.07	0.43
A1	12	0.23	0.36	-0.51	0.75	0.22	0.07	0.08	0.33
AF	20	-0.30	0.71	-1.75	0.86	0.31	0.07	0.18	0.42
MG	18	-0.16	0.59	-0.93	1.21	0.30	0.10	0.07	0.43
MR	8	-0.54	0.67	-1.35	0.63	0.31	0.05	0.25	0.40
NS	17	-0.36	0.70	-2.06	0.56	0.29	0.06	0.19	0.43
PS	13	-0.45	0.64	-1.25	0.82	0.33	0.06	0.21	0.39

[1] A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS=Number Sense, PS=Probability & Statistics

Table 6.B.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent
A	Use	29	40
B	Use	29	40
C	Use	14	19
D	Review	1	1
Total		73	100

Table 6.B.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent
A	Use	32	40
B	Use	23	29
C	Use	23	29
D	Review	1	1
F	Do not use	1	1
Total		80	100

Table 6.B.5: Operational Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	FM007193	L195B003	0	22	A	S	S	S	S	A	A	A	C-
MATH	VC023929	M20646	0	52	A	S	S	S	S	C+	A	A	A
MATH	VC138625	M32158	0	91	A	S	S	S	S	C+	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.
S Indicates that DIF was not performed due to insufficient sample size.

Table 6.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	0	0	1	1	1	1
B-	0	0	0	0	0	0	0	0	0	0	5	7	1	1	1	1	3	4		
A	73	100	0	0	0	0	0	0	0	0	65	89	72	99	72	99	68	93		
B+	0	0	0	0	0	0	0	0	0	0	3	4	0	0	0	0	1	1		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	73	100	73	100	73	100	73	100	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	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.7: Distribution of Operational Item DIF Classifications—Mathematics

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C DIF Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	3	4	0	0	0	0	0	0	0	0	8	10	1	1	0	0	0	0		
A	75	94	0	0	0	0	0	0	0	0	69	86	79	99	79	99	80	100		
B+	2	3	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0		
C+[2]	0	0	0	0	0	0	0	0	0	0	2	3	0	0	0	0	0	0	2	3
Small N[3]	0	0	80	100	80	100	80	100	80	100	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.B.8: Listing of CR Item Statistics—ELA

Accession number	VC124327
CAHSEE ID	L3OSA480
Polyserial correlation	0.68
IRT b-value	0.8809
Step category 1[1]	2.0581
Step category 2	1.3899
Step category 3	3.4929
Step category 4	-1.1097
Step category 5	-0.7403
Step category 6	-2.9516
Step category 7	-2.1394
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	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.C: Summary Statistics for Operational Items—October 2008

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	-0.03	0.69	-1.82	1.34	0.35	0.07	0.15	0.49
RC	18	-0.21	0.68	-1.82	0.66	0.36	0.08	0.17	0.46
RL	20	-0.15	0.57	-0.92	0.93	0.35	0.06	0.26	0.45
RW	7	-0.34	0.75	-1.32	0.98	0.36	0.05	0.27	0.41
WA	1	0.91	-	0.91	0.91	0.66	-	0.66	0.66
WC	15	-0.02	0.76	-1.52	1.30	0.35	0.09	0.15	0.49
WS	12	0.50	0.55	-0.67	1.34	0.33	0.07	0.22	0.43

[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	-0.23	0.76	-2.10	1.22	0.32	0.07	0.14	0.47
A1	12	0.37	0.62	-0.59	1.22	0.29	0.08	0.14	0.41
AF	20	-0.31	0.54	-1.15	0.68	0.34	0.07	0.22	0.47
MG	18	0.03	0.74	-1.65	1.06	0.33	0.08	0.21	0.47
MR	8	-0.27	0.53	-1.08	0.67	0.32	0.05	0.24	0.39
NS	17	-0.48	0.91	-2.10	1.00	0.29	0.05	0.17	0.36
PS	13	-0.69	0.54	-1.80	0.28	0.35	0.06	0.19	0.43

[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
A	Use	22	30
B	Use	23	32
C	Use	28	38
Total		73	100

Table 6.C.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	38	48
B	Use	20	25
C	Use	22	28
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 #	CAHSEE ID	Form	Item #	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
ELA	FM004202	L132B001	0	22	C-	A	C-	A	B-	B-	A	B-	C-
ELA	VC016614	L00SA098	0	69	A	A	A	A	A	A	A	A	C+
ELA	VC017359	L10SA335	0	78	A	A	A	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.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]	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	2	3
B-	1	1	0	0	5	7	1	1	3	4	3	4	0	0	2	3	4	5		
A	71	97	73	100	61	84	69	95	69	95	68	93	73	100	71	97	66	90		
B+	0	0	0	0	6	8	2	3	1	1	2	3	0	0	0	0	1	1		
C+	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	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.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-	2	3	0	0	2	3	2	3	1	1	1	1	0	0	1	1	1	1		
A	78	98	80	100	76	95	75	94	76	95	76	95	80	100	79	99	79	99		
B+	0	0	0	0	2	3	3	4	3	4	3	4	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[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	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	**VC020396**
CAHSEE ID	**L30SA331**
Polyserial correlation	0.73
IRT b-value	0.9091
Step category 1[1]	0.7006
Step category 2	1.7600
Step category 3	3.5774
Step category 4	-0.7531
Step category 5	-0.5900
Step category 6	-2.3077
Step category 7	-2.3873
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.D: Summary Statistics for Operational Items—November 2008

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	-0.17	0.75	-1.70	1.49	0.38	0.07	0.19	0.51
RC	18	-0.25	0.79	-1.25	1.49	0.38	0.09	0.19	0.51
RL	20	-0.47	0.51	-1.55	0.83	0.41	0.06	0.28	0.50
RW	7	-0.12	0.62	-1.06	0.79	0.40	0.02	0.37	0.43
WA	1	0.83	-	0.83	0.83	0.66	-	0.66	0.66
WC	15	-0.18	0.96	-1.70	1.43	0.37	0.07	0.22	0.50
WS	12	0.33	0.58	-0.72	1.22	0.36	0.05	0.28	0.43

[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	-0.19	0.68	-2.08	1.13	0.34	0.07	0.16	0.49
A1	12	0.24	0.45	-0.44	0.96	0.32	0.07	0.16	0.42
AF	20	-0.32	0.69	-1.59	0.77	0.36	0.05	0.19	0.42
MG	18	-0.08	0.63	-1.29	1.13	0.35	0.07	0.22	0.46
MR	8	-0.06	0.59	-1.37	0.54	0.32	0.06	0.23	0.38
NS	17	-0.24	0.76	-2.08	0.79	0.33	0.07	0.21	0.49
PS	13	-0.48	0.65	-1.42	0.37	0.34	0.06	0.20	0.43

[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	35	48
B	Use	28	38
C	Use	10	14
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	42	53
B	Use	30	38
C	Use	8	10
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 #	CAHSEE ID	Form	Item #	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	FM005219	L177A002	0	1	C-	A	A	A	A	A	A	A	A
ELA	FM005223	L177A007	0	3	A	A	B-	B-	C-	B-	A	B-	A
ELA	VC392080	L506H002	0	17	B-	A	C-	A	A	B-	A	A	C-
ELA	VC254905	L451G002	0	28	A	A	C-	A	A	B-	A	A	A
ELA	VC017374	L10SA351	0	73	A	A	C-	B-	A	B-	B-	B-	A
ELA	VC016607	L00SA084	0	79	A	A	C+	A	B+	B+	A	A	A
MATH	VC024214	M21002	0	11	C-	A	A	A	A	A	A	A	A
MATH	VC023975	M20696	0	91	A	A	B+	A	C+	B+	A	B+	B+

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	3	4	0	0	1	1	0	0	0	0	0	0	1	1	5	7
B-	2	3	0	0	4	5	3	4	2	3	6	8	1	1	2	3	1	1		
A	69	95	73	100	60	82	69	95	66	90	63	86	72	99	71	97	68	93		
B+	1	1	0	0	5	7	1	1	4	5	4	5	0	0	0	0	3	4		
C+[2]	0	0	0	0	1	1	0	0	0	0	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

	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-	1	1	0	0	3	4	0	0	1	1	0	0	0	0	0	0	0	0		
A	77	96	80	100	75	94	79	99	77	96	78	98	80	100	79	99	79	99		
B+	1	1	0	0	2	3	1	1	1	1	2	3	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]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	2	3

1 Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
2 Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
3 Small N indicates that DIF analysis was not performed due to insufficient sample size.
4 Total 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	**VC020398**
CAHSEE ID	**L30SA333**
Polyserial correlation	0.74
IRT b-value	0.8296
Step category 1[1]	1.8231
Step category 2	1.6779
Step category 3	3.5128
Step category 4	-0.9896
Step category 5	-0.8299
Step category 6	-2.7086
Step category 7	-2.4858
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	A
DIF category, White-Filipino	A
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	A

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.
[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

Appendix 6.E: Summary Statistics for Operational Items—December 2008

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	-0.17	0.68	-1.56	1.12	0.35	0.07	0.17	0.48
RC	18	-0.21	0.63	-1.14	0.85	0.36	0.08	0.17	0.48
RL	20	-0.33	0.63	-1.56	0.99	0.36	0.07	0.24	0.47
RW	7	-0.45	0.71	-1.44	0.54	0.31	0.07	0.24	0.45
WA	1	0.58	-	0.58	0.58	0.59	-	0.59	0.59
WC	15	-0.15	0.76	-1.33	1.12	0.34	0.06	0.20	0.42
WS	12	0.24	0.59	-0.88	1.11	0.34	0.08	0.20	0.44

[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	-0.11	0.65	-1.98	1.37	0.28	0.07	0.11	0.43
A1	12	0.45	0.44	-0.37	1.16	0.23	0.08	0.12	0.36
AF	20	-0.24	0.50	-1.15	0.62	0.30	0.08	0.11	0.43
MG	18	-0.05	0.86	-1.98	1.37	0.28	0.07	0.16	0.43
MR	8	-0.28	0.52	-1.05	0.62	0.31	0.04	0.27	0.41
NS	17	-0.20	0.53	-1.05	0.96	0.28	0.05	0.20	0.38
PS	13	-0.38	0.61	-1.24	0.74	0.31	0.05	0.21	0.41

[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
A	Use	26	36
B	Use	22	30
C	Use	25	34
Total		73	100

Table 6.E.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent
A	Use	33	41
B	Use	32	40
C	Use	15	19
Total		80	100

Table 6.E.5: Operational Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	VC018617	L226D003	0	2	C-	S	S	S	S	S	C-	A	C-
ELA	VC392202	L569G012	0	14	A	S	S	S	S	S	A	A	C+
ELA	VC392368	L581G011	0	34	A	S	S	S	S	S	A	A	C+
MATH	VC023358	M14115	0	4	C-	S	S	S	S	S	A	A	A
MATH	VC392775	M40016	0	11	C-	S	S	S	S	S	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.
S Indicates that DIF was not performed due to insufficient sample size.

Table 6.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]	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1
B-	5	7	0	0	0	0	0	0	0	0	0	0	3	4	1	1	5	7		
A	65	89	0	0	0	0	0	0	0	0	0	0	67	92	72	99	59	81		
B+	2	3	0	0	0	0	0	0	0	0	0	0	2	3	0	0	6	8		
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	0	0	0	0	0	0		
Total	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.

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]	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3
B-	0	0	0	0	0	0	0	0	0	0	0	0	2	3	3	4	2	3		
A	73	91	0	0	0	0	0	0	0	0	0	0	78	98	76	95	77	96		
B+	5	6	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	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.E.8: Listing of CR Item Statistics—ELA

Accession number	**VC367716**
CAHSEE ID	**L30SA248**
Polyserial correlation	0.66
IRT b-value	0.5765
Step category 1[1]	2.2457
Step category 2	1.5143
Step category 3	3.4574
Step category 4	-1.1265
Step category 5	-0.8643
Step category 6	-2.6448
Step category 7	-2.5818
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.F: Summary Statistics for Operational Items—February 2009

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	-0.07	0.70	-1.59	1.66	0.45	0.08	0.25	0.60
RC	18	-0.23	0.77	-1.36	1.66	0.42	0.09	0.25	0.54
RL	20	-0.14	0.58	-1.59	0.88	0.45	0.06	0.31	0.51
RW	7	-0.27	0.68	-1.34	0.47	0.42	0.10	0.32	0.60
WA	1	0.91	-	0.91	0.91	0.74	-	0.74	0.74
WC	15	-0.07	0.77	-1.55	1.42	0.46	0.06	0.34	0.56
WS	12	0.32	0.63	-0.65	1.20	0.49	0.07	0.36	0.58

[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	-0.22	0.73	-3.06	1.43	0.46	0.08	0.21	0.62
A1	12	0.38	0.39	-0.11	1.19	0.48	0.07	0.37	0.56
AF	20	-0.39	0.56	-1.61	0.55	0.47	0.07	0.31	0.57
MG	18	-0.03	0.87	-1.24	1.43	0.45	0.08	0.29	0.62
MR	8	-0.15	0.51	-0.97	0.35	0.45	0.06	0.36	0.54
NS	17	-0.46	0.83	-3.06	0.56	0.44	0.10	0.25	0.62
PS	13	-0.48	0.52	-1.19	0.35	0.45	0.09	0.21	0.56

[1] A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS=Number Sense, PS=Probability & Statistics

Table 6.F.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent
A	Use	37	51
B	Use	21	29
C	Use	14	19
D	Review	1	1
Total		73	100

Table 6.F.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	28	35
B	Use	31	39
C	Use	19	24
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.F.5: Operational Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	FM003752	L152A001	0	40	A	A	B-	A	C-	B-	A	A	A
ELA	VC392718	L5OSA673	0	73	A	A	C-	A	B-	B-	A	A	A
ELA	VC392697	L5OSA649	0	77	A	A	C-	C-	C-	C-	A	C-	A
MATH	VC023326	M14039	0	33	A	A	C-	A	B-	B-	A	A	B-

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

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient - English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	2	3	1	1	2	3	1	1	0	0	1	1	0	0	3	4
B-	4	5	0	0	6	8	1	1	5	7	6	8	1	1	0	0	2	3		
A	69	95	73	100	60	82	71	97	62	85	63	86	71	97	72	99	70	96		
B+	0	0	0	0	5	7	0	0	4	5	3	4	1	1	0	0	1	1		
C+[2]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	3	4

1 Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
2 Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
3 Small N indicates that DIF analysis was not performed due to insufficient sample size.
4 Total percentage may differ slightly from the sum of the category percentages due to rounding.

Table 6.F.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	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
B-	4	5	0	0	6	8	0	0	4	5	6	8	0	0	0	0	1	1		
A	74	93	80	100	71	89	80	100	76	95	74	93	80	100	80	100	79	99		
B+	2	3	0	0	2	3	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]	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	VC367717
CAHSEE ID	L30SA249
Polyserial correlation	0.81
IRT b-value	0.9129
Step category 1[1]	1.7479
Step category 2	1.7151
Step category 3	3.6608
Step category 4	-0.6450
Step category 5	-0.7571
Step category 6	-2.7914
Step category 7	-2.9303
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	B
DIF category, White-Combined Asian	A
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups[2]	B

[1] Step categories refer to the parameters describing each item category in the polytomous item calibrations.
[2] This refers to the most extreme DIF category found among all focal groups for which a comparison was made. Positive DIF categories favor the focal group, and negative DIF categories favor the reference group.

Appendix 6.G: Summary Statistics for Operational Items—March 2009

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	-0.08	0.70	-1.47	1.51	0.42	0.09	0.22	0.58
RC	18	-0.10	0.81	-1.28	1.51	0.44	0.09	0.28	0.58
RL	20	-0.38	0.67	-1.47	1.19	0.43	0.10	0.26	0.58
RW	7	-0.08	0.39	-0.65	0.63	0.36	0.07	0.24	0.43
WA	1	0.85	-	0.85	0.85	0.75	-	0.75	0.75
WC	15	-0.05	0.71	-1.19	1.05	0.43	0.06	0.34	0.55
WS	12	0.33	0.55	-0.61	1.12	0.42	0.09	0.22	0.51

[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	-0.22	0.73	-1.74	1.39	0.46	0.08	0.22	0.66
A1	12	0.35	0.56	-0.73	0.97	0.46	0.09	0.31	0.58
AF	20	-0.36	0.64	-1.58	0.74	0.48	0.07	0.33	0.58
MG	18	-0.11	0.91	-1.59	1.39	0.46	0.09	0.22	0.58
MR	8	-0.53	0.53	-1.37	0.23	0.42	0.06	0.33	0.51
NS	17	-0.46	0.75	-1.74	0.87	0.44	0.09	0.35	0.66
PS	13	-0.39	0.46	-1.01	0.35	0.43	0.07	0.29	0.52

[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	30	41
B	Use	17	23
C	Use	25	34
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
A	Use	33	41
B	Use	11	14
C	Use	31	39
D	Review	5	6
Total		80	100

Table 6.G.5: Operational Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	VC019867	L269E016	0	5	B-	A	B-	B-	C-	B-	A	A	A
ELA	VC392296	L576G007	0	15	A	A	C-	B-	C-	C-	B-	A	A
ELA	VC020996	L376E002	0	29	A	A	C-	A	C-	C-	A	A	A
ELA	VC017380	L10SA358	0	69	B+	A	C+	A	C+	B+	A	A	A
ELA	VC392724	L5OSA680	0	70	A	A	C-	C-	C-	C-	A	C-	A
ELA	VC391984	L20SA115	0	78	A	A	C-	C-	C-	C-	A	A	A

Note: Items with C values (positive and negative) were reviewed by the DIF review committee prior to scoring.

Table 6.G.6: Distribution of Operational Item DIF Classifications—ELA

	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		English Proficient-English Learner		Total C Items Across All Analyses[1]	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[2]	0	0	0	0	4	5	2	3	5	7	4	5	0	0	1	1	0	0	5	7
B-	3	4	0	0	2	3	2	3	2	3	3	4	1	1	0	0	1	1		
A	69	95	73	100	64	88	68	93	65	89	63	86	72	99	72	99	69	95		
B+	1	1	0	0	2	3	1	1	0	0	3	4	0	0	0	0	3	4		
C+[2]	0	0	0	0	1	1	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.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	
DIF category	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B-	4	5	0	0	5	6	0	0	4	5	5	6	0	0	0	0	1	1		
A	74	93	80	100	73	91	79	99	75	94	74	93	80	100	80	100	78	98		
B+	2	3	0	0	2	3	1	1	1	1	1	1	0	0	0	0	1	1		
C+[1]	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]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	0	0

1 Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
2 Small N indicates that DIF analysis was not performed due to insufficient sample size.
3 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	**VC020406**
CAHSEE ID	**L30SA341**
Polyserial correlation	0.82
IRT b-value	0.8486
Step category 1[1]	1.7516
Step category 2	1.4605
Step category 3	3.5311
Step category 4	-0.6437
Step category 5	-0.5767
Step category 6	-2.6711
Step category 7	-2.8518
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.H: Summary Statistics for Operational Items—May 2009

Table 6.H.1: Summary of Operational Item Statistics—ELA

		IRT b value				Point-biserial/Pearson correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	73	-0.11	0.65	-1.44	1.50	0.38	0.06	0.19	0.51
RC	18	-0.12	0.69	-1.23	1.50	0.37	0.07	0.19	0.49
RL	20	-0.28	0.63	-1.44	0.89	0.39	0.06	0.30	0.47
RW	7	-0.64	0.51	-1.39	0.14	0.39	0.07	0.28	0.50
WA	1	0.69	-	0.69	0.69	0.67	-	0.67	0.67
WC	15	-0.09	0.58	-1.18	0.91	0.38	0.07	0.28	0.51
WS	12	0.39	0.43	-0.52	1.11	0.38	0.06	0.27	0.48

[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	-0.23	0.58	-1.57	0.86	0.35	0.07	0.19	0.50
A1	12	0.22	0.51	-1.00	0.86	0.32	0.03	0.28	0.37
AF	20	-0.31	0.53	-1.51	0.78	0.37	0.05	0.27	0.44
MG	18	-0.21	0.56	-1.29	0.82	0.35	0.09	0.20	0.50
MR	8	-0.05	0.31	-0.57	0.49	0.36	0.06	0.27	0.44
NS	17	-0.25	0.57	-1.57	0.80	0.33	0.09	0.19	0.48
PS	13	-0.53	0.58	-1.46	0.30	0.35	0.05	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.H.3: IRT Model Data Fit Distribution of Operational Items—ELA

IRT review category	Status	Item count	Percent
A	Use	25	34
B	Use	23	32
C	Use	25	34
Total		73	100

Table 6.H.4: IRT Model Data Fit Distribution of Operational Items—Mathematics

IRT review category	Status	Item count	Percent
A	Use	33	41
B	Use	22	28
C	Use	25	31
Total		80	100

Table 6.H.5: Operational Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	VC026228	L190A003	0	2	A	A	B-	A	A	A	A	A	C-
ELA	VC026214	L188B001	0	34	C-	B-	C-	C-	A	C-	B-	C-	A

Note: Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.

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]	1	1	0	0	1	1	1	1	0	0	1	1	0	0	1	1	1	1	2	3
B-	1	1	1	1	4	5	1	1	3	4	0	0	1	1	0	0	3	4		
A	71	97	72	99	68	93	70	96	67	92	72	99	72	99	72	99	67	92		
B+	0	0	0	0	0	0	1	1	3	4	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	0	0	0	0
Small N[3]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[4]	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100	2	3

1 Items that are identified with C DIF in more than one comparison are counted once in the total C DIF column.
2 Items with C DIF values (positive or negative) were reviewed by the DIF review committee prior to scoring.
3 Small N indicates that DIF analysis was not performed due to insufficient sample size.
4 Total 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	
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-	2	3	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0		
A	76	95	80	100	79	99	76	95	77	96	79	99	80	100	78	98	80	100		
B+	2	3	0	0	0	0	3	4	2	3	0	0	0	0	1	1	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[1]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Total[2]	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	0	0

1 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.H.8: Listing of CR Item Statistics—ELA

Accession number	**VC172643**
CAHSEE ID	**L30SA478**
Polyserial correlation	0.74
IRT b Value	0.6867
Step category 1[1]	1.9227
Step category 2	1.5149
Step category 3	3.4279
Step category 4	-0.8211
Step category 5	-0.6608
Step category 6	-2.9075
Step category 7	-2.4761
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.I: Summary Statistics for Field-Test Items—February 2009

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	628	0.60	0.84	-1.78	3.29	0.39	0.11	0.09	0.63
RC	176	0.61	0.85	-1.31	3.29	0.40	0.11	0.09	0.60
RL	182	0.66	0.78	-1.01	2.76	0.39	0.11	0.09	0.63
RW	57	0.61	0.83	-1.53	2.37	0.39	0.11	0.09	0.56
WC	93	0.11	0.86	-1.78	2.44	0.42	0.10	0.13	0.60
WS	120	0.86	0.74	-0.92	3.00	0.37	0.11	0.10	0.58

[1] RC=Reading Comprehension, RL=Literary Response & Analysis, RW=Word Analysis, WA=Writing Applications, WC=Writing Conventions, WS=Writing Strategies

Table 6.I.2: Summary of Field-Test Item Statistics—Mathematics

		IRT b value				Point-biserial correlation			
Content area[1]	Number of items	Mean	Std. Dev.	Minimum	Maximum	Mean	Std. Dev.	Minimum	Maximum
Overall	428	0.52	0.89	-2.43	3.09	0.40	0.11	0.08	0.64
A1	69	0.82	0.71	-0.99	2.76	0.39	0.13	0.11	0.59
AF	96	0.43	0.88	-1.77	2.60	0.41	0.11	0.08	0.59
MG	86	0.84	0.88	-1.56	2.72	0.39	0.11	0.13	0.58
MR	62	0.76	0.87	-1.36	3.09	0.36	0.13	0.08	0.63
NS	97	0.47	0.86	-2.22	3.09	0.42	0.12	0.10	0.64
PS	80	0.07	0.89	-2.43	2.40	0.39	0.10	0.14	0.59

[1] A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS=Number Sense, PS=Probability & Statistics

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

IRT review category	Status	Item count	Percent
A	Use	123	19
B	Use	184	28
C	Use	221	34
D	Review	42	7
F	Do not use	58	9
N/A	Dropped from Calibration	18	3
Total		646	100

Table 6.I.4: IRT Model Data Fit Distribution of Field-Test Items—Mathematics

IRT review category	Status	Item count	Percent[1]
A	Use	89	20
B	Use	94	21
C	Use	180	41
D	Review	20	5
F	Do not use	45	10
N/A	Dropped from Calibration	16	4
Total		444	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 #	CAHSEE ID	Form	Item #	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	VE045404	L638H007	2	48	C+	S	S	S	S	S	A	S	A
ELA	VE045658	L650H004	4	47	A	S	S	S	S	S	B-	S	C-
ELA	VE046477	L6OSA1184	6	75	A	S	S	S	S	S	A	S	C-
ELA	VE045576	L645H001	10	46	C-	S	S	S	S	S	A	S	A
ELA	VE046478	L6OSA1185	15	75	A	S	S	S	S	C-	A	S	A
ELA	VE045758	L658H001	18	46	A	S	S	S	S	A	A	S	C-
ELA	VE045761	L658H004	19	47	C-	S	S	S	S	B-	B-	S	B-
ELA	VE045844	L664H009	22	50	C-	S	S	S	S	S	A	S	A
ELA	VE045683	L652H004	27	47	A	S	S	S	S	S	A	S	C-
ELA	VE046335	L698H009	28	46	A	S	S	S	S	S	C+	A	B+
ELA	VE044990	L621J001	29	46	A	S	S	S	S	S	A	S	C+
ELA	VE045126	L627H003	35	47	A	S	S	S	S	A	A	A	C-
ELA	VE045125	L627H002	36	51	C-	S	S	S	S	S	A	A	A
ELA	VE045614	L647H001	41	46	A	S	S	S	S	S	A	S	C-
ELA	VE045618	L647H005	41	48	A	S	S	S	S	S	A	S	C-
ELA	VE046025	L677H005	44	48	B-	S	S	S	S	A	A	S	C-
ELA	VE044532	L603J001	47	46	A	S	S	S	S	S	C-	C-	A
ELA	VE044534	L603J003	48	46	A	S	S	S	S	S	A	S	C-
ELA	VE045981	L674H017	51	47	A	S	S	S	S	S	C-	S	B-
ELA	VE045533	L643H006	53	48	B-	S	S	S	S	S	B-	S	C-
ELA	VE045971	L673H014	59	47	C-	S	S	S	S	S	A	S	A
ELA	VE046475	L6OSA1182	72	75	A	S	S	S	S	S	A	S	C+
ELA	VE045813	L662H006	73	48	A	S	S	S	S	S	A	S	C-
ELA	VE045808	L662H001	74	46	A	S	S	S	S	S	C-	S	C-
ELA	VE045728	L655H011	75	49	A	S	S	S	S	S	A	A	C-
ELA	VE045727	L655H010	76	48	C+	S	S	S	S	S	A	S	A
ELA	VE046098	L682H009	77	46	A	S	S	S	S	S	C-	S	A
MATH	VE048468	M51367	6	35	A	S	C-	S	S	C-	B-	B-	B-

Test	Accession #	CAHSEE ID	Form	Item #	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
MATH	VE047167	M50062	13	21	A	S	B+	S	S	C+	B+	B+	A
MATH	VE047173	M50068	15	21	A	S	A	S	S	A	B-	C-	C-
MATH	VC025841	M30648	15	58	C-	S	A	S	S	A	A	A	A
MATH	VE048896	M51794	17	35	A	S	C-	S	S	B-	A	A	B-
MATH	VE047186	M50081	18	20	A	S	C+	S	S	B+	A	A	A
MATH	VE047178	M50073	19	58	C+	S	B+	S	S	B+	B+	B+	A
MATH	VE047262	M50158	21	36	A	S	C-	S	S	B-	B-	A	A
MATH	VE047190	M50085	22	18	C-	S	A	S	S	A	A	A	A
MATH	VE047191	M50086	23	19	C-	S	A	S	S	A	A	A	B-
MATH	VE047200	M50095	27	18	C-	S	A	S	S	A	A	A	A
MATH	VE048453	M51352	28	36	A	S	C-	S	S	B-	B-	B-	B-
MATH	VE047192	M50087	35	18	C-	S	A	S	S	A	A	B-	A
MATH	VE047718	M50615	35	20	A	S	C-	S	S	B-	A	A	C-
MATH	VE048215	M51115	37	36	A	S	C+	S	S	A	A	A	A

Note: S indicates that DIF was not performed due to insufficient sample size.

Table 6.I.6: Distribution of Field-Test 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-	5	1	0	0	0	0	0	0	0	0	1	0	4	1	1	0	13	2	22	3
B-	18	3	0	0	0	0	0	0	0	0	4	1	31	5	6	1	45	7		
A	602	93	0	0	20	3	0	0	0	0	120	19	603	93	146	23	578	89		
B+	19	3	0	0	3	0	0	0	0	0	7	1	7	1	2	0	8	1		
C+	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	5	1
Small N[2]	0	0	646	100	623	96	646	100	646	100	514	80	0	0	491	76	0	0		
Total[3]	646	100	646	100	646	100	646	100	646	100	646	100	646	100	646	100	646	100	27	4

[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 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

	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-	5	1	0	0	5	1	0	0	0	0	1	0	0	0	1	0	2	0	11	2
B-	8	2	0	0	18	4	0	0	0	0	19	4	16	4	18	4	20	5		
A	424	95	0	0	383	86	0	0	0	0	403	91	425	96	422	95	421	95		
B+	6	1	0	0	36	8	0	0	0	0	20	5	3	1	3	1	1	0		
C+	1	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	4	1
Small N[2]	0	0	444	100	0	0	444	100	444	100	0	0	0	0	0	0	0	0		
Total[3]	444	100	444	100	444	100	444	100	444	100	444	100	444	100	444	100	444	100	15	3

[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 percentages may differ slightly from the sum of the category percentages due to rounding

Appendix 6.J: Summary Statistics for Field-Test Items—March 2009

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	970	0.58	0.90	-1.94	3.93	0.37	0.13	-0.15	0.59
RC	274	0.64	0.74	-1.49	2.92	0.38	0.12	-0.03	0.59
RL	223	0.71	0.85	-1.45	3.25	0.35	0.13	-0.06	0.59
RW	64	0.44	0.89	-1.15	2.43	0.39	0.12	0.05	0.57
WC	288	0.30	1.04	-1.94	3.84	0.38	0.13	-0.15	0.59
WS	121	0.96	0.80	-0.60	3.93	0.34	0.13	0.00	0.57

[1] RC=Reading Comprehension, RL=Literary Response & Analysis, RW=Word Analysis, WA=Writing Applications, 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	672	0.66	0.93	-2.06	4.02	0.39	0.13	-0.17	0.66
A1	100	1.17	0.77	-1.36	2.95	0.33	0.14	-0.05	0.59
AF	163	0.53	0.92	-1.76	2.93	0.41	0.11	-0.01	0.60
MG	141	0.86	0.98	-1.60	4.02	0.36	0.15	-0.17	0.62
MR	92	0.77	0.85	-1.46	2.58	0.35	0.13	-0.01	0.62
NS	147	0.60	0.86	-1.17	3.45	0.41	0.12	0.00	0.66
PS	121	0.27	0.88	-2.06	3.01	0.40	0.09	0.08	0.58

[1] A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning (Items in this category are also classified under one of the other identified strands), NS=Number Sense, PS=Probability & Statistics

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

IRT review category	Status	Item count	Percent[1]
A	Use	143	15
B	Use	236	24
C	Use	412	42
D	Review	61	6
F	Do not use	87	9
N/A	Dropped from Calibration	31	3
Total		970	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
A	Use	138	21
B	Use	162	24
C	Use	251	37
D	Review	27	4
F	Do not use	77	11
N/A	Dropped from Calibration	17	3
Total		672	100

Table 6.J.5: Field-Test Items Containing Significant DIF

Test	Accession #	CAHSEE ID	Form	Item #	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	VE046796	L6OSA782	1	74	A	S	C-	S	S	C-	A	A	A
ELA	VE045358	L636J003	2	47	C-	S	S	S	S	A	B-	A	A
ELA	VE045359	L636J004	2	49	C-	S	S	S	S	A	A	A	A
ELA	VE045364	L636J009	3	50	C-	S	S	S	S	A	A	S	A
ELA	VE046482	L6OSA1189	3	74	A	S	S	S	S	C+	A	S	A
ELA	VE046836	L6OSA822	5	73	A	S	S	S	S	C+	A	C+	A
ELA	VE046452	L6OSA1159	7	74	A	S	S	S	S	A	A	C-	B-
ELA	VE046744	L6OSA730	8	74	A	S	S	S	S	C+	A	A	A
ELA	VE046743	L6OSA729	9	73	A	S	S	S	S	C+	A	S	A
ELA	VE045110	L626J001	11	46	C-	S	S	S	S	A	A	S	A
ELA	VE045117	L626J008	12	49	A	S	S	S	S	C+	A	A	A
ELA	VE045452	L639J014	14	49	A	S	S	S	S	C+	A	S	A
ELA	VE046458	L6OSA1165	14	73	A	S	S	S	S	A	C-	S	C-
ELA	VE046454	L6OSA1161	16	73	C+	S	S	S	S	A	A	A	A
ELA	VE046747	L6OSA733	17	73	A	S	C+	S	S	B+	A	A	A
ELA	VE046748	L6OSA734	17	74	A	S	C+	S	S	B+	A	A	A
ELA	VE046750	L6OSA736	18	74	B+	S	S	S	S	C+	A	A	A
ELA	VE046379	L6OSA1028	19	73	A	S	S	S	S	A	B-	C-	A
ELA	VE046411	L6OSA1066	20	74	C+	S	S	S	S	A	A	A	A
ELA	VE046108	L683H001	22	46	A	S	S	S	S	C-	B-	A	A
ELA	VE046118	L683H011	22	48	A	S	S	S	S	A	C-	A	A
ELA	VE046753	L6OSA739	27	74	A	S	S	S	S	C+	A	A	A
ELA	VE045195	L629J001	29	46	A	S	C-	S	S	C-	A	A	B-
ELA	VE045198	L629J004	31	47	A	S	S	S	S	A	A	C-	A
ELA	VE046723	L6OSA709	33	74	A	S	S	S	S	A	C-	S	B-
ELA	VE046724	L6OSA710	34	73	A	S	S	S	S	C-	C-	C-	B-
ELA	VE045329	L635J001	36	46	A	S	S	S	S	A	A	S	C-

Test	Accession #	CAHSEE ID	Form	Item #	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	VE046760	L6OSA746	36	73	B+	S	S	S	S	C+	A	S	A
ELA	VE045331	L635J003	37	46	A	S	S	S	S	C+	A	A	A
ELA	VE045341	L635J013	37	50	A	S	S	S	S	A	A	C-	A
ELA	VE046726	L6OSA712	40	73	A	S	S	S	S	C-	B-	A	A
ELA	VE046729	L6OSA715	41	74	A	S	S	S	S	A	C-	S	A
ELA	VE046731	L6OSA717	42	74	A	S	S	S	S	A	A	S	C-
ELA	VE045225	L630J003	43	47	C+	S	S	S	S	A	A	S	A
ELA	VE044769	L614H001	46	46	B-	S	S	S	S	A	B-	S	C-
ELA	VE044777	L614H009	47	47	C-	S	S	S	S	A	A	A	A
ELA	VE044770	L614H002	48	46	C-	S	S	S	S	B+	A	S	B+
ELA	VE046073	L681H001	49	46	A	S	S	S	S	C-	A	S	A
ELA	VE046076	L681H004	49	47	A	S	S	S	S	B-	B-	S	C-
ELA	VE046733	L6OSA719	49	74	A	S	S	S	S	C-	A	S	B-
ELA	VE046717	L6OSA703	50	74	A	S	S	S	S	C-	C-	A	C-
ELA	VE046736	L6OSA722	51	73	A	S	S	S	S	A	A	S	C-
ELA	VE044817	L615J009	52	49	C+	S	S	S	S	A	A	S	A
ELA	VE045757	L657H014	55	46	C-	S	S	S	S	B-	B-	B-	A
ELA	VE045756	L657H013	56	46	C-	S	S	S	S	A	A	A	A
ELA	VE046809	L6OSA795	57	73	A	S	S	S	S	A	A	A	C-
ELA	VE046778	L6OSA764	59	73	A	S	S	S	S	A	C-	A	C-
ELA	VE046780	L6OSA766	60	74	A	S	S	S	S	A	C-	C-	A
ELA	VE046813	L6OSA799	62	73	A	S	S	S	S	C+	A	S	A
ELA	VE044959	L620J005	71	46	C-	S	S	S	S	A	A	S	A
ELA	VE044958	L620J004	71	49	A	S	S	S	S	C+	A	S	A
ELA	VE044965	L620J012	71	50	A	S	S	S	S	B-	C-	S	B-
ELA	VE046872	L6OSA858	71	73	A	S	S	S	S	C+	A	S	A
ELA	VE046873	L6OSA859	71	74	A	S	S	S	S	C+	A	S	A
ELA	VE044703	L611J011	75	47	A	S	S	S	S	C-	A	S	A

Test	Accession #	CAHSEE ID	Form	Item #	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	VE044698	L611J006	75	48	A	S	S	S	S	C-	B-	S	B-
ELA	VE046304	L696H006	79	49	A	S	S	S	S	A	C-	A	B-
ELA	VE046922	L6OSA908	81	74	B+	S	S	S	S	A	A	S	C-
ELA	VE046923	L6OSA909	82	73	B+	S	S	S	S	A	A	S	C-
ELA	VE044655	L610H001	84	46	C-	S	A	S	S	B-	C-	B-	A
ELA	VE044753	L613J002	89	47	A	S	S	S	S	A	A	A	C-
ELA	VE045459	L640H007	94	49	A	S	S	S	S	A	C-	S	C-
ELA	VE046313	L697H001	97	48	C-	S	S	S	S	B-	A	S	A
ELA	VE046937	L6OSA923	97	73	A	S	S	S	S	C+	A	S	B+
ELA	VE045383	L637H014	102	46	A	S	A	S	S	A	B-	B-	C-
ELA	VE045378	L637H009	102	50	A	S	C+	S	S	C+	A	A	A
ELA	VE046909	L6OSA895	106	73	A	S	S	S	S	C+	A	S	A
ELA	VE046915	L6OSA901	111	74	A	S	S	S	S	C+	A	S	A
ELA	VE044666	L610J001	113	46	C-	S	S	S	S	A	B-	S	A
ELA	VE045555	L644H002	115	46	A	S	S	S	S	C-	A	S	A
ELA	VE046951	L6OSA937	120	74	A	S	S	S	S	B-	C-	S	A
ELA	VE046042	L678H010	122	50	A	S	S	S	S	A	B-	S	C-
ELA	VE046128	L684H005	126	48	A	S	S	S	S	B-	C-	C-	A
ELA	VE046127	L684H004	127	47	C-	S	S	S	S	B-	A	S	A
ELA	VE046626	L6OSA1334	127	74	C+	S	S	S	S	A	A	S	A
ELA	VE045954	L672H001	128	46	B-	S	S	S	S	A	C-	S	B-
ELA	VE045956	L672H003	128	47	A	S	S	S	S	C+	B+	S	A
ELA	VE044738	L613H001	132	46	A	S	S	S	S	C-	A	A	B-
ELA	VE046690	L6OSA1397	132	73	A	S	S	S	S	C+	A	A	A
ELA	VE046689	L6OSA1396	132	74	A	S	S	S	S	C-	A	B-	A
ELA	VE046688	L6OSA1395	133	73	A	S	S	S	S	C+	A	B+	A
ELA	VE046444	L6OSA1151	134	73	A	S	S	S	S	C-	A	B-	A
ELA	VE046687	L6OSA1394	134	74	A	S	S	S	S	C-	A	A	A

Test	Accession #	CAHSEE ID	Form	Item #	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	VE045915	L669H012	136	50	A	S	S	S	S	A	A	S	A
ELA	VE046442	L6OSA1149	139	73	A	S	S	S	S	C+	B-	S	A
ELA	VE046443	L6OSA1150	139	74	A	S	S	S	S	C-	A	S	A
ELA	VE044919	L619H006	140	47	A	S	S	S	S	A	B-	S	C-
ELA	VE046450	L6OSA1157	141	74	A	S	S	S	S	C-	A	S	A
ELA	VE046568	L6OSA1275	144	74	A	S	S	S	S	A	A	S	C+
MATH	VE047349	M50245	3	19	A	S	A	S	S	A	A	C-	A
MATH	VE047220	M50115	3	57	C-	S	C-	S	S	C-	C-	C-	A
MATH	VE048242	M51142	12	57	A	S	C-	S	S	C-	C-	B-	B-
MATH	VE047368	M50264	14	36	A	S	C-	S	S	C-	B-	A	A
MATH	VE047369	M50265	16	36	A	S	A	S	S	A	A	A	C-
MATH	VE047264	M50160	17	36	A	S	A	S	S	A	B-	C-	A
MATH	VE047201	M50096	19	20	C-	S	A	S	S	A	A	A	A
MATH	VE047379	M50275	20	19	C-	S	A	S	S	A	A	A	A
MATH	VE047457	M50354	21	18	C-	S	A	S	S	A	A	B-	A
MATH	VE047372	M50268	24	57	A	S	A	S	S	A	C-	A	C-
MATH	VE047338	M50234	25	21	C-	S	A	S	S	B-	B-	B-	A
MATH	VE047344	M50240	27	19	C-	S	A	S	S	A	A	A	A
MATH	VE047375	M50271	27	36	A	S	B-	S	S	B-	B-	A	C-
MATH	VE047819	M50717	29	57	A	S	A	S	S	A	A	C-	A
MATH	VE048219	M51119	33	79	A	S	C+	S	S	C+	A	A	A
MATH	VE048742	M51641	42	19	A	S	C+	S	S	B+	A	A	A
MATH	VE047723	M50621	44	58	A	S	C-	S	S	B-	A	A	A
MATH	VC173154	M31055	46	20	B-	S	A	S	S	A	A	B-	C-
MATH	VE048156	M51056	46	60	A	S	C+	S	S	B+	A	A	A
MATH	VE047126	M50021	47	36	A	S	C+	S	S	B+	A	A	A
MATH	VE048616	M51515	48	79	A	S	C+	S	S	B+	A	A	A
MATH	VC024722	M21652	49	80	A	S	C+	S	S	C+	A	A	A

Test	Accession #	CAHSEE ID	Form	Item #	Male-Female	White-American Indian	White-Asian	White-Pacific Islander	White-Filipino	White-Combined Asian	White-Hispanic	White-African American	Native English-English Learner
MATH	VE048696	M51595	52	18	A	S	C+	S	S	A	A	A	A
MATH	VE047525	M50422	56	21	A	S	B-	S	S	B-	B-	C-	C-

Note: S indicates that DIF was not performed due to insufficient sample size.

Table 6.J.6: Distribution of Field-Test 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 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-	13	1	0	0	2	0	0	0	0	0	17	2	15	2	7	1	16	2	58	6
B-	19	2	0	0	2	0	0	0	0	0	44	5	61	6	32	3	68	7		
A	889	92	0	0	41	4	0	0	0	0	827	85	886	91	431	44	881	91		
B+	44	5	0	0	2	0	0	0	0	0	60	6	8	1	7	1	4	0		
C+	5	1	0	0	3	0	0	0	0	0	22	2	0	0	1	0	1	0	29	3
Small N2	0	0	970	100	920	95	970	100	970	100	0	0	0	0	492	51	0	0		
Total3	970	100	970	100	970	100	970	100	970	100	970	100	970	100	970	100	970	100	87	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 percentages 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-	6	1	0	0	4	1	0	0	0	0	3	0	3	0	5	1	5	1	17	3
B-	33	5	0	0	24	4	0	0	0	0	22	3	15	2	27	4	34	5		
A	622	93	0	0	583	87	0	0	0	0	620	92	653	97	640	95	632	94		
B+	11	2	0	0	54	8	0	0	0	0	25	4	1	0	0	0	1	0		
C+	0	0	0	0	7	1	0	0	0	0	2	0	0	0	0	0	0	0	7	1
Small N[2]	0	0	672	100	0	0	672	100	672	100	0	0	0	0	0	0	0	0		
Total[3]	672	100	672	100	672	100	672	100	672	100	672	100	672	100	672	100	672	100	24	4

[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 percentages 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 2008

	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	0.67	0.33	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.83	0.40	0.54	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.83	0.39	0.53	0.65	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.71	0.30	0.37	0.52	0.49	1.00	-	-	-	-	-	-	-
Writing Conventions	0.76	0.40	0.41	0.51	0.50	0.50	1.00	-	-	-	-	-	-
Mathematics	0.58	0.36	0.38	0.50	0.48	0.49	0.49	1.00	-	-	-	-	-
Probability and Statistics	0.53	0.32	0.36	0.47	0.45	0.44	0.42	0.76	1.00	-	-	-	-
Number Sense	0.49	0.30	0.35	0.42	0.41	0.39	0.41	0.79	0.54	1.00	-	-	-
Algebra & Functions	0.52	0.33	0.34	0.44	0.42	0.43	0.45	0.85	0.56	0.58	1.00	-	-
Measurement & Geometry	0.48	0.28	0.31	0.42	0.39	0.41	0.40	0.80	0.51	0.51	0.58	1.00	-
Algebra 1	0.28	0.19	0.14	0.22	0.21	0.27	0.26	0.60	0.30	0.36	0.43	0.38	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	47.09	1.99	4.14	9.83	10.92	5.05	8.18	37.06	6.73	8.25	9.65	8.11	4.32
SD	12.44	0.53	1.63	3.37	3.61	2.37	3.12	11.23	2.65	3.06	3.44	3.24	2.15
Reliability	0.87	0.26	0.45	0.68	0.68	0.55	0.68	0.87	0.62	0.62	0.65	0.64	0.45
SEM	4.40	0.46	1.21	1.91	2.03	1.59	1.76	4.12	1.63	1.89	2.04	1.95	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 13,369 examinees completing the ELA section.
Correlations for the Mathematics section are reported for 13,235 examinees completing the Mathematics section.
Correlations between ELA and Mathematics are reported for 5,993 examinees taking both sections.

Table 6.K.2: Intercorrelations and Reliability Estimates by Section—October 2008

	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	0.85	0.44	0.59	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.86	0.44	0.58	0.71	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.76	0.38	0.43	0.58	0.59	1.00	-	-	-	-	-	-	-
Writing Conventions	0.78	0.43	0.46	0.57	0.58	0.58	1.00	-	-	-	-	-	-
Mathematics	0.64	0.41	0.43	0.56	0.53	0.54	0.56	1.00	-	-	-	-	-
Probability and Statistics	0.59	0.38	0.43	0.54	0.51	0.47	0.50	0.75	1.00	-	-	-	-
Number Sense	0.52	0.33	0.36	0.45	0.43	0.42	0.45	0.79	0.51	1.00	-	-	-
Algebra & Functions	0.57	0.37	0.39	0.49	0.47	0.49	0.51	0.87	0.58	0.59	1.00	-	-
Measurement & Geometry	0.51	0.33	0.33	0.45	0.42	0.45	0.46	0.83	0.50	0.56	0.64	1.00	-
Algebra1	0.42	0.28	0.26	0.35	0.33	0.37	0.39	0.71	0.37	0.47	0.56	0.54	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	49.81	2.03	4.36	10.73	11.75	5.42	8.39	38.45	7.57	8.87	10.03	7.65	4.32
SD	14.04	0.70	1.67	3.62	3.88	2.53	3.13	12.22	2.89	2.95	3.79	3.39	2.31
Reliability	0.89	0.29	0.53	0.73	0.73	0.61	0.70	0.89	0.69	0.61	0.71	0.69	0.56
SEM	4.63	0.59	1.14	1.87	2.01	1.58	1.71	4.04	1.60	1.84	2.05	1.89	1.53

Note: All correlations are for raw scores. Students who took Braille, Large Print, and Audio CD forms are not included in the computation.
Correlations for the ELA section are reported for 43,931 examinees completing the ELA section.
Correlations for the Mathematics section are reported for 44,943 examinees completing the Mathematics section.
Correlations between ELA and Mathematics are reported for 25,920 examinees taking both sections.

Table 6.K.5: Intercorrelations and Reliability Estimates by Section—February 2009

	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	0.77	0.49	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.90	0.58	0.68	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.91	0.59	0.69	0.80	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.88	0.57	0.65	0.75	0.77	1.00	-	-	-	-	-	-	-
Writing Conventions	0.87	0.58	0.62	0.71	0.73	0.75	1.00	-	-	-	-	-	-
Mathematics	0.81	0.57	0.62	0.72	0.74	0.72	0.74	1.00	-	-	-	-	-
Probability and Statistics	0.75	0.52	0.59	0.68	0.69	0.67	0.67	0.86	1.00	-	-	-	-
Number Sense	0.71	0.51	0.55	0.64	0.65	0.63	0.65	0.90	0.73	1.00	-	-	-
Algebra & Functions	0.77	0.54	0.58	0.69	0.70	0.69	0.70	0.93	0.77	0.79	1.00	-	-
Measurement & Geometry	0.72	0.51	0.54	0.64	0.65	0.65	0.65	0.90	0.72	0.75	0.80	1.00	-
Algebra 1	0.66	0.49	0.49	0.58	0.60	0.60	0.61	0.86	0.65	0.72	0.76	0.73	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	62.05	2.40	5.18	13.19	14.46	7.77	10.62	51.77	9.06	11.58	13.60	11.02	6.51
SD	16.19	0.64	1.63	3.58	4.26	3.08	3.41	16.83	3.05	3.81	4.58	4.09	3.30
Reliability	0.94	0.44	0.60	0.79	0.83	0.79	0.81	0.95	0.77	0.81	0.84	0.81	0.80
SEM	3.99	0.48	1.03	1.64	1.76	1.41	1.50	3.71	1.45	1.67	1.82	1.76	1.48

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 195,015 examinees completing the ELA section.
Correlations for the Mathematics section are reported for 196,319 examinees completing the Mathematics section.
Correlations between ELA and Mathematics are reported for 168,323 examinees taking both sections.

Table 6.K.6: Intercorrelations and Reliability Estimates by Section—March 2009

	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	0.75	1.00	-	-	-	-	-	-	-	-	-	-	-
Word Analysis	0.71	0.46	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.90	0.57	0.62	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.90	0.58	0.62	0.79	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.85	0.55	0.53	0.71	0.71	1.00	-	-	-	-	-	-	-
Writing Conventions	0.87	0.58	0.55	0.71	0.72	0.72	1.00	-	-	-	-	-	-
Mathematics	0.78	0.56	0.53	0.70	0.67	0.69	0.70	1.00	-	-	-	-	-
Probability and Statistics	0.70	0.50	0.48	0.64	0.61	0.61	0.63	0.85	1.00	-	-	-	-
Number Sense	0.70	0.50	0.48	0.64	0.61	0.61	0.63	0.89	0.73	1.00	-	-	-
Algebra & Functions	0.73	0.52	0.49	0.66	0.64	0.65	0.66	0.93	0.75	0.79	1.00	-	-
Measurement & Geometry	0.69	0.50	0.46	0.62	0.59	0.62	0.63	0.90	0.72	0.74	0.78	1.00	-
Algebra 1	0.64	0.49	0.42	0.57	0.54	0.58	0.58	0.86	0.65	0.70	0.76	0.73	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	63.59	2.46	5.14	13.10	15.45	7.97	10.84	54.21	9.24	12.14	14.07	11.79	6.97
SD	15.11	0.65	1.53	3.69	3.84	2.75	3.19	16.26	2.90	3.63	4.55	4.07	3.08
Reliability	0.93	0.44	0.50	0.80	0.81	0.72	0.77	0.95	0.75	0.80	0.85	0.82	0.77
SEM	3.98	0.48	1.08	1.63	1.65	1.46	1.52	3.63	1.46	1.62	1.76	1.71	1.48

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 372,536 examinees completing the ELA section.
Correlations for the Mathematics section are reported for 371,580 examinees completing the Mathematics section.
Correlations between ELA and Mathematics are reported for 339,010 examinees taking both sections.

Table 6.K.7: Intercorrelations and Reliability Estimates by Section—May 2009

	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	0.73	0.41	1.00	-	-	-	-	-	-	-	-	-	-
Reading Comprehension	0.86	0.49	0.61	1.00	-	-	-	-	-	-	-	-	-
Literary Responses & Analysis	0.88	0.50	0.63	0.73	1.00	-	-	-	-	-	-	-	-
Writing Strategies	0.80	0.45	0.51	0.63	0.63	1.00	-	-	-	-	-	-	-
Writing Conventions	0.82	0.49	0.53	0.60	0.64	0.63	1.00	-	-	-	-	-	-
Mathematics	0.70	0.48	0.49	0.62	0.61	0.61	0.59	1.00	-	-	-	-	-
Probability and Statistics	0.63	0.44	0.47	0.56	0.56	0.53	0.52	0.80	1.00	-	-	-	-
Number Sense	0.57	0.39	0.42	0.51	0.50	0.50	0.47	0.81	0.59	1.00	-	-	-
Algebra & Functions	0.64	0.44	0.44	0.56	0.56	0.56	0.55	0.89	0.65	0.62	1.00	-	-
Measurement & Geometry	0.60	0.41	0.42	0.53	0.52	0.54	0.52	0.85	0.59	0.59	0.69	1.00	-
Algebra 1	0.48	0.35	0.30	0.41	0.41	0.43	0.42	0.74	0.47	0.51	0.59	0.57	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	50.26	2.12	4.67	10.18	11.95	5.53	8.39	38.20	7.03	8.09	9.96	8.53	4.59
SD	15.01	0.61	1.74	3.73	4.15	2.76	3.40	13.55	2.81	3.44	4.07	3.63	2.47
Reliability	0.91	0.35	0.59	0.75	0.78	0.68	0.75	0.91	0.67	0.70	0.75	0.72	0.60
SEM	4.38	0.49	1.12	1.88	1.96	1.57	1.71	4.08	1.62	1.88	2.04	1.93	1.55

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 42,523 examinees completing the ELA section.
Correlations for the Mathematics section are reported for 22,139 examinees completing the Mathematics section.
Correlations between ELA and Mathematics are reported for 42,754 examinees taking both sections.

Table 6.K.8: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for ELA—February 2009

	Total MC N=195,091		Composite (MC + Essay) N=195,091	
	REL	SEM	REL	SEM
Gender				
Male	0.94	3.43	0.94	4.10
Female	0.94	3.27	0.94	3.84
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.39	0.93	4.09
Asian	0.95	3.12	0.94	3.84
Pacific Islander	0.93	3.44	0.93	4.13
Filipino	0.93	3.10	0.92	3.76
Hispanic or Latino	0.93	3.60	0.93	4.13
African American	0.94	3.56	0.93	4.18
White (not of Hispanic origin)	0.94	2.95	0.93	3.71
Language Fluency				
English Proficient Students	0.94	3.17	0.93	3.85
English Learner Students	0.88	3.88	0.88	4.37
Special Education Program Participation				
Students Receiving Services	0.92	3.82	0.91	4.47
Students Not Receiving Services	0.94	3.29	0.93	3.92

Table 6.K.9: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—February 2009

	Total N=196,387	
	REL	SEM
Gender		
Male	0.95	3.68
Female	0.95	3.73
Race/Ethnicity		
American Indian or Alaska Native	0.94	3.81
Asian	0.96	3.12
Pacific Islander	0.94	3.78
Filipino	0.94	3.44
Hispanic or Latino	0.94	3.91
African American	0.93	3.96
White (not of Hispanic origin)	0.95	3.41
Language Fluency		
English Proficient Students	0.95	3.59
English Learner Students	0.90	4.07
Special Education Program Participation		
Students Receiving Services	0.91	4.06
Students Not Receiving Services	0.95	3.66

Table 6.K.10: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for ELA—March 2009

	Total MC N=372,707		Composite (MC + Essay) N=372,707	
	REL	SEM	REL	SEM
Gender				
Male	0.94	3.40	0.93	4.09
Female	0.93	3.25	0.93	3.84
Race/Ethnicity				
American Indian or Alaska Native	0.94	3.34	0.93	4.09
Asian	0.94	2.98	0.93	3.71
Pacific Islander	0.93	3.40	0.92	4.03
Filipino	0.92	3.12	0.91	3.78
Hispanic or Latino	0.92	3.52	0.92	4.09
African American	0.93	3.53	0.93	4.17
White (not of Hispanic origin)	0.93	2.97	0.91	3.72
Language Fluency				
English Proficient Students	0.92	3.17	0.92	3.85
English Learner Students	0.89	3.82	0.88	4.34
Special Education Program Participation				
Students Receiving Services	0.92	3.80	0.91	4.48
Students Not Receiving Services	0.93	3.27	0.92	3.91

Table 6.K.11: Reliabilities (REL) and Standard Errors of Measurement (SEM) of Subgroups for Mathematics—March 2009

	Total N=371,704	
	REL	SEM
Gender		
Male	0.95	3.59
Female	0.95	3.65
Race/Ethnicity		
American Indian or Alaska Native	0.95	3.71
Asian	0.95	2.86
Pacific Islander	0.94	3.68
Filipino	0.94	3.34
Hispanic or Latino	0.94	3.82
African American	0.94	3.91
White (not of Hispanic origin)	0.95	3.33
Language Fluency		
English Proficient Students	0.95	3.51
English Learner Students	0.92	4.03
Special Education Program Participation		
Students Receiving Services	0.92	4.05
Students Not Receiving Services	0.95	3.58

Appendix 6.L: Rater Agreement Analyses

Table 6.L.1: Agreement of First and Second Ratings on the ELA Essay Item—July 2008

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	423	0	0	0	0	423	3
1	0	596	362	2	0	960	7
2	0	326	9185	851	2	10364	78
3	0	0	836	696	34	1566	12
4	0	0	4	27	25	56	0
Total	423	922	10387	1576	61	13369	100
Percent	3	7	78	12	0	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	10925	82	82
1	2436	18	100
2	8	0	100
3	0	0	100
4	0	0	100

Table 6.L.2: Summary Statistics for the ELA Essay Item—July 2008

	First Rating	Second Rating
Mean	1.99	1.99
Standard Deviation	0.58	0.58
Mean Absolute Difference Between First and Second Ratings: 0.18		
Correlation of First and Second Ratings: 0.72		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	205
Illegible	0
Off Topic	199
Cartoon / Inappropriate	18
Not in English	1
Total	**423**

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 2008

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	2857	0	0	0	0	2857	7
1	0	1080	877	6	0	1963	4
2	0	893	25926	3618	66	30503	69
3	0	12	3749	3835	343	7939	18
4	0	0	61	363	245	669	2
Total	2857	1985	30613	7822	654	43931	100
Percent	7	5	70	18	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	33943	77	77
1	9843	22	100
2	145	0	100
3	0	0	100
4	0	0	100

Table 6.L.4: Summary Statistics for the ELA Essay Item—October 2008

	First Rating	Second Rating
Mean	2.04	2.03
Standard Deviation	0.74	0.74
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.78		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	924
Illegible	0
Off Topic	1859
Cartoon / Inappropriate	64
Not in English	10
Total	**2857**

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 2008

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	3756	0	15	2	0	3773	3
1	39	3656	2463	9	0	6167	6
2	39	2364	67218	8483	163	78267	72
3	3	13	8703	9228	805	18752	17
4	0	0	161	838	687	1686	2
Total	3837	6033	78560	18560	1655	108645	100
Percent	4	6	72	17	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	84545	78	78
1	23695	22	100
2	400	0	100
3	5	0	100
4	0	0	100

Table 6.L.6: Summary Statistics for the ELA Essay Item—November 2008

	First Rating	Second Rating
Mean	2.08	2.08
Standard Deviation	0.65	0.65
Mean Absolute Difference Between First and Second Ratings: 0.23		
Correlation of First and Second Ratings: 0.72		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	2451
Illegible	0
Off Topic	1149
Cartoon / Inappropriate	182
Not in English	30
Total	**3812**

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 2008

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	71	0	0	0	0	71	2
1	0	110	70	1	0	181	5
2	0	83	2326	268	2	2679	76
3	0	1	283	267	21	572	16
4	0	0	1	23	17	41	1
Total	71	194	2680	559	40	3544	100
Percent	2	5	76	16	1	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	2791	79	79
1	748	21	100
2	5	0	100
3	0	0	100
4	0	0	100

Table 6.L.8: Summary Statistics for the ELA Essay Item—December 2008

	First Rating	Second Rating
Mean	2.09	2.09
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	44
Illegible	0
Off Topic	26
Cartoon / Inappropriate	1
Not in English	0
Total	**71**

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 2009

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	3092	0	0	0	0	3092	2
1	0	2868	2212	26	2	5108	3
2	0	2174	82241	21527	630	106572	55
3	0	35	21595	42882	5503	70015	36
4	0	0	724	5605	3899	10228	5
Total	3092	5077	106772	70040	10034	195015	100
Percent	2	3	55	36	5	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	134982	69	69
1	58616	30	99
2	1415	1	100
3	2	0	100
4	0	0	100

Table 6.L.10: Summary Statistics for the ELA Essay Item—February 2009

	First Rating	Second Rating
Mean	2.41	2.40
Standard Deviation	0.70	0.70
Mean Absolute Difference Between First and Second Ratings: 0.32		
Correlation of First and Second Ratings: 0.67		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	2183
Illegible	0
Off Topic	753
Cartoon / Inappropriate	127
Not in English	29
Total	**3092**

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 2009

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	5279	0	0	0	0	5279	1
1	0	5456	3672	47	7	9182	2
2	0	3557	141616	41254	1155	187582	50
3	0	34	41217	95210	12062	148523	40
4	0	1	1235	12459	8275	21970	6
Total	5279	9048	187740	148970	21499	372536	100
Percent	1	2	50	40	6	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	255836	69	69
1	114221	31	99
2	2471	1	100
3	8	0	100
4	0	0	100

Table 6.L.12: Summary Statistics for the ELA Essay Item—March 2009

	First Rating	Second Rating
Mean	2.46	2.46
Standard Deviation	0.71	0.71
Mean Absolute Difference Between First and Second Ratings: 0.32		
Correlation of First and Second Ratings: 0.67		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	3922
Illegible	0
Off Topic	1105
Cartoon / Inappropriate	200
Not in English	52
Total	**5279**

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 2009

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	1288	0	0	0	0	1288	3
1	0	1486	907	5	0	2398	6
2	0	977	24934	3697	53	29661	69
3	0	6	3733	4632	363	8734	20
4	0	2	38	324	309	673	2
Total	1288	2471	29612	8658	725	42754	100
Percent	3	6	69	20	2	100	

Differences Between First and Second Ratings on Essay

Difference	Frequency	Percent	Cumulative Percent
0	32649	76	76
1	10001	23	100
2	102	0	100
3	2	0	100
4	0	0	100

Table 6.L.14: Summary Statistics for the ELA Essay Item—May 2009

	First Rating	Second Rating
Mean	2.12	2.12
Standard Deviation	0.66	0.66
Mean Absolute Difference Between First and Second Ratings: 0.24		
Correlation of First and Second Ratings: 0.72		

Summary of Essays Receiving Final Score of Zero

Essay	**N**
Blank	713
Illegible	0
Off Topic	526
Cartoon / Inappropriate	36
Not in English	13
Total	**1288**

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 2008

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	13,171	0.56	0.23324	71.64
Rater (R)	1	0.15	0.00000	0.00
PR,*e*	13,171	0.09	0.09231	28.36
Generalizability Coefficient				0.83
Dependability Coefficient				0.83

Table 6.M.2: Generalizability Results—October 2008

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	42,782	0.97	0.42288	78.00
Rater (R)	1	0.37	0.00000	0
PR,*e*	42,782	0.0.12	0.11928	22.00
Generalizability Coefficient				0.88
Dependability Coefficient				0.88

Table 6.M.3: Generalizability Results—November 2008

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	105,792	0.73	0.30434	72.19
Rater (R)	1	0.31	0.00000	0.00
PR,*e*	105,792	0.12	0.11727	27.81
Generalizability Coefficient				0.84
Dependability Coefficient				0.84

Table 6.M.4: Generalizability Results—December 2008

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	3,444	0.54	0.21719	66.56
Rater (R)	1	0.08	0.00000	0.00
PR,e	3,444	0.11	0.10915	33.45
Generalizability Coefficient				0.80
Dependability Coefficient				0.80

Table 6.M.5: Generalizability Results—February 2009

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	191,369	0.81	0.32337	66.10
Rater (R)	1	0.24	0.00000	0.00
PR,*e*	191,369	0.17	0.16583	33.90
Generalizability Coefficient				0.80
Dependability Coefficient				0.80

Table 6.M.6: Generalizability Results—March 2009

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	367,815	0.83	0.32909	66.28
Rater (R)	1	0.19	0.00000	0.00
PR,*e*	367,815	0.17	0.16741	33.72
Generalizability Coefficient				0.80
Dependability Coefficient				0.80

Table 6.M.7: Generalizability Results—May 2009

Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	41,393	0.74	0.30895	71.57
Rater (R)	1	0.01	0.00000	0.00
PR,*e*	41,393	0.12	0.12271	28.43
Generalizability Coefficient				0.83
Dependability Coefficient				0.83

Appendix 6.N: Decision Classification Reliability Analyses

Table 6.N.1: ESEA Reliability Classifications—July 2008

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	0.00	0.02
Proficient (68-75)	0.00	0.02	0.01	0.03
Below Proficient (0-67)	0.00	0.01	0.94	0.95
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.99
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.97

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.00	0.01	0.01	0.02
Proficient (68-75)	0.00	0.02	0.02	0.03
Below Proficient (0-67)	0.00	0.02	0.93	0.95
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.99
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	0.03	0.01	0.04
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	0.04
Below Proficient (0-58)	0.00	0.02	0.94	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 2008

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.25	0.05	0.30
0-53	0.05	0.65	0.70
Estimated Proportion Correctly Classified			0.90

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
54-90	0.24	0.06	0.30
0-53	0.08	0.63	0.70
Estimated Proportion Consistently Classified			0.87

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.24	0.05	0.29
0-42	0.04	0.67	0.71
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Score	Pass
43-80	0.23	0.06	0.29
0-42	0.06	0.65	0.71
Estimated Proportion Consistently Classified			0.88

Table 6.N.3: ESEA Reliability Classifications—October 2008

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.02	0.01	0.01	0.04
Proficient (69-75)	0.00	0.03	0.02	0.05
Below Proficient (0-68)	0.00	0.02	0.89	0.91
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Correctly Classified for Proficient or Above vs. Below				0.96

[1] True Score

Decision Consistency

Classification on Alternate Form				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.02	0.01	0.01	0.04
Proficient (69-75)	0.00	0.02	0.02	0.05
Below Proficient (0-68)	0.00	0.03	0.88	0.91
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.00	0.01	0.01	0.02
Proficient (58-71)	0.00	0.05	0.01	0.05
Below Proficient (0-57)	0.00	0.02	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 (72-80)	0.00	0.01	0.00	0.02
Proficient (58-71)	0.00	0.04	0.01	0.05
Below Proficient (0-57)	0.00	0.03	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 2008

English-Language Arts
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.33	0.04	0.37
0-54	0.05	0.58	0.63
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
55-90	0.31	0.06	0.37
0-54	0.07	0.56	0.63
Estimated Proportion Consistently Classified			0.87

Mathematics
Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.29	0.04	0.33
0-42	0.04	0.63	0.67
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.28	0.06	0.33
0-42	0.06	0.61	0.67
Estimated Proportion Consistently Classified			0.89

Table 6.N.5: ESEA Reliability Classifications—November 2008

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	0.00	0.05
Proficient (70-76)	0.00	0.04	0.02	0.07
Below Proficient (0-69)	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				0.95

[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 (70-76)	0.01	0.04	0.02	0.07
Below Proficient (0-69)	0.00	0.04	0.84	0.88
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.93

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	0.91
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 (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	0.03	0.88	0.91
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Table 6.N.6: Pass/Not Pass Classifications—November 2008

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.36	0.04	0.40
0-56	0.04	0.56	0.60
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.34	0.06	0.40
0-56	0.06	0.54	0.60
Estimated Proportion Consistently Classified			0.88

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
42-80	0.35	0.03	0.38
0-41	0.07	0.55	0.62
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	0.33	0.04	0.38
0-41	0.09	0.53	0.62
Estimated Proportion Consistently Classified			0.87

Table 6.N.7: ESEA Reliability Classifications—December 2008

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 (70-76)	0.00	0.03	0.01	0.04
Below Proficient (0-69)	0.00	0.01	0.91	0.93
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 (70-76)	0.00	0.02	0.02	0.04
Below Proficient (0-69)	0.00	0.03	0.90	0.93
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (72-80)	0.00	0.00	0.01	0.01
Proficient (57-71)	0.00	0.02	0.02	0.05
Below Proficient (0-56)	0.00	0.00	0.94	0.94
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 (72-80)	0.00	0.00	0.01	0.01
Proficient (57-71)	0.00	0.02	0.02	0.05
Below Proficient (0-56)	0.00	0.02	0.93	0.94
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 2008

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.27	0.03	0.30
0-56	0.05	0.65	0.70
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.25	0.05	0.30
0-56	0.07	0.63	0.70
Estimated Proportion Consistently Classified			0.88

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
41-80	0.32	0.04	0.36
0-40	0.08	0.57	0.64
Estimated Proportion Correctly Classified			0.88

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
41-80	0.30	0.06	0.36
0-40	0.11	0.54	0.64
Estimated Proportion Consistently Classified			0.84

Table 6.N.9: ESEA Reliability Classifications—February 2009

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (76-90)	0.21	0.04	0.00	0.25
Proficient (69-75)	0.03	0.10	0.04	0.17
Below Proficient (0-68)	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)	0.20	0.04	0.01	0.25
Proficient (69-75)	0.04	0.08	0.05	0.17
Below Proficient (0-68)	0.00	0.04	0.54	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)	0.13	0.02	0.00	0.15
Proficient (58-71)	0.02	0.21	0.03	0.26
Below Proficient (0-57)	0.00	0.03	0.56	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	0.55	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 2009

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.64	0.02	0.67
0-55	0.04	0.30	0.33
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.63	0.04	0.67
0-55	0.05	0.29	0.33
Estimated Proportion Consistently Classified			0.91

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.64	0.02	0.67
0-42	0.03	0.30	0.33
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.63	0.04	0.67
0-42	0.05	0.29	0.33
Estimated Proportion Consistently Classified			0.92

Table 6.N.11: ESEA Reliability Classifications—March 2009

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (77-90)	0.19	0.04	0.00	0.23
Proficient (69-76)	0.03	0.15	0.04	0.22
Below Proficient (0-68)	0.00	0.04	0.51	0.55
Estimated Proportion Correctly Classified for Advanced vs. Proficient or Below				0.93
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 (77-90)	0.18	0.05	0.00	0.23
Proficient (69-76)	0.05	0.12	0.05	0.22
Below Proficient (0-68)	0.00	0.05	0.50	0.55
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.90
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)	0.14	0.03	0.00	0.17
Proficient (58-71)	0.02	0.25	0.03	0.30
Below Proficient (0-57)	0.00	0.03	0.50	0.53
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)	0.14	0.03	0.00	0.17
Proficient (58-71)	0.03	0.22	0.05	0.30
Below Proficient (0-57)	0.00	0.04	0.49	0.53
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.94
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.91

Table 6.N.12: Pass/Not Pass Classifications—March 2009

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.70	0.03	0.73
0-55	0.03	0.24	0.27
Estimated Proportion Correctly Classified			0.94

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
56-90	0.68	0.04	0.73
0-55	0.04	0.23	0.27
Estimated Proportion Consistently Classified			0.92

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.71	0.02	0.74
0-42	0.03	0.23	0.26
Estimated Proportion Correctly Classified			0.95

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.70	0.04	0.74
0-42	0.04	0.22	0.26
Estimated Proportion Consistently Classified			0.92

Table 6.N.13: ESEA Reliability Classifications—May 2009

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 (70-76)	0.00	0.04	0.02	0.06
Below Proficient (0-69)	0.00	0.02	0.87	0.89
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 (70-76)	0.01	0.03	0.02	0.06
Below Proficient (0-69)	0.00	0.03	0.86	0.89
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.97
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.94

Mathematics

Decision Accuracy

Classification on All-Forms Average[1]				
Placement (Raw Score)	Category			
	Advanced	Proficient	Below Proficient	Total
Advanced (73-80)	0.00	0.01	0.01	0.02
Proficient (59-72)	0.00	0.05	0.03	0.08
Below Proficient (0-58)	0.00	0.00	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 (73-80)	0.00	0.01	0.01	0.02
Proficient (59-72)	0.00	0.05	0.03	0.08
Below Proficient (0-58)	0.00	0.02	0.89	0.90
Estimated Proportion Consistently Classified for Advanced vs. Proficient or Below				0.98
Estimated Proportion Consistently Classified for Proficient or Above vs. Below				0.95

Table 6.N.14: Pass/Not Pass Classifications—May 2009

English-Language Arts

Decision Accuracy

Classification on All-Forms Average[1]			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.30	0.03	0.33
0-56	0.05	0.62	0.67
Estimated Proportion Correctly Classified			0.92

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
57-90	0.29	0.04	0.33
0-56	0.07	0.61	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
43-80	0.30	0.01	0.31
0-42	0.08	0.61	0.69
Estimated Proportion Correctly Classified			0.91

[1] True Score

Decision Consistency

Classification on Alternate Form			
Placement	Category		
(Raw Score)	Pass	Not Pass	Total
43-80	0.28	0.03	0.31
0-42	0.10	0.59	0.69
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 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	501.4882	450	23
88	479.6769	450	27
87	466.9779	450	26
86	457.5006	450	23
85	449.6730	450	21
84	442.8813	443	17
83	436.8058	437	17
82	431.2807	431	17
81	426.1948	426	15
80	421.4793	421	15
79	417.0847	417	14
78	412.9699	413	14
77	409.0980	409	13
76	405.4406	405 (Advanced)	13
75	401.9764	402	12
74	398.6775	399	12
73	395.5245	396	11
72	392.5007	393	11
71	389.5930	390	11
70	386.7887	387	11
69	384.0752	384	10
68	381.4429	381 (Proficient)	10
67	378.8856	379	10
66	376.3958	376	10
65	373.9670	374	10
64	371.5926	372	9
63	369.2707	369	9
62	366.9940	367	9
61	364.7596	365	9
60	362.5621	363	9
59	360.4009	360	9
58	358.2720	358	9
57	356.1706	356	9
56	354.0950	354	9
55	352.0442	352	9
54	350.0127	350 (Pass)	9
53	348.0042	348	9
52	346.0110	346	8
51	344.0349	344	8
50	342.0700	342	8
49	340.1172	340	8
48	338.1727	338	8
47	336.2376	336	8
46	334.3077	334	8
45	332.3855	332	8
44	330.4673	330	8
43	328.5516	329	8
42	326.6389	327	8
41	324.7271	325	9
40	322.8165	323	9
39	320.9049	321	9
38	318.9962	319	9
37	317.0835	317	9
36	315.1724	315	9
35	313.2584	313	9
34	311.3437	311	9
33	309.4272	309	9
32	307.5080	308	9
31	305.5862	306	9
30	303.6606	304	9
29	301.7321	302	9
28	299.7965	300	10
27	297.8517	298	10
26	295.8960	296	10
25	293.9255	294	10
24	291.9367	292	10
23	289.9266	290	10
22	287.8877	288	10
21	285.8120	286	11
20	283.6937	284	11
19	281.5226	282	11
18	279.2874	279	11
17	276.9630	277	11
16	274.5461	275	12
15	272.0191	275	12
14	269.3594	275	12
13	266.5415	275	13
12	263.5335	275	13
11	260.2955	275	14
10	256.7662	275	14
9	252.8732	275	15
8	248.5410	275	15
7	243.6509	275	16
6	238.031	275	21
5	231.4166	275	24
4	223.3696	275	27
3	213.0617	275	28
2	198.6427	275	26
1	174.1998	275	19
0	147.2238	275	8

Table 6.O.2: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, July 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	492.5628	450	20
78	469.3664	450	22
77	455.6077	450	19
76	445.6886	446	17
75	437.8708	438	16
74	431.3780	431	14
73	425.7957	426 (Advanced)	13
72	420.8776	421	12
71	416.4651	416	12
70	412.4467	412	11
69	408.7505	409	11
68	405.3154	405	10
67	402.1010	402	10
66	399.0717	399	10
65	396.1988	396	9
64	393.4615	393	9
63	390.8445	391	9
62	388.3328	388	9
61	385.9109	386	9
60	383.5708	384	9
59	381.3017	381 (Proficient)	8
58	379.0980	379	8
57	376.9494	377	8
56	374.8541	375	8
55	372.8034	373	8
54	370.7940	371	8
53	368.8222	369	8
52	366.8786	367	8
51	364.9677	365	8
50	363.0855	363	8
49	361.2224	361	8
48	359.3811	359	8
47	357.5556	358	8
46	355.7457	356	8
45	353.9484	354	8
44	352.1607	352	8
43	350.3794	350 (Pass)	8
42	348.6064	349	8
41	346.8351	347	8
40	345.0655	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.2961	343	8
38	341.5238	342	8
37	339.7475	340	8
36	337.9631	338	8
35	336.1702	336	8
34	334.3662	334	8
33	332.5500	333	8
32	330.7179	331	8
31	328.8671	329	8
30	326.9956	327	8
29	325.0988	325	8
28	323.1755	323	8
27	321.2239	321	8
26	319.2366	319	8
25	317.2115	317	8
24	315.1445	315	8
23	313.0296	313	8
22	310.8629	311	8
21	308.6387	309	9
20	306.3495	306	9
19	303.9858	304	9
18	301.5385	302	9
17	298.9996	299	9
16	296.3530	296	9
15	293.5872	294	10
14	290.6830	291	10
13	287.6175	288	10
12	284.3650	284	10
11	280.8906	281	11
10	277.1502	277	11
9	273.0880	275	12
8	268.6270	275	12
7	263.6547	275	13
6	258.0145	275	14
5	251.4607	275	14
4	243.5772	275	20
3	233.5855	275	20
2	219.7346	275	18
1	196.4412	275	13
0	171.7079	275	10

Table 6.O.3: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, October 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	501.9302	450	23
88	478.8967	450	27
87	465.0876	450	26
86	454.9495	450	23
85	446.7440	447	21
84	439.7542	440	17
83	433.6027	434	17
82	428.0848	428	17
81	423.0571	423	15
80	418.4305	418	14
79	414.1394	414	14
78	410.1304	410	13
77	406.3613	406	13
76	402.8086	403 (Advanced)	12
75	399.4370	399	12
74	396.2237	396	12
73	393.1490	393	11
72	390.1977	390	11
71	387.3564	387	11
70	384.6101	385	11
69	381.9491	382 (Proficient)	10
68	379.3638	379	10
67	376.8496	377	10
66	374.3972	374	10
65	372.0004	372	10
64	369.6558	370	10
63	367.3586	367	9
62	365.1029	365	9
61	362.8858	363	9
60	360.7047	361	9
59	358.5573	359	9
58	356.4386	356	9
57	354.3477	354	9
56	352.2820	352	9
55	350.2369	350 (Pass)	9
54	348.2173	348	9
53	346.2157	346	9
52	344.2358	344	9
51	342.2703	342	9
50	340.3217	340	9
49	338.3854	338	9
48	336.4651	336	9
47	334.5553	335	9
46	332.6600	333	9
45	330.7777	331	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.9060	329	9
43	327.0467	327	9
42	325.1991	325	9
41	323.3636	323	9
40	321.5395	322	9
39	319.7301	320	9
38	317.9312	318	9
37	316.1469	316	9
36	314.3742	314	9
35	312.6140	313	9
34	310.8647	311	9
33	309.1261	309	9
32	307.3953	307	9
31	305.6716	306	9
30	303.9523	304	9
29	302.2348	302	9
28	300.5164	301	9
27	298.7909	299	10
26	297.0547	297	10
25	295.3029	295	10
24	293.5224	294	10
23	291.7135	292	10
22	289.8694	290	10
21	287.9804	288	10
20	286.0361	286	10
19	284.0272	284	11
18	281.9313	282	11
17	279.7375	280	11
16	277.4317	277	11
15	274.9959	275	12
14	272.4075	275	12
13	269.6383	275	12
12	266.6545	275	12
11	263.4162	275	13
10	259.8713	275	13
9	255.9602	275	14
8	251.5971	275	14
7	246.6594	275	15
6	240.9797	275	21
5	234.2954	275	24
4	226.1667	275	27
3	215.7601	275	28
2	201.1885	275	26
1	176.4029	275	19
0	147.2238	275	12

Table 6.O.4: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, October 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	494.0938	450	20
78	471.1714	450	22
77	457.4493	450	19
76	447.5204	448	17
75	439.6718	440	16
74	433.1381	433	14
73	427.5082	428	13
72	422.5380	423 (Advanced)	12
71	418.0724	418	12
70	413.9978	414	11
69	410.2457	410	11
68	406.7532	407	10
67	403.4811	403	10
66	400.3949	400	10
65	397.4658	397	10
64	394.6711	395	9
63	391.9963	392	9
62	389.4289	389	9
61	386.9510	387	9
60	384.5554	385	9
59	382.2307	382	9
58	379.9722	380 (Proficient)	8
57	377.7692	378	8
56	375.6193	376	8
55	373.5151	374	8
54	371.4521	371	8
53	369.4271	369	8
52	367.4316	367	8
51	365.4676	365	8
50	363.5323	364	8
49	361.6177	362	8
48	359.7239	360	8
47	357.8471	358	8
46	355.9857	356	8
45	354.1374	354	8
44	352.2989	352	8
43	350.4672	350 (Pass)	8
42	348.6437	349	8
41	346.8222	347	8
40	345.0025	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.1830	343	8
38	341.3606	341	8
37	339.5344	340	8
36	337.6995	338	8
35	335.8567	336	8
34	334.0020	334	8
33	332.1361	332	8
32	330.2524	330	8
31	328.3510	328	8
30	326.4277	326	8
29	324.4785	324	8
28	322.5044	323	8
27	320.4993	320	8
26	318.4581	318	8
25	316.3801	316	8
24	314.2574	314	8
23	312.0880	312	8
22	309.8656	310	9
21	307.5847	308	9
20	305.2381	305	9
19	302.8138	303	9
18	300.3093	300	9
17	297.7078	298	9
16	295.0006	295	9
15	292.1727	292	10
14	289.2042	289	10
13	286.0748	286	10
12	282.7565	283	11
11	279.2162	279	11
10	275.4086	275	11
9	271.2816	275	12
8	266.7550	275	12
7	261.7180	275	13
6	256.0194	275	14
5	249.4138	275	14
4	241.4914	275	20
3	231.4711	275	20
2	217.6263	275	18
1	194.6224	275	13
0	171.7079	275	10

Table 6.O.5: Operational, Large Print, Audio Books and Audio CD Conversions—ELA, November 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	502.1436	450	23
88	479.6632	450	27
87	466.1528	450	26
86	456.1224	450	23
85	447.8572	448	21
84	440.7007	441	17
83	434.3447	434	17
82	428.5795	429	17
81	423.2886	423	16
80	418.3954	418	15
79	413.8418	414	14
78	409.5771	410	14
77	405.5627	406 (Advanced)	13
76	401.7715	402	13
75	398.1694	398	12
74	394.7326	395	12
73	391.4418	391	12
72	388.2822	388	11
71	385.2368	385	11
70	382.3014	382 (Proficient)	11
69	379.4581	379	11
68	376.6995	377	10
67	374.0155	374	10
66	371.3988	371	10
65	368.8477	369	10
64	366.3522	366	10
63	363.9146	364	10
62	361.5258	362	9
61	359.1846	359	9
60	356.8839	357	9
59	354.6212	355	9
58	352.3942	352	9
57	350.1983	350 (Pass)	9
56	348.0352	348	9
55	345.9015	346	9
54	343.7950	344	9
53	341.7101	342	9
52	339.6460	340	9
51	337.6020	338	9
50	335.5740	336	9
49	333.5632	334	9
48	331.5680	332	9
47	329.5865	330	9
46	327.6163	328	9
45	325.6573	326	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	323.7074	324	9
43	321.7660	322	9
42	319.8347	320	9
41	317.9093	318	9
40	315.9934	316	9
39	314.0823	314	9
38	312.1795	312	9
37	310.2809	310	9
36	308.3893	308	9
35	306.5017	307	9
34	304.6206	305	9
33	302.7439	303	9
32	300.8735	301	9
31	299.0048	299	9
30	297.1373	297	9
29	295.2710	295	9
28	293.3973	293	10
27	291.5190	292	10
26	289.6337	290	10
25	287.7358	288	10
24	285.8216	286	10
23	283.8878	284	10
22	281.9207	282	10
21	279.9082	280	10
20	277.8530	278	11
19	275.7481	276	11
18	273.5862	275	11
17	271.3201	275	11
16	268.9581	275	12
15	266.4911	275	12
14	263.9025	275	12
13	261.1207	275	13
12	258.1365	275	13
11	254.9308	275	13
10	251.4593	275	14
9	247.6573	275	14
8	243.3426	275	15
7	238.4781	275	15
6	232.8943	275	21
5	226.3321	275	24
4	218.3565	275	27
3	208.1476	275	28
2	193.8478	275	26
1	170.8174	275	19
0	147.2238	275	8

Table 6.O.6: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, November 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	493.5740	450	20
78	470.5534	450	22
77	456.8771	450	19
76	447.0061	447	17
75	439.2210	439	16
74	432.7530	433	14
73	427.1904	427	13
72	422.2978	422 (Advanced)	12
71	417.9068	418	12
70	413.9042	414	11
69	410.2226	410	11
68	406.7987	407	10
67	403.5936	404	10
66	400.5730	401	10
65	397.7083	398	9
64	394.9770	395	9
63	392.3634	392	9
62	389.8568	390	9
61	387.4389	387	9
60	385.1013	385	9
59	382.8344	383	8
58	380.6314	381 (Proficient)	8
57	378.4846	378	8
56	376.3878	376	8
55	374.3371	374	8
54	372.3259	372	8
53	370.3513	370	8
52	368.4077	368	8
51	366.4913	366	8
50	364.6023	365	8
49	362.7367	363	8
48	360.8875	361	8
47	359.0565	359	8
46	357.2382	357	8
45	355.4326	355	8
44	353.6365	354	8
43	351.8468	352	8
42	350.0608	350 (Pass)	8
41	348.2809	348	8
40	346.4983	346	8
39	344.7155	345	8
38	342.9296	343	8
37	341.1376	341	8
36	339.3389	339	8
35	337.5281	338	8
34	335.7068	336	8
33	333.8700	334	8
32	332.0191	332	8
31	330.1467	330	8
30	328.2536	328	8
29	326.3341	326	8
28	324.3852	324	8
27	322.4082	322	8
26	320.3955	320	8
25	318.3421	318	8
24	316.2473	316	8
23	314.1023	314	8
22	311.9046	312	8
21	309.6485	310	9
20	307.3265	307	9
19	304.9310	305	9
18	302.4481	302	9
17	299.8739	300	9
16	297.1881	297	9
15	294.3823	294	10
14	291.4371	291	10
13	288.3290	288	10
12	285.0331	285	11
11	281.5136	282	11
10	277.7274	278	11
9	273.6172	275	12
8	269.1077	275	12
7	264.0859	275	13
6	258.3945	275	14
5	251.7886	275	14
4	243.8517	275	20
3	233.8053	275	20
2	219.8976	275	18
1	196.5427	275	13
0	171.7079	275	10

Table 6.O.7: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, December 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	498.8820	450	23
88	475.5735	450	27
87	461.7370	450	26
86	451.6031	450	23
85	443.4083	443	21
84	436.4127	436	17
83	430.2563	430	17
82	424.7144	425	17
81	419.6568	420	15
80	414.9963	415	15
79	410.6655	411	14
78	406.6245	407	13
77	402.8301	403 (Advanced)	13
76	399.2405	399	12
75	395.8289	396	12
74	392.5731	393	12
73	389.4557	389	11
72	386.4600	386	11
71	383.5708	384	11
70	380.7752	381 (Proficient)	11
69	378.0683	378	10
68	375.4363	375	10
67	372.8746	373	10
66	370.3794	370	10
65	367.9451	368	10
64	365.5593	366	10
63	363.2220	363	9
62	360.9272	361	9
61	358.6738	359	9
60	356.4556	356	9
59	354.2709	354	9
58	352.1168	352	9
57	349.9880	350 (Pass)	9
56	347.8906	348	9
55	345.8144	346	9
54	343.7603	344	9
53	341.7233	342	9
52	339.7023	340	9
51	337.6961	338	9
50	335.7013	336	8
49	333.7179	334	8
48	331.7450	332	8
47	329.7789	330	8
46	327.8188	328	8
45	325.8635	326	8
44	323.9103	324	8
43	321.9584	322	8
42	320.0079	320	8
41	318.0556	318	8
40	316.1030	316	9
39	314.1460	314	9
38	312.1871	312	9
37	310.2214	310	9
36	308.2514	308	9
35	306.2743	306	9
34	304.2918	304	9
33	302.3033	302	9
32	300.3089	300	9
31	298.3056	298	9
30	296.2954	296	9
29	294.2772	294	9
28	292.2499	292	9
27	290.2149	290	10
26	288.1701	288	10
25	286.1120	286	10
24	284.0410	284	10
23	281.9445	282	10
22	279.8211	280	10
21	277.6699	278	11
20	275.4874	275	11
19	273.2699	275	11
18	270.9940	275	11
17	268.6348	275	11
16	266.2059	275	12
15	263.6971	275	12
14	261.0521	275	12
13	258.2460	275	13
12	255.2869	275	13
11	252.1491	275	14
10	248.6745	275	14
9	244.8727	275	15
8	240.6853	275	16
7	236.0208	275	16
6	230.5285	275	21
5	224.0106	275	24
4	216.0539	275	27
3	205.8044	275	28
2	191.2652	275	26
1	169.5217	275	19
0	147.2238	275	8

Table 6.O.8: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, December 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	495.3897	450	20
78	473.7720	450	22
77	459.7902	450	19
76	449.7902	450	17
75	441.9238	442	16
74	435.3997	435	15
73	429.7907	430	13
72	424.8496	425 (Advanced)	12
71	420.4167	420	12
70	416.3803	416	11
69	412.6669	413	11
68	409.2175	409	10
67	405.9865	406	10
66	402.9435	403	10
65	400.0655	400	10
64	397.3274	397	9
63	394.7044	395	9
62	392.1849	392	9
61	389.7594	390	9
60	387.4127	387	9
59	385.1379	385	9
58	382.9267	383	8
57	380.7731	381 (Proficient)	8
56	378.6709	379	8
55	376.6137	377	8
54	374.5992	375	8
53	372.6207	373	8
52	370.6756	371	8
51	368.7601	369	8
50	366.8682	367	8
49	365.0017	365	8
48	363.1592	363	8
47	361.3303	361	8
46	359.5181	360	8
45	357.7173	358	8
44	355.9274	356	8
43	354.1460	354	8
42	352.3702	352	8
41	350.5970	351 (Pass)	8
40	348.8276	349	8
39	347.0568	347	8
38	345.2835	345	8
37	343.5065	344	8
36	341.7229	342	8
35	339.9312	340	8
34	338.1276	338	8
33	336.3109	336	8
32	334.4790	334	8
31	332.6300	333	8
30	330.7605	331	8
29	328.8663	329	8
28	326.9461	327	8
27	324.9942	325	8
26	323.0101	323	8
25	320.9896	321	8
24	318.9253	319	8
23	316.8149	317	8
22	314.6517	315	8
21	312.4300	312	9
20	310.1446	310	9
19	307.7869	308	9
18	305.3481	305	9
17	302.8136	303	9
16	300.1781	300	9
15	297.4204	297	10
14	294.5227	295	10
13	291.4622	291	10
12	288.2130	288	10
11	284.7437	285	11
10	281.0099	281	11
9	276.9552	277	12
8	272.5066	275	12
7	267.5609	275	13
6	261.9644	275	14
5	255.4304	275	14
4	247.4940	275	20
3	237.4025	275	20
2	223.3506	275	18
1	199.5553	275	13
0	171.7079	275	10

Table 6.O.9: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, February 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	507.7123	450	23
88	489.4439	450	27
87	473.9295	450	26
86	462.6499	450	23
85	453.4113	450	21
84	445.4236	445	17
83	438.3321	438	17
82	431.9630	432	17
81	426.1883	426	16
80	420.9183	421	15
79	416.0808	416	15
78	411.6111	412	14
77	407.4503	407	13
76	403.5668	404 (Advanced)	13
75	399.9134	400	12
74	396.4668	396	12
73	393.1915	393	12
72	390.0648	390	11
71	387.0693	387	11
70	384.1870	384	11
69	381.4046	381 (Proficient)	10
68	378.7133	379	10
67	376.1024	376	10
66	373.5644	374	10
65	371.0903	371	10
64	368.6786	369	10
63	366.3172	366	9
62	364.0076	364	9
61	361.7410	362	9
60	359.5171	360	9
59	357.3284	357	9
58	355.1719	355	9
57	353.0458	353	9
56	350.9448	351 (Pass)	9
55	348.8707	349	9
54	346.8188	347	9
53	344.7883	345	9
52	342.7744	343	9
51	340.7779	341	9
50	338.7923	339	9
49	336.8243	337	8
48	334.8631	335	8
47	332.9137	333	8
46	330.9739	331	8
45	329.0395	329	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	327.1115	327	8
43	325.1885	325	9
42	323.2700	323	9
41	321.3543	321	9
40	319.4451	319	9
39	317.5355	318	9
38	315.6325	316	9
37	313.7290	314	9
36	311.8301	312	9
35	309.9330	310	9
34	308.0393	308	9
33	306.1474	306	9
32	304.2588	304	9
31	302.3728	302	9
30	300.4895	300	9
29	298.6050	299	9
28	296.7197	297	10
27	294.8324	295	10
26	292.9385	293	10
25	291.0372	291	10
24	289.1178	289	10
23	287.1774	287	10
22	285.2126	285	10
21	283.2164	283	10
20	281.1828	281	11
19	279.1062	279	11
18	276.9499	277	11
17	274.7185	275	11
16	272.4011	275	12
15	269.9808	275	12
14	267.4374	275	12
13	264.7079	275	13
12	261.7823	275	13
11	258.6311	275	13
10	255.2056	275	14
9	251.4414	275	14
8	247.2511	275	15
7	242.4951	275	15
6	236.9774	275	21
5	230.4689	275	24
4	222.5336	275	27
3	212.3443	275	28
2	198.0589	275	26
1	173.7830	275	19
0	147.2238	275	8

Table 6.O.10: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, February 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	493.8873	450	20
78	470.8512	450	22
77	457.1232	450	19
76	447.1903	447	17
75	439.3403	439	16
74	432.8066	433	14
73	427.1789	427	13
72	422.2143	422 (Advanced)	12
71	417.7564	418	12
70	413.6900	414	11
69	409.9473	410	11
68	406.4647	406	10
67	403.2038	403	10
66	400.1295	400	10
65	397.2128	397	10
64	394.4314	394	9
63	391.7710	392	9
62	389.2185	389	9
61	386.7559	387	9
60	384.3766	384	9
59	382.0685	382	9
58	379.8275	380 (Proficient)	8
57	377.6422	378	8
56	375.5109	376	8
55	373.4254	373	8
54	371.3818	371	8
53	369.3767	369	8
52	367.4012	367	8
51	365.4580	365	8
50	363.5438	364	8
49	361.6507	362	8
48	359.7788	360	8
47	357.9242	358	8
46	356.0854	356	8
45	354.2600	354	8
44	352.4448	352	8
43	350.6369	351 (Pass)	8
42	348.8370	349	8
41	347.0402	347	8
40	345.2450	345	8
39	343.4504	343	8
38	341.6534	342	8
37	339.8525	340	8
36	338.0441	338	8
35	336.2270	336	8
34	334.3992	334	8
33	332.5592	333	8
32	330.7033	331	8
31	328.8286	329	8
30	326.9331	327	8
29	325.0119	325	8
28	323.0643	323	8
27	321.0876	321	8
26	319.0746	319	8
25	317.0235	317	8
24	314.9292	315	8
23	312.7863	313	8
22	310.5905	311	8
21	308.3355	308	9
20	306.0139	306	9
19	303.6151	304	9
18	301.1317	301	9
17	298.5524	299	9
16	295.8631	296	9
15	293.0506	293	10
14	290.0947	290	10
13	286.9726	287	10
12	283.6568	284	11
11	280.1117	280	11
10	276.2915	276	11
9	272.1401	275	12
8	267.5770	275	12
7	262.4876	275	13
6	256.7155	275	14
5	250.0118	275	14
4	241.9610	275	20
3	231.7929	275	20
2	217.7875	275	18
1	194.5036	275	13
0	171.7079	275	10

Table 6.O.11: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, March 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	505.9316	450	23
88	485.7696	450	27
87	470.6862	450	26
86	459.6321	450	23
85	450.5874	450	21
84	442.7976	443	17
83	435.9017	436	17
82	429.7160	430	17
81	424.1056	424	16
80	418.9886	419	15
79	414.3027	414	14
78	409.9620	410	14
77	405.9157	406 (Advanced)	13
76	402.1329	402	13
75	398.5695	399	12
74	395.1958	395	12
73	391.9873	392	11
72	388.9245	389	11
71	385.9880	386	11
70	383.1621	383	11
69	380.4316	380 (Proficient)	10
68	377.7923	378	10
67	375.2272	375	10
66	372.7351	373	10
65	370.3051	370	10
64	367.9337	368	10
63	365.6092	366	9
62	363.3314	363	9
61	361.0935	361	9
60	358.8949	359	9
59	356.7292	357	9
58	354.5942	355	9
57	352.4876	352	9
56	350.4042	350 (Pass)	9
55	348.3475	348	9
54	346.3107	346	9
53	344.2957	344	9
52	342.2965	342	9
51	340.3142	340	9
50	338.3436	338	9
49	336.3876	336	9
48	334.4405	334	8
47	332.5058	333	8
46	330.5802	331	8
45	328.6617	329	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	326.7513	327	9
43	324.8471	325	9
42	322.9494	323	9
41	321.0562	321	9
40	319.1720	319	9
39	317.2891	317	9
38	315.4146	315	9
37	313.5426	314	9
36	311.6768	312	9
35	309.8150	310	9
34	307.9579	308	9
33	306.1040	306	9
32	304.2540	304	9
31	302.4067	302	9
30	300.5619	301	9
29	298.7145	299	9
28	296.8646	297	10
27	295.0107	295	10
26	293.1471	293	10
25	291.2658	291	10
24	289.3678	289	10
23	287.4452	287	10
22	285.4931	285	10
21	283.5044	284	10
20	281.4724	281	11
19	279.3821	279	11
18	277.2121	277	11
17	274.9637	275	11
16	272.6229	275	11
15	270.1727	275	12
14	267.5881	275	12
13	264.8111	275	13
12	261.8385	275	13
11	258.6344	275	13
10	255.1517	275	14
9	251.3268	275	14
8	247.0738	275	15
7	242.2797	275	15
6	236.7175	275	21
5	230.1677	275	24
4	222.2013	275	27
3	211.9947	275	28
2	197.7124	275	26
1	173.4794	275	19
0	147.2238	275	8

Table 6.O.12: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, March 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	494.1159	450	20
78	471.2111	450	22
77	457.4816	450	19
76	447.5490	448	17
75	439.6983	440	16
74	433.1632	433	14
73	427.5325	428	13
72	422.5618	423 (Advanced)	12
71	418.0960	418	12
70	414.0215	414	11
69	410.2696	410	11
68	406.7775	407	10
67	403.5059	404	10
66	400.4202	400	10
65	397.4917	397	10
64	394.6977	395	9
63	392.0235	392	9
62	389.4568	389	9
61	386.9795	387	9
60	384.5844	385	9
59	382.2604	382	9
58	380.0024	380 (Proficient)	8
57	377.8000	378	8
56	375.6506	376	8
55	373.5469	374	8
54	371.4844	371	8
53	369.4600	369	8
52	367.4651	367	8
51	365.5017	366	8
50	363.5670	364	8
49	361.6532	362	8
48	359.7603	360	8
47	357.8846	358	8
46	356.0243	356	8
45	354.1773	354	8
44	352.3405	352	8
43	350.5106	351 (Pass)	8
42	348.6892	349	8
41	346.8702	347	8
40	345.0533	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.2370	343	8
38	341.4183	341	8
37	339.5963	340	8
36	337.7663	338	8
35	335.9288	336	8
34	334.0803	334	8
33	332.2210	332	8
32	330.3451	330	8
31	328.4521	328	8
30	326.5383	327	8
29	324.5996	325	8
28	322.6370	323	8
27	320.6448	321	8
26	318.6178	319	8
25	316.5550	317	8
24	314.4495	314	8
23	312.2984	312	8
22	310.0962	310	8
21	307.8373	308	9
20	305.5147	306	9
19	303.1169	303	9
18	300.6400	301	9
17	298.0696	298	9
16	295.3952	295	9
15	292.6027	293	10
14	289.6728	290	10
13	286.5844	287	10
12	283.3104	283	11
11	279.8173	280	11
10	276.0603	276	11
9	271.9861	275	12
8	267.5158	275	12
7	262.5380	275	13
6	256.8991	275	14
5	250.3536	275	14
4	242.4882	275	20
3	232.5310	275	20
2	218.7455	275	18
1	195.6011	275	13
0	171.7079	275	10

Table 6.O.13: Operational, Large Print, Audio Books, and Audio CD Conversions—ELA, May 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	500.5602	450	23
88	478.0638	450	27
87	464.5843	450	26
86	454.5483	450	23
85	446.2799	446	21
84	439.1029	439	17
83	432.7014	433	17
82	426.8902	427	17
81	421.564	422	15
80	416.6537	417	15
79	412.1040	412	14
78	407.8664	408	14
77	403.9059	404 (Advanced)	13
76	400.1892	400	13
75	396.6814	397	12
74	393.3661	393	12
73	390.2108	390	11
72	387.1960	387	11
71	384.3023	384	11
70	381.5156	382 (Proficient)	11
69	378.8254	379	10
68	376.2205	376	10
67	373.6928	374	10
66	371.2321	371	10
65	368.8359	369	10
64	366.4925	366	9
63	364.2011	364	9
62	361.9520	362	9
61	359.7492	360	9
60	357.5824	358	9
59	355.4479	355	9
58	353.3441	353	9
57	351.2659	351 (Pass)	9
56	349.2137	349	9
55	347.1846	347	9
54	345.1757	345	9
53	343.1844	343	9
52	341.2091	341	9
51	339.2457	339	8
50	337.2967	337	8
49	335.3550	335	8
48	333.4243	333	8
47	331.5021	332	8
46	329.5854	330	8
45	327.6728	328	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	325.7643	326	8
43	323.8574	324	8
42	321.9517	322	8
41	320.0479	320	8
40	318.1432	318	9
39	316.2389	316	9
38	314.3325	314	9
37	312.4258	312	9
36	310.5166	311	9
35	308.6068	309	9
34	306.6921	307	9
33	304.7765	305	9
32	302.8576	303	9
31	300.9387	301	9
30	299.0146	299	9
29	297.0864	297	9
28	295.1553	295	9
27	293.2146	293	10
26	291.264	291	10
25	289.3032	289	10
24	287.3258	287	10
23	285.3291	285	10
22	283.309	283	10
21	281.2601	281	10
20	279.1803	279	11
19	277.0317	277	11
18	274.8282	275	11
17	272.5596	275	11
16	270.2139	275	12
15	267.7607	275	12
14	265.1591	275	12
13	262.4122	275	13
12	259.4931	275	13
11	256.3704	275	14
10	252.9382	275	14
9	249.1515	275	15
8	244.9446	275	15
7	240.2013	275	16
6	234.7533	275	21
5	228.3441	275	24
4	220.552	275	27
3	210.5405	275	28
2	196.3871	275	26
1	172.3720	275	19
0	147.2238	275	8

Table 6.O.14: Operational, Large Print, Audio Books, and Audio CD Conversions—Mathematics, May 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	491.6093	450	20
78	468.2823	450	22
77	454.4847	450	19
76	444.5592	445	17
75	436.7487	437	16
74	430.2705	430	14
73	424.7067	425 (Advanced)	13
72	419.8103	420	12
71	415.4169	415	12
70	411.4258	411	11
69	407.7534	408	11
68	404.3439	404	10
67	401.1549	401	10
66	398.1505	398	10
65	395.3025	395	9
64	392.5917	393	9
63	390.0011	390	9
62	387.5142	388	9
61	385.1188	385	9
60	382.8041	383	9
59	380.5618	381 (Proficient)	8
58	378.3833	378	8
57	376.2618	376	8
56	374.1923	374	8
55	372.1681	372	8
54	370.1856	370	8
53	368.2392	368	8
52	366.3230	366	8
51	364.4396	364	8
50	362.5830	363	8
49	360.7469	361	8
48	358.9323	359	8
47	357.1341	357	8
46	355.3517	355	8
45	353.5822	354	8
44	351.8222	352	8
43	350.0695	350 (Pass)	8
42	348.3258	348	8
41	346.5829	347	7
40	344.8426	345	7
39	343.1025	343	7
38	341.3596	341	8
37	339.6132	340	8
36	337.8585	338	8
35	336.0960	336	8
34	334.3224	334	8
33	332.5371	333	8
32	330.7360	331	8
31	328.9165	329	8
30	327.0769	327	8
29	325.2123	325	8
28	323.3213	323	8
27	321.4028	321	8
26	319.4491	319	8
25	317.4577	317	8
24	315.4256	315	8
23	313.3454	313	8
22	311.2143	311	8
21	309.0263	309	8
20	306.7739	307	9
19	304.4487	304	9
18	302.0389	302	9
17	299.5402	300	9
16	296.9333	297	9
15	294.2086	294	9
14	291.3467	291	10
13	288.3242	288	10
12	285.1162	285	10
11	281.6870	282	11
10	277.9934	278	11
9	273.9776	275	12
8	269.5646	275	12
7	264.6410	275	13
6	259.0477	275	14
5	252.5389	275	14
4	244.695	275	20
3	234.7300	275	20
2	220.8737	275	18
1	197.4720	275	13
0	171.7079	275	10

Table 6.O.15: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—July,2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	502.9092	450	23
88	481.5591	450	27
87	468.1929	450	26
86	458.2292	450	23
85	450.0319	450	21
84	442.9393	443	17
83	436.6174	437	17
82	430.8931	431	17
81	425.6481	426	16
80	420.8059	421	15
79	416.3097	416	14
78	412.1103	412	14
77	408.1632	408	13
76	404.4443	404 (Advanced)	13
75	400.9192	401	12
74	397.5615	398	12
73	394.3505	394	12
72	391.2694	391	11
71	388.3049	388	11
70	385.442	385	11
69	382.6704	383	10
68	379.9789	380 (Proficient)	10
67	377.3696	377	10
66	374.8266	375	10
65	372.3440	372	10
64	369.9176	370	10
63	367.5443	368	9
62	365.2165	365	9
61	362.9322	363	9
60	360.6879	361	9
59	358.4815	358	9
58	356.3076	356	9
57	354.1651	354	9
56	352.0509	352	9
55	349.9607	350 (Pass)	9
54	347.8999	348	9
53	345.859	346	9
52	343.8398	344	9
51	341.8371	342	9
50	339.8507	340	9
49	337.879	338	9
48	335.9204	336	8
47	333.9738	334	8
46	332.0400	332	8
45	330.1154	330	9

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
44	328.1994	328	9
43	326.2927	326	9
42	324.3927	324	9
41	322.4998	322	9
40	320.6127	321	9
39	318.7332	319	9
38	316.8571	317	9
37	314.9873	315	9
36	313.1208	313	9
35	311.2585	311	9
34	309.3990	309	9
33	307.5410	308	9
32	305.6837	306	9
31	303.8256	304	9
30	301.9662	302	9
29	300.101	300	9
28	298.2275	298	10
27	296.3436	296	10
26	294.4455	294	10
25	292.5282	293	10
24	290.5892	291	10
23	288.6233	289	10
22	286.6180	287	10
21	284.5623	285	10
20	282.4525	282	11
19	280.2824	280	11
18	278.0397	278	11
17	275.7116	276	11
16	273.2852	275	12
15	270.7409	275	12
14	268.0321	275	12
13	265.1557	275	13
12	262.0796	275	13
11	258.7678	275	14
10	255.1744	275	14
9	251.2367	275	15
8	246.8695	275	15
7	241.9606	275	16
6	236.3439	275	21
5	229.7086	275	24
4	221.6674	275	27
3	211.4008	275	28
2	197.0857	275	26
1	172.8883	275	19
0	147.2238	275	8

Table 6.O.16: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics—July,2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	491.6116	450	20
78	468.2854	450	22
77	454.4884	450	19
76	444.5631	445	17
75	436.7526	437	16
74	430.2741	430	14
73	424.7098	425 (Advanced)	13
72	419.8127	420	12
71	415.4184	415	12
70	411.4261	411	11
69	407.7523	408	11
68	404.3412	404	10
67	401.1504	401	10
66	398.1439	398	10
65	395.2936	395	9
64	392.5803	393	9
63	389.9870	390	9
62	387.4972	387	9
61	385.0987	385	9
60	382.7805	383	9
59	380.5348	381 (Proficient)	8
58	378.3524	378	8
57	376.2269	376	8
56	374.1532	374	8
55	372.1246	372	8
54	370.1375	370	8
53	368.1860	368	8
52	366.2648	366	8
51	364.3763	364	8
50	362.5141	363	8
49	360.6723	361	8
48	358.8517	359	8
47	357.0474	357	8
46	355.2587	355	8
45	353.4826	353	8
44	351.7156	352	8
43	349.9559	350 (Pass)	8
42	348.2051	348	8
41	346.4546	346	8
40	344.7070	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.9592	343	8
38	341.2083	341	8
37	339.4538	339	8
36	337.6906	338	8
35	335.9197	336	8
34	334.1372	334	8
33	332.3432	332	8
32	330.5326	331	8
31	328.7039	329	8
30	326.8544	327	8
29	324.9797	325	8
28	323.0791	323	8
27	321.1499	321	8
26	319.1854	319	8
25	317.1835	317	8
24	315.1398	315	8
23	313.0485	313	8
22	310.9058	311	8
21	308.7060	309	8
20	306.4415	306	9
19	304.1032	304	9
18	301.6814	302	9
17	299.1691	299	9
16	296.5493	297	9
15	293.8111	294	10
14	290.9352	291	10
13	287.8987	288	10
12	284.6762	285	10
11	281.2326	281	11
10	277.5241	278	11
9	273.4941	275	12
8	269.0665	275	12
7	264.1285	275	13
6	258.5223	275	14
5	252.0020	275	14
4	244.1499	275	20
3	234.1832	275	20
2	220.3405	275	18
1	196.9992	275	13
0	171.7079	275	10

Table 6.O.17: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—October, November and December, 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	501.9302	450	23
88	478.8967	450	27
87	465.0876	450	26
86	454.9495	450	23
85	446.7440	447	21
84	439.7542	440	17
83	433.6027	434	17
82	428.0848	428	17
81	423.0571	423	15
80	418.4305	418	14
79	414.1394	414	14
78	410.1304	410	13
77	406.3613	406	13
76	402.8086	403 (Advanced)	12
75	399.4370	399	12
74	396.2237	396	12
73	393.1490	393	11
72	390.1977	390	11
71	387.3564	387	11
70	384.6101	385	11
69	381.9491	382 (Proficient)	10
68	379.3638	379	10
67	376.8496	377	10
66	374.3972	374	10
65	372.0004	372	10
64	369.6558	370	10
63	367.3586	367	9
62	365.1029	365	9
61	362.8858	363	9
60	360.7047	361	9
59	358.5573	359	9
58	356.4386	356	9
57	354.3477	354	9
56	352.2820	352	9
55	350.2369	350 (Pass)	9
54	348.2173	348	9
53	346.2157	346	9
52	344.2358	344	9
51	342.2703	342	9
50	340.3217	340	9
49	338.3854	338	9
48	336.4651	336	9
47	334.5553	335	9
46	332.6600	333	9
45	330.7777	331	9
44	328.9060	329	9
43	327.0467	327	9
42	325.1991	325	9
41	323.3636	323	9
40	321.5395	322	9
39	319.7301	320	9
38	317.9312	318	9
37	316.1469	316	9
36	314.3742	314	9
35	312.6140	313	9
34	310.8647	311	9
33	309.1261	309	9
32	307.3953	307	9
31	305.6716	306	9
30	303.9523	304	9
29	302.2348	302	9
28	300.5164	301	9
27	298.7909	299	10
26	297.0547	297	10
25	295.3029	295	10
24	293.5224	294	10
23	291.7135	292	10
22	289.8694	290	10
21	287.9804	288	10
20	286.0361	286	10
19	284.0272	284	11
18	281.9313	282	11
17	279.7375	280	11
16	277.4317	277	11
15	274.9959	275	12
14	272.4075	275	12
13	269.6383	275	12
12	266.6545	275	12
11	263.4162	275	13
10	259.8713	275	13
9	255.9602	275	14
8	251.5971	275	14
7	246.6594	275	15
6	240.9797	275	21
5	234.2954	275	24
4	226.1667	275	27
3	215.7601	275	28
2	201.1885	275	26
1	176.4029	275	19
0	147.2238	275	12

Table 6.O.18: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— October, November, and December, 2008

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	494.0938	450	20
78	471.1714	450	22
77	457.4493	450	19
76	447.5204	448	17
75	439.6718	440	16
74	433.1381	433	14
73	427.5082	428	13
72	422.5380	423 (Advanced)	12
71	418.0724	418	12
70	413.9978	414	11
69	410.2457	410	11
68	406.7532	407	10
67	403.4811	403	10
66	400.3949	400	10
65	397.4658	397	10
64	394.6711	395	9
63	391.9963	392	9
62	389.4289	389	9
61	386.9510	387	9
60	384.5554	385	9
59	382.2307	382	9
58	379.9722	380 (Proficient)	8
57	377.7692	378	8
56	375.6193	376	8
55	373.5151	374	8
54	371.4521	371	8
53	369.4271	369	8
52	367.4316	367	8
51	365.4676	365	8
50	363.5323	364	8
49	361.6177	362	8
48	359.7239	360	8
47	357.8471	358	8
46	355.9857	356	8
45	354.1374	354	8
44	352.2989	352	8
43	350.4672	350 (Pass)	8
42	348.6437	349	8
41	346.8222	347	8
40	345.0025	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.1830	343	8
38	341.3606	341	8
37	339.5344	340	8
36	337.6995	338	8
35	335.8567	336	8
34	334.0020	334	8
33	332.1361	332	8
32	330.2524	330	8
31	328.3510	328	8
30	326.4277	326	8
29	324.4785	324	8
28	322.5044	323	8
27	320.4993	320	8
26	318.4581	318	8
25	316.3801	316	8
24	314.2574	314	8
23	312.0880	312	8
22	309.8656	310	9
21	307.5847	308	9
20	305.2381	305	9
19	302.8138	303	9
18	300.3093	300	9
17	297.7078	298	9
16	295.0006	295	9
15	292.1727	292	10
14	289.2042	289	10
13	286.0748	286	10
12	282.7565	283	11
11	279.2162	279	11
10	275.4086	275	11
9	271.2816	275	12
8	266.7550	275	12
7	261.7180	275	13
6	256.0194	275	14
5	249.4138	275	14
4	241.4914	275	20
3	231.4711	275	20
2	217.6263	275	18
1	194.6224	275	13
0	171.7079	275	10

Table 6.O.19: Braille, LP-Braille, and LP-Braille CD Conversions—ELA—February, March, and May 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450	15
89	507.7123	450	23
88	489.4439	450	27
87	473.9295	450	26
86	462.6499	450	23
85	453.4113	450	21
84	445.4236	445	17
83	438.3321	438	17
82	431.9630	432	17
81	426.1883	426	16
80	420.9183	421	15
79	416.0808	416	15
78	411.6111	412	14
77	407.4503	407	13
76	403.5668	404 (Advanced)	13
75	399.9134	400	12
74	396.4668	396	12
73	393.1915	393	12
72	390.0648	390	11
71	387.0693	387	11
70	384.1870	384	11
69	381.4046	381 (Proficient)	10
68	378.7133	379	10
67	376.1024	376	10
66	373.5644	374	10
65	371.0903	371	10
64	368.6786	369	10
63	366.3172	366	9
62	364.0076	364	9
61	361.7410	362	9
60	359.5171	360	9
59	357.3284	357	9
58	355.1719	355	9
57	353.0458	353	9
56	350.9448	351 (Pass)	9
55	348.8707	349	9
54	346.8188	347	9
53	344.7883	345	9
52	342.7744	343	9
51	340.7779	341	9
50	338.7923	339	9
49	336.8243	337	8
48	334.8631	335	8
47	332.9137	333	8
46	330.9739	331	8
45	329.0395	329	8
44	327.1115	327	8
43	325.1885	325	9
42	323.2700	323	9
41	321.3543	321	9
40	319.4451	319	9
39	317.5355	318	9
38	315.6325	316	9
37	313.7290	314	9
36	311.8301	312	9
35	309.9330	310	9
34	308.0393	308	9
33	306.1474	306	9
32	304.2588	304	9
31	302.3728	302	9
30	300.4895	300	9
29	298.6050	299	9
28	296.7197	297	10
27	294.8324	295	10
26	292.9385	293	10
25	291.0372	291	10
24	289.1178	289	10
23	287.1774	287	10
22	285.2126	285	10
21	283.2164	283	10
20	281.1828	281	11
19	279.1062	279	11
18	276.9499	277	11
17	274.7185	275	11
16	272.4011	275	12
15	269.9808	275	12
14	267.4374	275	12
13	264.7079	275	13
12	261.7823	275	13
11	258.6311	275	13
10	255.2056	275	14
9	251.4414	275	14
8	247.2511	275	15
7	242.4951	275	15
6	236.9774	275	21
5	230.4689	275	24
4	222.5336	275	27
3	212.3443	275	28
2	198.0589	275	26
1	173.7830	275	19
0	147.2238	275	8

Table 6.O.20: Braille, LP-Braille, and LP-Braille CD Conversions—Mathematics— February, March, and May 2009

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	14
79	493.8873	450	20
78	470.8512	450	22
77	457.1232	450	19
76	447.1903	447	17
75	439.3403	439	16
74	432.8066	433	14
73	427.1789	427	13
72	422.2143	422 (Advanced)	12
71	417.7564	418	12
70	413.6900	414	11
69	409.9473	410	11
68	406.4647	406	10
67	403.2038	403	10
66	400.1295	400	10
65	397.2128	397	10
64	394.4314	394	9
63	391.7710	392	9
62	389.2185	389	9
61	386.7559	387	9
60	384.3766	384	9
59	382.0685	382	9
58	379.8275	380 (Proficient)	8
57	377.6422	378	8
56	375.5109	376	8
55	373.4254	373	8
54	371.3818	371	8
53	369.3767	369	8
52	367.4012	367	8
51	365.4580	365	8
50	363.5438	364	8
49	361.6507	362	8
48	359.7788	360	8
47	357.9242	358	8
46	356.0854	356	8
45	354.2600	354	8
44	352.4448	352	8
43	350.6369	351 (Pass)	8
42	348.8370	349	8
41	347.0402	347	8
40	345.2450	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.4504	343	8
38	341.6534	342	8
37	339.8525	340	8
36	338.0441	338	8
35	336.2270	336	8
34	334.3992	334	8
33	332.5592	333	8
32	330.7033	331	8
31	328.8286	329	8
30	326.9331	327	8
29	325.0119	325	8
28	323.0643	323	8
27	321.0876	321	8
26	319.0746	319	8
25	317.0235	317	8
24	314.9292	315	8
23	312.7863	313	8
22	310.5905	311	8
21	308.3355	308	9
20	306.0139	306	9
19	303.6151	304	9
18	301.1317	301	9
17	298.5524	299	9
16	295.8631	296	9
15	293.0506	293	10
14	290.0947	290	10
13	286.9726	287	10
12	283.6568	284	11
11	280.1117	280	11
10	276.2915	276	11
9	272.1401	275	12
8	267.5770	275	12
7	262.4876	275	13
6	256.7155	275	14
5	250.0118	275	14
4	241.9610	275	20
3	231.7929	275	20
2	217.7875	275	18
1	194.5036	275	13
0	171.7079	275	10

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 error 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. Since then, Assembly Bill (AB) 1609 passed in October 2001. The bill prohibited ninth-graders from taking the CAHSEE in order to have a census testing of all tenth-graders. In July 2003, the SBE made the passing of both CAHSEE ELA and mathematics examinations a diploma requirement for the Class of 2006 and adopted revised test blueprints for the CAHSEE. The SBE directed the CDE to conduct a standard setting on the first census administration of the CAHSEE, which was in spring 2003.

In September 2003, a standard-setting workshop was carried out using CAHSEE test items that were constructed to meet the new blueprints and difficulty specifications. ETS used the Bookmark Method, a procedure for setting passing scores that 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 is typically a three-round standard-setting process in which panelists are instructed to set one passing score at the “just sufficient” level on the test they reviewed (ELA or mathematics). During the standard setting workshop, the panelists were asked to review the item booklet, in which items were ordered from easiest to hardest based on item difficulty, and then to find the point that defined the knowledge and skills needed to just pass the CAHSEE.

Two standard-setting panels for each subject were convened for the workshop. One panel in each subject was comprised primarily of California educators (e.g., English or mathematics teachers), while the other panel was comprised primarily 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 went through a rigorous training process, 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 feel students have demonstrated sufficient knowledge and skills in that subject area. Panelists typically place the first bookmark independently and then receive information on how their bookmark placement compares with those of their peers. There is then a small-group discussion followed by a second bookmark placement. Finally, there is 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.

First, the February 2004 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.[7]

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

Second, 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

[7] 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 score corresponding to the theta level closest to the values of .5356 for ELA and -.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 2008–09 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 2008 administration, 29 percent of examinees who took the ELA examination received a raw score of 54 or higher and thereby passed the ELA. Of those who passed, 5 percent of examinees received an ELA score of 68 or higher and passed the proficient cut point, while 2 percent scored 76 or above and passed the advanced cut point. Among examinees who took the mathematics exam, 28 percent passed by attaining a raw score of 43 or higher. Of those who passed, 4 percent passed the proficient cut point with a raw score of 59 or higher, while 1 percent passed the advanced cut 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 2008	Pass	54	29	43	28
	Proficient And Above	68	5	59	4
	Advanced And Above	76	2	73	1
October 2008	Pass	55	36	43	32
	Proficient And Above	69	8	58	7
	Advanced And Above	76	4	72	1
November 2008	Pass	57	39	42	36
	Proficient And Above	70	11	58	9
	Advanced And Above	77	5	72	2
December 2008	Pass	57	29	41	34
	Proficient And Above	70	7	57	5
	Advanced And Above	77	3	72	1
February 2009	Pass	56	65	43	65
	Proficient And Above	69	41	58	40
	Advanced And Above	76	25	72	15
March 2009	Pass	56	72	43	72
	Proficient And Above	69	44	58	46
	Advanced And Above	77	22	72	17
May 2009	Pass	57	32	43	30
	Proficient And Above	70	11	59	9
	Advanced And Above	77	5	73	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 ETS's mission 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 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—Provide 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—Describe 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—Identify the reporting strands for each test and the number of items in each strand.

- Matching Criteria for MC and Writing Answer Documents—Describe the method used to match students' writing and MC results.

The scoring specifications are reviewed and revised by the CDE, ETS, and Pearson each year. After a version that all 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, MC item data, and, if applicable, data from the Pre-ID process. Pearson sends the data file images to ETS for use in the ETS *Online Scoring Network*™ (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 Score Key Management (SKM) system. Prior to the scoring, information on all test items and 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 ETS 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 examinees' scores on the MC test items. In ELA, the total raw score equals the weighted sum of examinees' scores on both the MC items and the writing task. The weighting scheme for raw score 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 that have been tested in different administrations in a given content area can be compared. However, the raw scores of these examinees 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

The 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 10th 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,[8] it is considered discrepant, and a scoring leader or content specialist provides a third score, which

[8] If an essay received an NS from a reader and a score of 1, 2, 3, or 4 from another reader, scores are considered discrepant and the essay will receive a score from the third reader.

becomes the score of record. Otherwise, the writing score is obtained by averaging two scores.

Score Verification Procedures

ETS takes various measures to ascertain that the scoring keys are applied to the student responses as expected and the student scores are computed accurately.

Scoring Key Verification Process

Scoring keys, provided in the form planners, are produced by ETS 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. Various checks are performed before keys are finalized, as 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 at ETS.

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[9] in the form planners are 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 ETS 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, a follow-up process is carried out to determine if 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 score is based on the average of the two ratings. The next sections provide details of the process employed by ETS to score the writing tasks.

[9] 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 ETS 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 score 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. As well, each response is systematically routed for first and second reads and discrepancy resolutions, thereby assuring that they are 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 ETS 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, 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 ETS 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 rationale, 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 prescored 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, providing 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. Included in ETS's recruiting and training plans is 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 consist 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 powerful tool 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 of one test form to another. As a result of this process, scoring tables are produced. These tables map the current administration's raw score to an appropriate scale score. ETS's Scoring, Reporting, and Technology (SR&T) division utilizes these tables to generate scale scores for each student.

After score conversion is completed, both the SR&T 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, CAHSEE scores for a given content area are aggregated at the school, independently testing 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

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 numbers 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 Appendixes 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 Appendixes 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 was 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 Appendixes 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.[10] 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.

[10] Students' economic status was 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).

Similar summary statistics for demographic groups are presented in Tables 7 and 8 in Appendixes 8.A to 8.G. These summary statistics display the percentages of examinees classified as Below Proficient, Proficient, and Advanced according to the ESEA classifications. Selected percentiles, scale score means, and standard deviations for the subgroups are presented for all students in Tables 9 and 10 for ELA and mathematics, respectively.

Table 8.1 provides a listing of the frequency distribution and demographic summary tables found in the Appendices. To simplify the presentation of these data, all tables for this section are located in 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' achievement and if the students fulfill one or both part(s) of the CAHSEE for the 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 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 ◦ LEA-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; the school district forwards the appropriate reports to test sites or, in the case of the CAHSEE Student and Parent Report, sends the reports to the child's parents or guardians and forwards a copy to the student's school or test site. 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 students with disabilities and English learners 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 used a modification and earned the equivalent of a passing score on one or both parts of the CAHSEE did not pass, but may have been 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 Online 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 (Academic Performance Index) and federal (Adequate Yearly Progress) accountability reporting timeline requirements. CD-ROMs or DVDs containing these data files with encryption software are delivered to the CDE so that the CDE and LEAs 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 students with disabilities, no student receives 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 grade10 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 score scales are different for each content area. Comparisons between raw scores and cluster scores should be limited to comparisons within not only content area but also test administration. Comparing scores obtained in different content areas should be avoided because the results are not on the same scale. 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 2008

Table 8.A.1: Frequency Distributions, ELA—July 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	15	0	15	99
440-449	10	0	25	99
430-439	35	0	60	99
420-429	47	0	107	99
410-419	52	0	159	99
400-409	96	1	255	98
390-399	198	1	453	97
380-389	214	2	667	95
370-379	470	4	1137	92
360-369	1055	8	2192	84
350-359[1]	1761	13	3953	70
340-349	2218	17	6171	54
330-339	2241	17	8412	37
320-329	1769	13	10181	24
310-319	1344	10	11525	14
300-309	1074	8	12599	6
290-299	451	3	13050	2
280-289	169	1	13219	1
275-279	154	1	13373	0

[1]Passing Score = 350

Table 8.A.2: Frequency Distributions, Mathematics—July 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	28	0	28	99
440-449	14	0	42	99
430-439	22	0	64	99
420-429	31	0	95	99
410-419	46	0	141	99
400-409	81	1	222	98
390-399	121	1	343	97
380-389	216	2	559	96
370-379	443	3	1002	92
360-369	834	6	1836	86
350-359[1]	1978	15	3814	71
340-349	2746	21	6560	50
330-339	2366	18	8926	33
320-329	1999	15	10925	17
310-319	1419	11	12344	7
300-309	614	5	12958	2
290-299	225	2	13183	0
280-289	32	0	13215	0
275-279	22	0	13237	0

[1]Passing Score = 350

Table 8.A.3: Frequency Distributions, ELA for ESEA—July 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	15	0	15	99
440-449	10	0	25	99
430-439	35	0	60	99
420-429	47	0	107	99
410-419	52	0	159	99
403-409[1]	52	0	211	98
390-402	242	2	453	97
380-389[2]	214	2	667	95
370-379	470	4	1137	92
360-369	1055	8	2192	84
350-359	1761	13	3953	70
340-349	2218	17	6171	54
330-339	2241	17	8412	37
320-329	1769	13	10181	24
310-319	1344	10	11525	14
300-309	1074	8	12599	6
290-299	451	3	13050	2
280-289	169	1	13219	1
275-279	154	1	13373	0

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

Table 8.A.4: Frequency Distributions, Mathematics for ESEA—July 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	28	0	28	99
440-449	14	0	42	99
430-439	22	0	64	99
422-429[1]	12	0	76	99
410-421	65	0	141	99
400-409	81	1	222	98
390-399	121	1	343	97
380-389[2]	216	2	559	96
370-379	443	3	1002	92
360-369	834	6	1836	86
350-359	1978	15	3814	71
340-349	2746	21	6560	50
330-339	2366	18	8926	33
320-329	1999	15	10925	17
310-319	1419	11	12344	7
300-309	614	5	12958	2
290-299	225	2	13183	0
280-289	32	0	13215	0
275-279	22	0	13237	0

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

Table 8.A.5: Demographic Summary for All Examinees ELA—July 2008

	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	13373	3953	30	9420	70	337	59	55	55	42	55	2.0
Grade[3]												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-	-
Twelfth	10964	2902	26	8062	74	335	58	53	53	40	53	2.0
Adult Education	2409	1051	44	1358	56	349	65	62	61	50	61	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	7361	2100	29	5261	71	336	59	54	54	41	53	2.0
Female	6004	1851	31	4153	69	339	59	55	56	43	56	2.0
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	77	26	34	51	66	342	66	56	58	46	54	1.9
Asian	1406	333	24	1073	76	335	52	50	52	44	57	1.9
Pacific Islander	89	30	34	59	66	340	63	54	56	43	57	2.1
Filipino	280	95	34	185	66	342	60	55	56	49	58	2.0
Hispanic or Latino	8633	2357	27	6276	73	336	58	54	54	41	53	2.0
African American	1594	522	33	1072	67	338	62	56	56	40	54	2.0
White (not of Hispanic origin)	1095	513	47	582	53	349	69	61	62	47	60	2.1
Unknown	199	77	39	122	61	343	63	58	57	44	56	2.0
Language Fluency												
English-Only Students	3927	1391	35	2536	65	340	63	57	57	41	55	2.0
Initially Fluent English Proficient (IFEP)	421	164	39	257	61	341	63	58	57	42	57	2.0
Reclassified Fluent English Proficient (RFEP)	444	208	47	236	53	346	66	60	60	45	61	2.1
English-Learner Students	6528	1253	19	5275	81	331	54	50	50	40	52	1.9
Unknown	2053	937	46	1116	54	350	66	63	62	50	61	2.1
Economically Disadvantaged												
No	2021	691	34	1330	66	340	62	56	56	42	57	2.0
Yes	8152	2000	25	6152	75	334	57	52	52	40	53	2.0
Unknown	3200	1262	39	1938	61	345	63	60	59	47	58	2.0
Special Education Program Participation												
Students Receiving Services	2521	460	18	2061	82	327	54	48	49	36	47	1.8
Students Not Receiving Services	10852	3493	32	7359	68	340	60	56	56	44	56	2.0

[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 12 and Adult Education students may take the July administration of the CAHSEE.

Table 8.A.6: Demographic Summary for All Examinees, Mathematics—July 2008

	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	13237	3814	29	9423	71	340	52	49	48	45	36
Grade[3]											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-	-
Twelfth	10732	2802	26	7930	74	338	50	47	47	44	35
Adult Education	2505	1012	40	1493	60	347	57	53	53	51	39
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	6095	1754	29	4341	71	339	51	49	48	45	35
Female	7134	2056	29	5078	71	340	52	48	49	45	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity											
American Indian or Alaska Native	106	23	22	83	78	337	54	46	47	43	32
Asian	488	186	38	302	62	351	54	55	54	51	47
Pacific Islander	103	22	21	81	79	337	45	47	47	43	39
Filipino	227	82	36	145	64	344	51	51	52	47	43
Hispanic or Latino	8584	2376	28	6208	72	339	51	48	48	45	36
African American	2220	513	23	1707	77	336	50	47	46	42	34
White (not of Hispanic origin)	1250	525	42	725	58	348	59	54	53	50	37
Unknown	259	87	34	172	66	342	54	52	49	45	36
Language Fluency											
English-Only Students	5145	1428	28	3717	72	339	52	49	47	44	34
Initially Fluent English Proficient (IFEP)	607	211	35	396	65	342	54	51	50	46	37
Reclassified Fluent English Proficient (RFEP)	923	334	36	589	64	344	56	51	51	49	38
English-Learner Students	4347	925	21	3422	79	336	47	45	46	42	36
Unknown	2215	916	41	1299	59	348	58	54	53	51	39
Economically Disadvantaged											
No	2467	766	31	1701	69	341	53	50	49	45	37
Yes	7367	1812	25	5555	75	337	49	47	47	43	35
Unknown	3403	1236	36	2167	64	345	56	51	51	49	38
Special Education Program Participation											
Students Receiving Services	2145	344	16	1801	84	330	44	43	41	38	32
Students Not Receiving Services	11092	3470	31	7622	69	342	53	50	50	46	37

[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 12 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 2008

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		13373	12706	95	456	3	211	2	667	5
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	10964	10612	97	266	2	86	1	352	3
	Adult Education	2409	2094	87	190	8	125	5	315	13
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	7361	6980	95	264	4	117	2	381	5
	Female	6004	5718	95	192	3	94	2	286	5
	Unknown	-	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	77	71	92	3	4	3	4	6	8
	Asian	1406	1356	96	33	2	17	1	50	4
	Pacific Islander	89	85	96	4	4			4	4
	Filipino	280	265	95	12	4	3	1	15	5
	Hispanic or Latino	8633	8320	96	235	3	78	1	313	4
	African American	1594	1505	94	70	4	19	1	89	6
	White (not of Hispanic origin)	1095	926	85	89	8	80	7	169	15
	Unknown	199	178	89	10	5	11	6	21	11
Language Fluency	English-Only Students	3927	3660	93	186	5	81	2	267	7
	Initially Fluent English Proficient	421	392	93	23	5	6	1	29	7
	Reclassified Fluent English Proficient	444	414	93	23	5	7	2	30	7
	English-Learner Students	6528	6464	99	59	1	5	0	64	1
	Unknown	2053	1776	87	165	8	112	5	277	13
Economically Disadvantaged	No	2021	1903	94	76	4	42	2	118	6
	Yes	8152	7956	98	164	2	32	0	196	2
	Unknown	3200	2847	89	216	7	137	4	353	11
Special Education Program	Receiving Services	2521	2480	98	34	1	7	0	41	2
	Not Receiving Services	10852	10226	94	422	4	204	2	626	6

[1] Results for groups with less than 11 students are not reported.
[2] As of 2007, only Grade 12 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 2008

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		13237	12678	96	483	4	76	1	559	4
Grade[2]	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	-	-	-	-	-	-	-	-	-
	Twelfth	10732	10412	97	281	3	39	0	320	3
	Adult Education	2505	2266	90	202	8	37	1	239	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	6095	5781	95	272	4	42	1	314	5
	Female	7134	6889	97	211	3	34	0	245	3
	Unknown	-	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	106	103	97	3	3	0	0	3	3
	Asian	488	414	85	44	9	30	6	74	15
	Pacific Islander	103	98	95	5	5	0	0	5	5
	Filipino	227	214	94	13	6	0	0	13	6
	Hispanic or Latino	8584	8326	97	240	3	18	0	258	3
	African American	2220	2161	97	54	2	5	0	59	3
	White (not of Hispanic origin)	1250	1122	90	108	9	20	2	128	10
	Unknown	259	240	93	16	6	3	1	19	7
Language Fluency	English-Only Students	5145	4944	96	176	3	25	0	201	4
	Initially Fluent English Proficient	607	586	97	16	3	5	1	21	3
	Reclassified Fluent English Proficient	923	899	97	23	2	1	0	24	3
	English-Learner Students	4347	4255	98	81	2	11	0	92	2
	Unknown	2215	1994	90	187	8	34	2	221	10
Economically Disadvantaged	No	2467	2344	95	100	4	23	1	123	5
	Yes	7367	7198	98	156	2	13	0	169	2
	Unknown	3403	3136	92	227	7	40	1	267	8
Special Education Program	Receiving Services	2145	2103	98	38	2	4	0	42	2
	Not Receiving Services	11092	10575	95	445	4	72	1	517	5

[1] Results for groups with less than 11 students are not reported.
[2] As of 2007, only Grade 12 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 2008

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	279	298	321	336	352	379	413	337	26	13373
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	279	296	319	336	350	374	399	335	24	10964
Adult Education	282	306	330	346	365	405	437	349	29	2409
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	294	319	336	352	381	413	336	27	7361
Female	288	302	325	338	352	379	413	339	24	6004
Unknown	279	279	309	334	350	372	372	329	30	8
Race/Ethnicity										
American Indian or Alaska Native	290	292	325	338	358	399	450	342	30	77
Asian	279	294	319	334	348	374	413	335	25	1406
Pacific Islander	294	304	323	340	352	376	402	340	22	89
Filipino	286	302	325	340	356	383	413	342	24	280
Hispanic or Latino	282	298	321	336	350	374	402	336	24	8633
African American	275	294	321	338	354	381	405	338	27	1594
White (not of Hispanic origin)	275	298	329	346	369	413	443	349	33	1095
Unknown	275	294	319	342	360	405	437	343	33	199
Language Fluency										
English-Only Students	275	296	323	340	356	387	421	340	28	3927
Initially Fluent English Proficient (IFEP)	284	300	325	342	358	387	417	341	26	421
Reclassified Fluent English Proficient (RFEP)	275	304	334	348	358	384	417	346	24	444
English-Learner Students	279	296	317	332	346	365	379	331	21	6528
Unknown	284	308	332	346	365	405	437	350	29	2053
Economically Disadvantaged										
No	279	298	323	340	354	384	421	340	27	2021
Yes	279	296	319	334	348	369	393	334	23	8152
Unknown	275	300	327	342	360	402	431	345	30	3200
Special Education Program Participation										
Students Receiving Services	275	290	309	327	344	367	387	327	24	2521
Students Not Receiving Services	282	300	325	338	354	384	417	340	26	10852

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Table 8.A.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—July 2008

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	296	306	325	338	352	377	412	340	22	13237
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	-	-	-	-	-	-	-	-	-	-
Twelfth	294	306	323	338	350	373	402	338	21	10732
Adult Education	299	313	331	343	359	396	438	347	25	2505
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	294	306	323	338	352	381	416	339	24	6095
Female	296	309	327	340	352	375	409	340	21	7134
Unknown	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity										
American Indian or Alaska Native	299	309	325	337	349	375	381	337	19	106
Asian	299	311	329	343	361	426	450	351	34	488
Pacific Islander	299	309	323	334	349	379	381	337	20	103
Filipino	291	313	329	343	358	381	405	344	22	227
Hispanic or Latino	296	309	325	338	350	373	396	339	20	8584
African American	291	304	323	336	349	369	393	336	21	2220
White (not of Hispanic origin)	294	306	331	345	359	402	446	348	27	1250
Unknown	296	309	325	338	354	388	431	342	24	259
Language Fluency										
English-Only Students	294	306	323	338	350	377	405	339	22	5145
Initially Fluent English Proficient (IFEP)	296	311	329	343	354	377	409	342	22	607
Reclassified Fluent English Proficient (RFEP)	304	315	333	345	354	371	393	344	18	923
English-Learner Students	296	304	321	334	347	367	402	336	20	4347
Unknown	299	313	333	345	359	396	438	348	26	2215
Economically Disadvantaged										
No	296	309	327	340	352	379	421	341	23	2467
Yes	296	306	323	336	349	369	396	337	20	7367
Unknown	296	311	329	342	356	388	431	345	25	3403
Special Education Program Participation										
Students Receiving Services	291	299	315	329	343	365	393	330	21	2145
Students Not Receiving Services	296	309	327	340	352	379	416	342	22	11092

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.B: Frequency Distributions and Demographic Summaries—October 2008

Table 8.B.1: Frequency Distributions, ELA—October 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	161	0	161	99
440-449	180	0	341	99
430-439	109	0	450	99
420-429	300	1	750	98
410-419	571	1	1321	97
400-409	420	1	1741	96
390-399	1070	2	2811	94
380-389	1048	2	3859	91
370-379	2776	6	6635	85
360-369	3632	8	10267	77
350-359[1]	6102	14	16369	63
340-349	6733	15	23102	47
330-339	6096	14	29198	34
320-329	5684	13	34882	21
310-319	3447	8	38329	13
300-309	2866	7	41195	6
290-299	1660	4	42855	3
280-289	699	2	43554	1
275-279	401	1	43955	0

[1] Passing Score = 350

Table 8.B.2: Frequency Distributions, Mathematics—October 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	231	1	231	99
440-449	166	0	397	99
430-439	95	0	492	99
420-429	240	1	732	98
410-419	391	1	1123	98
400-409	454	1	1577	96
390-399	477	1	2054	95
380-389	1133	3	3187	93
370-379	1503	3	4690	90
360-369	4284	10	8974	80
350-359[1]	5973	13	14947	67
340-349	9158	20	24105	46
330-339	7384	16	31489	30
320-329	6136	14	37625	16
310-319	4544	10	42169	6
300-309	1921	4	44090	2
290-299	589	1	44679	1
280-289	159	0	44838	0
275-279	120	0	44958	0

[1] Passing Score = 350

Table 8.B.3: Frequency Distributions, ELA for ESEA—October 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	161	0	161	99
440-449	180	0	341	99
430-439	109	0	450	99
420-429	300	1	750	98
410-419	571	1	1321	97
403-409[1]	420	1	1741	96
390-402	1070	2	2811	94
380-389[2]	1048	2	3859	91
370-379	2776	6	6635	85
360-369	3632	8	10267	77
350-359	6102	14	16369	63
340-349	6733	15	23102	47
330-339	6096	14	29198	34
320-329	5684	13	34882	21
310-319	3447	8	38329	13
300-309	2866	7	41195	6
290-299	1660	4	42855	3
280-289	699	2	43554	1
275-279	401	1	43955	0

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

Table 8.B.4: Frequency Distributions, Mathematics for ESEA—October 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	231	1	231	99
440-449	166	0	397	99
430-439	95	0	492	99
422-429[1]	240	1	732	98
410-421	391	1	1123	98
400-409	454	1	1577	96
390-399	477	1	2054	95
380-389[2]	1133	3	3187	93
370-379	1503	3	4690	90
360-369	4284	10	8974	80
350-359	5973	13	14947	67
340-349	9158	20	24105	46
330-339	7384	16	31489	30
320-329	6136	14	37625	16
310-319	4544	10	42169	6
300-309	1921	4	44090	2
290-299	589	1	44679	1
280-289	159	0	44838	0
275-279	120	0	44958	0

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

Table 8.B.5: Demographic Summary for All Examinees, ELA—October 2008

	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	43955	16369	37	27586	63	342	62	60	59	45	56	2.0
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	18447	7537	41	10910	59	344	64	60	60	46	57	2.0
Twelfth	23197	7790	34	15407	66	340	60	58	57	44	55	2.0
Adult Education	2311	1042	45	1269	55	350	65	66	63	51	59	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	25376	8909	35	16467	65	340	63	59	58	43	54	2.0
Female	18543	7450	40	11093	60	345	62	60	60	47	59	2.1
Unknown	36	10	28	26	72	334	57	54	56	40	46	1.9
Race/Ethnicity												
American Indian or Alaska Native	383	176	46	207	54	346	70	64	62	45	56	2.0
Asian	2857	896	31	1961	69	339	58	56	55	45	58	2.0
Pacific Islander	376	158	42	218	58	346	68	61	60	47	57	2.2
Filipino	738	330	45	408	55	349	67	63	63	49	59	2.2
Hispanic or Latino	26372	8901	34	17471	66	339	59	58	57	44	55	2.0
African American	6069	2249	37	3820	63	341	65	59	59	44	53	2.0
White (not of Hispanic origin)	6475	3361	52	3114	48	355	72	67	65	50	61	2.2
Unknown	685	298	44	387	56	347	66	63	61	48	57	2.1
Language Fluency												
English-Only Students	18284	8225	45	10059	55	348	69	63	62	47	57	2.1
Initially Fluent English Proficient (IFEP)	1800	891	50	909	51	350	68	64	63	49	60	2.2
Reclassified Fluent English Proficient (RFEP)	2165	1276	59	889	41	354	71	67	66	51	64	2.2
English-Learner Students	19243	4906	25	14337	75	334	55	55	54	42	53	1.9
Unknown	2463	1071	43	1392	57	348	64	65	62	50	58	2.1
Economically Disadvantaged												
No	9685	4814	50	4871	50	352	70	65	64	50	60	2.2
Yes	28089	8983	32	19106	68	338	59	57	57	43	54	2.0
Unknown	6181	2572	42	3609	58	346	64	62	60	48	57	2.1
Special Education Program Participation												
Students Receiving Services	9330	1752	19	7578	81	327	55	51	50	36	46	1.8
Students Not Receiving Services	34625	14617	42	20008	58	346	64	62	61	48	59	2.1

[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

Table 8.B.6: Demographic Summary for All Examinees, Mathematics—October 2008

	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	44958	14947	33	30011	67	343	58	52	50	43	36
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	18342	6640	36	11702	64	344	58	53	51	43	37
Twelfth	24221	7478	31	16743	69	341	58	51	49	42	35
Adult Education	2395	829	35	1566	65	345	62	54	50	46	34
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	22214	7142	32	15072	68	341	57	52	49	42	35
Female	22703	7792	34	14911	66	344	59	52	51	43	37
Unknown	41	13	32	28	68	338	54	49	47	39	36
Race/Ethnicity											
American Indian or Alaska Native	464	169	36	295	64	343	61	52	50	43	34
Asian	1579	809	51	770	49	361	60	62	60	54	50
Pacific Islander	343	115	34	228	66	342	57	52	51	41	36
Filipino	703	296	42	407	58	350	61	57	55	46	41
Hispanic or Latino	26846	7993	30	18853	70	340	57	50	49	41	35
African American	7339	2050	28	5289	72	338	56	51	47	39	34
White (not of Hispanic origin)	6922	3242	47	3680	53	352	66	58	56	48	40
Unknown	762	273	36	489	64	345	61	54	50	44	37
Language Fluency											
English-Only Students	21412	7796	36	13616	64	344	61	53	51	43	36
Initially Fluent English Proficient (IFEP)	2234	919	41	1315	59	348	63	55	53	45	38
Reclassified Fluent English Proficient (RFEP)	3461	1504	43	1957	57	348	65	54	55	45	38
English-Learner Students	15155	3861	25	11294	75	338	52	49	48	41	35
Unknown	2696	867	32	1829	68	343	60	53	49	44	34
Economically Disadvantaged											
No	10804	4692	43	6112	57	350	64	56	55	46	40
Yes	27503	8020	29	19483	71	339	56	50	48	41	34
Unknown	6651	2235	34	4416	66	344	60	53	50	44	36
Special Education Program Participation											
Students Receiving Services	7925	1300	16	6625	84	329	48	46	41	35	30
Students Not Receiving Services	37033	13647	37	23386	63	345	61	54	52	44	37

[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 2008

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		43955	40096	91	2118	5	1741	4	3859	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	18447	16686	90	1024	6	737	4	1761	10
	Twelfth	23197	21482	93	875	4	840	4	1715	7
	Adult Education	2311	1928	83	219	9	164	7	383	17
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	25376	23459	92	1140	4	777	3	1917	8
	Female	18543	16603	90	978	5	962	5	1940	10
	Unknown	36	34	94	-	-	2	6	2	6
Race/Ethnicity	American Indian or Alaskan Native	383	336	88	26	7	21	5	47	12
	Asian	2857	2624	92	116	4	117	4	233	8
	Pacific Islander	376	341	91	20	5	15	4	35	9
	Filipino	738	630	85	52	7	56	8	108	15
	Hispanic or Latino	26372	24936	95	960	4	476	2	1436	5
	African American	6069	5630	93	280	5	159	3	439	7
	White (not of Hispanic origin)	6475	5031	78	601	9	843	13	1444	22
	Unknown	685	568	83	63	9	54	8	117	17
Language Fluency	English-Only Students	18284	15699	86	1275	7	1310	7	2585	14
	Initially Fluent English Proficient (IFEP)	1800	1545	86	154	9	101	6	255	14
	Reclassified Fluent English Proficient (RFEP)	2165	1907	88	181	8	77	4	258	12
	English-Learner Students	19243	18850	98	291	2	102	1	393	2
	Unknown	2463	2095	85	217	9	151	6	368	15
Economically Disadvantaged	No	9685	7987	82	759	8	939	10	1698	18
	Yes	28089	26750	95	931	3	408	1	1339	5
	Unknown	6181	5359	87	428	7	394	6	822	13
Special Education Program Participation	Receiving Services	9330	9171	98	134	1	25	0	159	2
	Not Receiving Services	34625	30925	89	1984	6	1716	5	3700	11

[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 2008

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		44958	41771	93	2455	5	732	2	3187	7
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	18342	16913	92	1133	6	296	2	1429	8
	Twelfth	24221	22679	94	1140	5	402	2	1542	6
	Adult Education	2395	2179	91	182	8	34	1	216	9
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	22214	20499	92	1329	6	386	2	1715	8
	Female	22703	21233	94	1125	5	345	2	1470	6
	Unknown	41	39	95	1	2	1	2	2	5
Race/Ethnicity	American Indian or Alaskan Native	464	427	92	34	7	3	1	37	8
	Asian	1579	1181	75	205	13	193	12	398	25
	Pacific Islander	343	318	93	20	6	5	1	25	7
	Filipino	703	607	86	73	10	23	3	96	14
	Hispanic or Latino	26846	25776	96	941	4	129	0	1070	4
	African American	7339	7038	96	272	4	29	0	301	4
	White (not of Hispanic origin)	6922	5757	83	837	12	328	5	1165	17
	Unknown	762	667	88	73	10	22	3	95	12
Language	English-Only Students	21412	19447	91	1504	7	461	2	1965	9
	Initially Fluent English Proficient (IFEP)	2234	2007	90	170	8	57	3	227	10
	Reclassified Fluent English Proficient	3461	3291	95	144	4	26	1	170	5
	English-Learner Students	15155	14548	96	457	3	150	1	607	4
	Unknown	2696	2478	92	180	7	38	1	218	8
Economically Disadvantaged	No	10804	9333	86	1033	10	438	4	1471	14
	Yes	27503	26402	96	952	3	149	1	1101	4
	Unknown	6651	6036	91	470	7	145	2	615	9
Special Program	Receiving Services	7925	7774	98	137	2	14	0	151	2
	Not Receiving Services	37033	33997	92	2318	6	718	2	3036	8

[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 2008

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	280	297	323	340	359	396	434	342	29	43955
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	280	297	323	342	361	396	434	344	30	18447
Twelfth	280	297	322	338	354	393	434	340	29	23197
Adult Education	284	304	329	346	367	410	434	350	32	2311
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	277	294	320	338	356	390	423	340	29	25376
Female	284	302	327	342	361	403	440	345	30	18543
Unknown	275	282	309	329	351	403	423	334	33	36
Race/Ethnicity										
American Indian or Alaska Native	275	299	325	346	363	403	423	346	30	383
Asian	282	295	320	338	354	396	440	339	30	2857
Pacific Islander	290	301	328	344	361	393	418	346	27	376
Filipino	286	301	329	346	367	414	440	349	32	738
Hispanic or Latino	282	297	322	338	354	382	410	339	26	26372
African American	275	294	323	340	359	390	418	341	28	6069
White (not of Hispanic origin)	277	297	329	350	377	428	450	355	38	6475
Unknown	275	295	323	344	367	410	450	347	36	685
Language Fluency										
English-Only Students	277	297	325	346	365	410	447	348	33	18284
Initially Fluent English Proficient (IFEP)	284	303	331	348	367	403	444	350	31	1800
Reclassified Fluent English Proficient (RFEP)	294	314	340	354	367	393	423	354	24	2165
English-Learner Students	280	295	318	335	350	370	390	334	23	19243
Unknown	284	302	327	344	367	406	434	348	31	2463
Economically Disadvantaged										
No	282	301	329	348	370	418	450	352	34	9685
Yes	280	295	322	338	354	379	406	338	26	28089
Unknown	277	297	325	342	363	406	440	346	32	6181
Special Education Program Participation										
Students Receiving Services	275	288	309	327	344	367	387	327	24	9330
Students Not Receiving Services	284	302	327	344	361	399	434	346	29	34625

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Table 8.B.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—October 2008

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	295	308	326	341	354	387	433	343	26	44958
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	295	308	326	341	358	389	433	344	26	18342
Twelfth	295	308	326	340	354	387	433	341	25	24221
Adult Education	298	310	328	341	356	395	428	345	25	2395
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	292	305	324	340	354	392	433	341	27	22214
Female	298	310	328	341	354	385	433	344	24	22703
Unknown	283	298	318	338	352	365	423	338	28	41
Race/Ethnicity										
American Indian or Alaska Native	275	305	326	341	358	387	414	343	26	464
Asian	298	312	334	350	380	448	450	361	39	1579
Pacific Islander	292	305	326	340	354	389	433	342	26	343
Filipino	300	312	330	345	362	410	448	350	29	703
Hispanic or Latino	295	308	326	340	352	376	407	340	22	26846
African American	289	305	323	338	352	376	403	338	23	7339
White (not of Hispanic origin)	295	308	332	349	367	418	450	352	32	6922
Unknown	289	305	326	341	358	403	448	345	30	762
Language Fluency										
English-Only Students	292	305	326	341	358	397	440	344	28	21412
Initially Fluent English Proficient (IFEP)	298	310	330	345	360	400	450	348	27	2234
Reclassified Fluent English Proficient (RFEP)	305	318	336	347	358	378	414	348	20	3461
English-Learner Students	295	305	323	336	350	374	418	338	23	15155
Unknown	295	308	326	340	354	392	428	343	26	2696
Economically Disadvantaged										
No	298	310	330	345	362	414	450	350	31	10804
Yes	295	305	324	340	352	376	407	339	22	27503
Unknown	295	308	326	340	356	397	440	344	27	6651
Special Education Program Participation										
Students Receiving Services	286	298	314	328	343	365	389	329	21	7925
Students Not Receiving Services	298	310	328	343	356	392	440	345	26	37033

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.C: Frequency Distributions and Demographic Summaries—November 2008

Table 8.C.1: Frequency Distributions, ELA—November 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	637	1	637	99
440-449	695	1	1332	99
430-439	440	0	1772	98
420-429	1002	1	2774	97
410-419	1954	2	4728	96
400-409	1577	1	6305	94
390-399	2744	3	9049	92
380-389	3614	3	12663	88
370-379	6817	6	19480	82
360-369	9430	9	28910	73
350-359[1]	14656	13	43566	60
340-349	15210	14	58776	46
330-339	13541	12	72317	33
320-329	10920	10	83237	23
310-319	8496	8	91733	16
300-309	6813	6	98546	9
290-299	5682	5	104228	4
280-289	2583	2	106811	2
275-279	1911	2	108722	0

[1] Passing Score = 350

Table 8.C.2: Frequency Distributions, Mathematics—November 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	680	1	680	99
440-449	319	0	999	99
430-439	681	1	1680	98
420-429	807	1	2487	98
410-419	1285	1	3772	97
400-409	1392	1	5164	95
390-399	2111	2	7275	93
380-389	2832	3	10107	91
370-379	5416	5	15523	86
360-369	9208	8	24731	77
350-359[1]	16540	15	41271	62
340-349	16689	15	57960	47
330-339	20525	19	78485	28
320-329	14690	13	93175	15
310-319	10617	10	103792	5
300-309	4208	4	108000	1
290-299	965	1	108965	0
280-289	282	0	109247	0
275-279	198	0	109445	0

[1] Passing Score = 350

Table 8.C.3: Frequency Distributions, ELA for ESEA—November 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	637	1	637	99
440-449	695	1	1332	99
430-439	440	0	1772	98
420-429	1002	1	2774	97
410-419	1954	2	4728	96
403-409[1]	773	1	5501	95
390-402	3548	3	9049	92
380-389[2]	3614	3	12663	88
370-379	6817	6	19480	82
360-369	9430	9	28910	73
350-359	14656	13	43566	60
340-349	15210	14	58776	46
330-339	13541	12	72317	33
320-329	10920	10	83237	23
310-319	8496	8	91733	16
300-309	6813	6	98546	9
290-299	5682	5	104228	4
280-289	2583	2	106811	2
275-279	1911	2	108722	0

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

Table 8.C.4: Frequency Distributions, Mathematics for ESEA—November 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	680	1	680	99
440-449	319	0	999	99
430-439	681	1	1680	98
422-429[1]	807	1	2487	98
410-421	1285	1	3772	97
400-409	1392	1	5164	95
390-399	2111	2	7275	93
380-389[2]	2832	3	10107	91
370-379	5416	5	15523	86
360-369	9208	8	24731	77
350-359	16540	15	41271	62
340-349	16689	15	57960	47
330-339	20525	19	78485	28
320-329	14690	13	93175	15
310-319	10617	10	103792	5
300-309	4208	4	108000	1
290-299	965	1	108965	0
280-289	282	0	109247	0
275-279	198	0	109445	0

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

Table 8.C.5: Demographic Summary for All Examinees, ELA—November 2008

	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	108722	43566	40	65156	60	343	58	61	65	49	59	2.1
Grade												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	70967	30884	44	40083	56	345	59	62	67	50	60	2.1
Twelfth	33819	10852	32	22967	68	338	55	58	62	46	56	2.0
Adult Education	3936	1830	46	2106	54	351	64	66	69	54	61	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	63090	24349	39	38741	61	341	59	60	64	48	57	2.0
Female	45595	19209	42	26386	58	346	58	61	67	51	61	2.2
Unknown	37	8	22	29	78	330	52	56	58	43	48	1.9
Race/Ethnicity												
American Indian or Alaska Native	892	418	47	474	53	348	65	64	69	51	59	2.1
Asian	6905	2683	39	4222	61	343	52	59	65	52	62	2.1
Pacific Islander	793	351	44	442	56	347	58	61	67	52	60	2.2
Filipino	1972	947	48	1025	52	352	62	64	69	55	65	2.2
Hispanic or Latino	68703	23889	35	44814	65	339	55	59	63	47	57	2.0
African American	11063	4409	40	6654	60	341	59	59	65	48	56	2.1
White (not of Hispanic origin)	17154	10301	60	6853	40	361	70	70	73	58	66	2.3
Unknown	1240	568	46	672	54	349	62	63	67	52	60	2.1
Language Fluency												
English-Only Students	44497	22267	50	22230	50	351	65	65	69	53	61	2.2
Initially Fluent English Proficient (IFEP)	4798	2621	55	2177	45	354	66	66	71	54	64	2.2
Reclassified Fluent English Proficient (RFEP)	7809	4762	61	3047	39	355	67	68	73	55	67	2.2
English-Learner Students	48017	12182	25	35835	75	332	50	55	60	44	55	1.9
Unknown	3601	1734	48	1867	52	352	65	66	70	55	61	2.1
Economically Disadvantaged												
No	25191	13680	54	11511	46	356	66	67	71	55	64	2.2
Yes	68029	22863	34	45166	66	337	54	58	63	46	57	2.0
Unknown	15502	7023	45	8479	55	348	62	63	68	52	60	2.1
Special Education Program Participation												
Students Receiving Services	22402	4955	22	17447	78	327	50	52	56	41	49	1.9
Students Not Receiving Services	86320	38611	45	47709	55	347	60	63	68	51	61	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.C.6: Demographic Summary for All Examinees, Mathematics—November 2008

	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	109445	41271	38	68174	62	345	55	49	51	46	40
Grade											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	70769	28318	40	42451	60	346	55	50	52	47	41
Twelfth	34816	11554	33	23262	67	343	54	47	49	45	38
Adult Education	3860	1399	36	2461	64	346	57	48	52	50	36
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	54460	20296	37	34164	63	345	54	50	51	46	38
Female	54939	20964	38	33975	62	345	56	48	52	47	41
Unknown	46	11	24	35	76	333	45	42	44	37	36
Race/Ethnicity											
American Indian or Alaska Native	967	378	39	589	61	345	57	49	51	47	38
Asian	4065	2425	60	1640	40	368	61	63	63	59	56
Pacific Islander	839	358	43	481	57	349	56	52	54	47	44
Filipino	1879	967	51	912	49	356	59	56	57	53	48
Hispanic or Latino	68687	23008	33	45679	67	341	53	47	50	45	38
African American	13490	4046	30	9444	70	339	52	46	47	42	37
White (not of Hispanic origin)	18157	9539	53	8618	47	356	63	56	58	53	44
Unknown	1361	550	40	811	60	348	57	52	53	48	40
Language Fluency											
English-Only Students	51164	20864	41	30300	59	347	57	51	52	47	40
Initially Fluent English Proficient (IFEP)	5690	2640	46	3050	54	352	60	53	56	50	43
Reclassified Fluent English Proficient (RFEP)	10653	5253	49	5400	51	350	60	51	56	50	43
English-Learner Students	38332	11149	29	27183	71	340	49	46	48	44	38
Unknown	3606	1365	38	2241	62	347	58	49	53	50	37
Economically Disadvantaged											
No	27836	13268	48	14568	52	353	60	54	56	51	44
Yes	65583	21575	33	44008	67	341	52	46	49	44	38
Unknown	16026	6428	40	9598	60	348	57	50	53	48	40
Special Education Program Participation											
Students Receiving Services	18969	3553	19	15416	81	331	45	42	42	37	32
Students Not Receiving Services	90476	37718	42	52758	58	348	57	50	53	49	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.C.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—November 2008

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		108722	96059	88	7162	7	5501	5	1266	12
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	70967	62063	87	5186	7	3718	5	8904	13
	Twelfth	33819	30781	91	1554	5	1484	4	3038	9
	Adult Education	3936	3215	82	422	11	299	8	721	18
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	63090	56437	89	4073	6	2580	4	6653	11
	Female	45595	39587	87	3088	7	2920	6	6008	13
	Unknown	37	35	95	1	3	1	3	2	5
Race/Ethnicity	American Indian or Alaskan Native	892	746	84	79	9	67	8	146	16
	Asian	6905	6059	88	426	6	420	6	846	12
	Pacific Islander	793	670	84	62	8	61	8	123	16
	Filipino	1972	1600	81	176	9	196	10	372	19
	Hispanic or Latino	68703	63976	93	3275	5	1452	2	4727	7
	African American	11063	9898	89	755	7	410	4	1165	11
	White (not of Hispanic origin)	17154	12114	71	2281	13	2759	16	5040	29
	Unknown	1240	996	80	108	9	136	11	244	20
Language Fluency	English-Only Students	44497	35841	81	4462	10	4194	9	8656	19
	Initially Fluent English Proficient	4798	3856	80	520	11	422	9	942	20
	Reclassified Fluent English Proficient	7809	6644	85	824	11	341	4	1165	15
	English-Learner Students	48017	46809	97	955	2	253	1	1208	3
	Unknown	3601	2909	81	401	11	291	8	692	19
Economically Disadvantaged	No	25191	19391	77	2735	11	3065	12	5800	23
	Yes	68029	63824	94	2985	4	1220	2	4205	6
	Unknown	15502	12844	83	1442	9	1216	8	2658	17
Special Education Program	Receiving Services	22402	21744	97	510	2	148	1	658	3
	Not Receiving Services	86320	74315	86	6652	8	5353	6	1200	14

[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 2008

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		109445	99338	91	7620	7	2487	2	10107	9
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	70769	63834	90	5307	7	1628	2	6935	10
	Twelfth	34816	32029	92	2008	6	779	2	2787	8
	Adult Education	3860	3475	90	305	8	80	2	385	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	54460	49014	90	4058	7	1388	3	5446	10
	Female	54939	50279	92	3561	6	1099	2	4660	8
	Unknown	46	45	98	1	2	-	-	1	2
Race/Ethnicity	American Indian or Alaskan Native	967	881	91	69	7	17	2	86	9
	Asian	4065	2750	68	669	16	646	16	1315	32
	Pacific Islander	839	741	88	78	9	20	2	98	12
	Filipino	1879	1527	81	241	13	111	6	352	19
	Hispanic or Latino	68687	65151	95	3118	5	418	1	3536	5
	African American	13490	12792	95	597	4	101	1	698	5
	White (not of Hispanic origin)	18157	14337	79	2707	15	1113	6	3820	21
	Unknown	1361	1159	85	141	10	61	4	202	15
Language Fluency	English-Only Students	51164	44957	88	4606	9	1601	3	6207	12
	Initially Fluent English Proficient (IFEP)	5690	4882	86	538	9	270	5	808	14
	Reclassified Fluent English Proficient (RFEP)	10653	9797	92	743	7	113	1	856	8
	English-Learner Students	38332	36496	95	1430	4	406	1	1836	5
	Unknown	3606	3206	89	303	8	97	3	400	11
Economically Disadvantaged	No	27836	23053	83	3299	12	1484	5	4783	17
	Yes	65583	62242	95	2876	4	465	1	3341	5
	Unknown	16026	14043	88	1445	9	538	3	1983	12
Special Education Program Participation	Receiving Services	18969	18502	98	427	2	40	0	467	2
	Not Receiving Services	90476	80836	89	7193	8	2447	3	9640	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 2008

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	292	320	342	362	406	441	343	33	108722
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	275	292	322	344	364	406	441	345	33	70967
Twelfth	275	288	316	336	355	398	441	338	33	33819
Adult Education	282	301	328	346	371	414	441	351	34	3936
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	288	318	340	362	398	434	341	33	63090
Female	280	295	324	344	364	410	448	346	33	45595
Unknown	275	276	314	330	348	385	410	330	30	37
Race/Ethnicity										
American Indian or Alaska Native	275	290	324	346	366	414	448	348	36	892
Asian	275	290	320	340	362	410	450	343	35	6905
Pacific Islander	275	295	322	344	369	418	450	347	36	793
Filipino	282	301	328	348	371	423	450	352	35	1972
Hispanic or Latino	275	292	318	338	357	388	418	339	29	68703
African American	275	286	318	342	362	398	429	341	33	11063
White (not of Hispanic origin)	275	295	334	359	388	434	450	361	41	17154
Unknown	275	288	318	346	371	423	450	349	40	1240
Language Fluency										
English-Only Students	275	292	326	350	374	423	450	351	38	44497
Initially Fluent English Proficient (IFEP)	275	297	332	352	374	418	448	354	35	4798
Reclassified Fluent English Proficient (RFEP)	286	308	340	355	371	402	429	355	27	7809
English-Learner Students	275	290	314	334	350	371	395	332	26	48017
Unknown	276	301	328	348	371	414	441	352	34	3601
Economically Disadvantaged										
No	278	295	330	352	377	429	450	356	38	25191
Yes	275	290	318	338	355	385	414	337	29	68029
Unknown	275	292	324	346	369	418	448	348	37	15502
Special Education Program Participation										
Students Receiving Services	275	280	305	326	346	374	395	327	28	22402
Students Not Receiving Services	278	295	326	346	366	410	448	347	33	86320

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Table 8.C.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—November 2008

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	297	307	326	341	357	398	439	345	27	109445
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	297	310	328	343	359	398	439	346	27	70769
Twelfth	294	307	324	339	355	395	439	343	27	34816
Adult Education	302	312	328	341	357	395	439	346	26	3860
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	294	307	324	341	359	401	447	345	29	54460
Female	300	310	328	343	357	395	439	345	26	54939
Unknown	307	307	318	328	348	368	387	333	19	46
Race/Ethnicity										
American Indian or Alaska Native	297	307	326	343	359	398	433	345	27	967
Asian	300	314	338	357	395	450	450	368	41	4065
Pacific Islander	300	312	330	345	363	404	439	349	28	839
Filipino	300	312	334	350	370	427	450	356	33	1879
Hispanic or Latino	297	307	326	339	354	381	410	341	23	68687
African American	294	305	322	338	352	381	414	339	24	13490
White (not of Hispanic origin)	297	312	334	350	372	427	450	356	34	18157
Unknown	291	307	326	343	363	418	447	348	32	1361
Language Fluency										
English-Only Students	294	307	326	343	361	407	447	347	30	51164
Initially Fluent English Proficient (IFEP)	300	312	332	346	365	418	450	352	31	5690
Reclassified Fluent English Proficient (RFEP)	305	318	336	348	361	387	422	350	22	10653
English-Learner Students	297	307	324	338	352	378	422	340	24	38332
Unknown	302	312	330	343	359	404	447	347	28	3606
Economically Disadvantaged										
No	300	312	332	348	368	422	450	353	32	27836
Yes	297	307	324	339	354	381	414	341	23	65583
Unknown	297	310	328	343	361	410	450	348	30	16026
Special Education Program Participation										
Students Receiving Services	288	300	316	328	345	368	395	331	22	18969
Students Not Receiving Services	300	312	330	345	361	404	447	348	27	90476

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.D: Frequency Distributions and Demographic Summaries—December 2008

Table 8.D.1: Frequency Distributions, ELA—December 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	8	0	8	99
440-449	9	0	17	99
430-439	18	1	35	99
420-429	20	1	55	98
410-419	26	1	81	98
400-409	32	1	113	97
390-399	55	2	168	95
380-389	90	3	258	93
370-379	158	4	416	88
360-369	241	7	657	81
350-359[1]	416	12	1073	70
340-349	552	16	1625	54
330-339	559	16	2184	38
320-329	482	14	2666	25
310-319	336	9	3002	15
300-309	247	7	3249	8
290-299	164	5	3413	4
280-289	78	2	3491	2
275-279	53	2	3544	0

[1]Passing Score = 350

Table 8.D.2: Frequency Distributions, Mathematics—December 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	14	0	14	99
440-449	5	0	19	99
430-439	13	0	32	99
420-429	8	0	40	99
410-419	17	0	57	98
400-409	42	1	99	97
390-399	41	1	140	96
380-389	61	2	201	94
370-379	129	4	330	91
360-369	379	11	709	80
350-359[1]	550	16	1259	64
340-349	831	23	2090	41
330-339	675	19	2765	22
320-329	458	13	3223	9
310-319	242	7	3465	2
300-309	62	2	3527	0
290-299	11	0	3538	0
280-289	5	0	3543	0
275-279	1	0	3544	0

[1]Passing Score = 350

Table 8.D.3: Frequency Distributions, ELA for ESEA—December 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	8	0	8	99
440-449	9	0	17	99
430-439	18	1	35	99
420-429	20	1	55	98
410-419	26	1	81	98
403-409[1]	32	1	113	97
390-402	55	2	168	95
380-389[2]	90	3	258	93
370-379	158	4	416	88
360-369	241	7	657	81
350-359	416	12	1073	70
340-349	552	16	1625	54
330-339	559	16	2184	38
320-329	482	14	2666	25
310-319	336	9	3002	15
300-309	247	7	3249	8
290-299	164	5	3413	4
280-289	78	2	3491	2
275-279	53	2	3544	0

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

Table 8.D.4: Frequency Distributions, Mathematics for ESEA—December 2008

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	14	0	14	99
440-449	5	0	19	99
430-439	13	0	32	99
422-429[1]	4	0	36	99
410-421	21	1	57	98
400-409	42	1	99	97
390-399	41	1	140	96
380-389[2]	61	2	201	94
370-379	129	4	330	91
360-369	379	11	709	80
350-359	550	16	1259	64
340-349	831	23	2090	41
330-339	675	19	2765	22
320-329	458	13	3223	9
310-319	242	7	3465	2
300-309	62	2	3527	0
290-299	11	0	3538	0
280-289	5	0	3543	0
275-279	1	0	3544	0

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

Table 8.D.5: Demographic Summary for All Examinees, ELA—December 2008

	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	3544	1073	30	2471	70	338	61	57	60	48	56	2.1
Grade[3]												
Tenth	-	-	-	-	-	-	-	-	-	-	-	-
Eleventh	410	180	44	230	56	346	66	61	65	52	60	2.2
Twelfth	2345	575	25	1770	75	333	59	55	57	45	54	2.0
Adult Education	789	318	40	471	60	347	66	63	64	54	60	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	1837	534	29	1303	71	335	62	56	58	46	55	2.1
Female	1701	539	32	1162	68	340	61	58	61	50	58	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity												
American Indian or Alaska Native	24	12	50	12	50	347	73	63	65	52	61	2.0
Asian	291	71	24	220	76	333	58	54	57	45	55	2.0
Pacific Islander	23	7	30	16	70	337	57	56	58	47	57	2.2
Filipino	45	14	31	31	69	340	61	61	59	47	61	2.2
Hispanic or Latino	2222	601	27	1621	73	335	60	56	58	47	55	2.1
African American	514	177	34	337	66	340	65	58	61	47	56	2.2
White (not of Hispanic origin)	363	172	47	191	53	352	70	65	67	54	63	2.2
Unknown	62	19	31	43	69	341	65	58	59	54	59	2.1
Language Fluency												
English-Only Students	1025	360	35	665	65	341	65	58	62	48	57	2.2
Initially Fluent English Proficient (IFEP)	110	35	32	75	68	341	63	59	60	48	57	2.2
Reclassified Fluent English Proficient (RFEP)	108	49	45	59	55	346	63	63	65	53	63	2.1
English-Learner Students	1549	315	20	1234	80	330	56	53	55	44	53	2.0
Unknown	752	314	42	438	58	348	67	64	65	54	60	2.2
Economically Disadvantaged												
No	495	193	39	302	61	345	66	60	64	50	59	2.2
Yes	2100	511	24	1589	76	333	59	54	57	45	54	2.0
Unknown	949	369	39	580	61	345	66	62	63	52	59	2.1
Special Education Program Participation												
Students Receiving Services	590	107	18	483	82	326	57	49	52	41	49	2.0
Students Not Receiving Services	2954	966	33	1988	67	340	62	59	61	49	57	2.1

[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] Grade 10 students can only take CAHSEE one time in the spring during the February, March or May administration

Table 8.D.6: Demographic Summary for All Examinees, Mathematics—December 2008

	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	3544	1259	36	2285	64	345	54	49	50	46	35
Grade[3]											
Tenth	-	-	-	-	-	-	-	-	-	-	-
Eleventh	337	152	45	185	55	350	55	54	54	46	40
Twelfth	2360	750	32	1610	68	343	52	47	49	45	34
Adult Education	847	357	42	490	58	349	57	51	52	50	36
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	1531	583	38	948	62	346	54	50	51	46	35
Female	2010	676	34	1334	66	345	53	47	50	46	35
Unknown	-	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity											
American Indian or Alaska Native	26	10	38	16	62	344	57	48	49	46	31
Asian	136	52	38	84	62	351	51	51	53	49	44
Pacific Islander	27	11	41	16	59	346	54	48	52	50	34
Filipino	47	19	40	28	60	349	52	52	57	48	37
Hispanic or Latino	2127	716	34	1411	66	344	52	47	50	46	35
African American	710	226	32	484	68	342	52	47	48	43	33
White (not of Hispanic origin)	405	202	50	203	50	355	64	55	56	50	37
Unknown	66	23	35	43	65	340	47	47	47	43	31
Language Fluency											
English-Only Students	1329	465	35	864	65	345	55	49	50	45	34
Initially Fluent English Proficient (IFEP)	137	65	47	72	53	352	59	53	55	49	39
Reclassified Fluent English Proficient (RFEP)	188	91	48	97	52	352	60	51	55	50	39
English-Learner Students	1086	283	26	803	74	340	47	44	47	43	35
Unknown	804	355	44	449	56	351	57	52	53	51	37
Economically Disadvantaged											
No	558	224	40	334	60	348	56	51	52	46	36
Yes	1976	615	31	1361	69	342	52	46	48	44	34
Unknown	1010	420	42	590	58	349	56	51	53	49	36
Special Education Program Participation											
Students Receiving Services	482	88	18	394	82	335	46	44	43	38	31
Students Not Receiving Services	3062	1171	38	1891	62	347	55	49	51	47	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] Grade 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Table 8.D.7: ESEA Demographic Summary for All CAHSEE Examinees, ELA—December 2008

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		3544	3286	93	145	4	113	3	258	7
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	410	357	87	28	7	25	6	53	13
	Twelfth	2345	2257	96	56	2	32	1	88	4
	Adult Education	789	672	85	61	8	56	7	117	15
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	1837	1711	93	78	4	48	3	126	7
	Female	1701	1569	92	67	4	65	4	132	8
	Unknown	-	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	24	21	88	2	8	1	4	3	13
	Asian	291	274	94	9	3	8	3	17	6
	Pacific Islander	23	20	87	2	9	1	4	3	13
	Filipino	45	43	96	1	2	1	2	2	4
	Hispanic or Latino	2222	2114	95	73	3	35	2	108	5
	African American	514	479	93	21	4	14	3	35	7
	White (not of Hispanic origin)	363	279	77	35	10	49	13	84	23
	Unknown	62	56	90	2	3	4	6	6	10
Language	English-Only Students	1025	931	91	55	5	39	4	94	9
	Initially Fluent English Proficient (IFEP)	110	95	86	8	7	7	6	15	14
	Reclassified Fluent English Proficient	108	100	93	6	6	2	2	8	7
	English-Learner Students	1549	1528	99	16	1	5	0	21	1
	Unknown	752	632	84	60	8	60	8	120	16
Economically Disadvantaged	No	495	431	87	32	6	32	6	64	13
	Yes	2100	2034	97	49	2	17	1	66	3
	Unknown	949	821	87	64	7	64	7	128	13
Special Program	Receiving Services	590	581	98	8	1	1	0	9	2
	Not Receiving Services	2954	2705	92	137	5	112	4	249	8

[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 2008

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		3544	3343	94	165	5	36	1	201	6
Grade	Tenth	-	-	-	-	-	-	-	-	-
	Eleventh	337	307	91	21	6	9	3	30	9
	Twelfth	2360	2276	96	70	3	14	1	84	4
	Adult Education	847	760	90	74	9	13	2	87	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	1531	1422	93	91	6	18	1	109	7
	Female	2010	1918	95	74	4	18	1	92	5
	Unknown	-	-	-	-	-	-	-	-	-
Race/Ethnicity	American Indian or Alaskan Native	26	24	92	2	8	-	-	2	8
	Asian	136	114	84	14	10	8	6	22	16
	Pacific Islander	27	26	96	1	4	-	-	1	4
	Filipino	47	43	91	4	9	-	-	4	9
	Hispanic or Latino	2127	2039	96	75	4	13	1	88	4
	African American	710	691	97	18	3	1	0	19	3
	White (not of Hispanic origin)	405	342	84	49	12	14	3	63	16
	Unknown	66	64	97	2	3	-	-	2	3
Language Fluency	English-Only Students	1329	1259	95	59	4	11	1	70	5
	Initially Fluent English Proficient	137	120	88	14	10	3	2	17	12
	Reclassified Fluent English	188	178	95	7	4	3	2	10	5
	English-Learner Students	1086	1070	99	11	1	5	0	16	1
	Unknown	804	716	89	74	9	14	2	88	11
Economically Disadvantaged	No	558	511	92	37	7	10	2	47	8
	Yes	1976	1918	97	51	3	7	0	58	3
	Unknown	1010	914	90	77	8	19	2	96	10
Special Education Program	Receiving Services	482	475	99	5	1	2	0	7	1
	Not Receiving Services	3062	2868	94	160	5	34	1	194	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 2008

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	292	320	336	352	389	425	338	29	3544
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	278	292	324	344	366	407	443	346	34	410
Twelfth	275	290	316	334	348	375	407	333	26	2345
Adult Education	280	302	324	342	363	411	436	347	32	789
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	288	318	334	352	386	420	335	29	1837
Female	284	296	322	338	354	393	436	340	29	1701
Unknown	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity										
American Indian or Alaska Native	286	290	323	349	366	396	450	347	35	24
Asian	275	286	316	332	348	386	407	333	28	291
Pacific Islander	294	294	312	332	352	393	415	337	33	23
Filipino	292	300	322	340	352	373	411	340	25	45
Hispanic or Latino	275	294	318	334	350	378	415	335	27	2222
African American	275	288	322	340	356	389	425	340	29	514
White (not of Hispanic origin)	275	294	324	346	375	430	450	352	39	363
Unknown	292	294	322	336	359	403	420	341	30	62
Language Fluency										
English-Only Students	275	294	322	340	356	396	436	341	31	1025
Initially Fluent English Proficient (IFEP)	275	288	318	340	352	403	436	341	35	110
Reclassified Fluent English Proficient (RFEP)	286	292	328	348	363	384	430	346	28	108
English-Learner Students	275	290	314	332	344	366	384	330	23	1549
Unknown	284	302	326	344	366	411	436	348	32	752
Economically Disadvantaged										
No	275	294	322	342	363	407	443	345	33	495
Yes	275	290	316	334	348	373	398	333	25	2100
Unknown	275	296	324	342	363	407	436	345	32	949
Special Education Program Participation										
Students Receiving Services	275	284	308	326	342	368	384	326	25	590
Students Not Receiving Services	275	294	320	338	354	396	430	340	29	2954

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Table 8.D.10: Examinee Demographics Showing Mean Scale Score at Each Percentile, Mathematics—December 2008

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	305	315	331	344	356	383	425	345	22	3544
Grade[4]										
Tenth	-	-	-	-	-	-	-	-	-	-
Eleventh	303	312	333	347	363	400	450	350	26	337
Twelfth	303	315	331	342	354	373	413	343	20	2360
Adult Education	305	317	334	345	361	395	430	349	24	847
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	303	315	331	344	358	390	425	346	23	1531
Female	305	317	333	344	354	379	416	345	21	2010
Unknown	-	-	-	-	-	-	-	-	-	-
Race/Ethnicity										
American Indian or Alaska Native	305	308	327	345	361	381	403	344	25	26
Asian	305	315	331	344	363	435	450	351	32	136
Pacific Islander	310	315	336	349	354	373	392	346	18	27
Filipino	308	312	338	345	358	385	413	349	20	47
Hispanic or Latino	308	317	331	342	354	375	413	344	20	2127
African American	297	312	331	341	352	373	390	342	18	710
White (not of Hispanic origin)	308	319	336	349	367	416	450	355	29	405
Unknown	295	305	327	338	354	373	403	340	20	66
Language Fluency										
English-Only Students	303	315	331	344	354	381	420	345	21	1329
Initially Fluent English Proficient (IFEP)	303	319	338	349	360	403	450	352	26	137
Reclassified Fluent English Proficient (RFEP)	317	327	340	349	361	381	442	352	20	188
English-Learner Students	303	312	329	338	351	367	395	340	18	1086
Unknown	308	319	334	345	363	397	435	351	25	804
Economically Disadvantaged										
No	305	317	333	345	358	392	450	348	24	558
Yes	303	315	331	342	352	371	406	342	19	1976
Unknown	308	317	334	345	360	395	435	349	24	1010
Special Education Program Participation										
Students Receiving Services	295	308	321	334	347	365	392	335	19	482
Students Not Receiving Services	308	317	333	344	356	385	430	347	22	3062

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.E: Frequency Distributions and Demographic Summaries—February 2009

Table 8.E.1: Frequency Distributions, ELA—February 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	7532	4	7532	96
440-449	3505	2	11037	94
430-439	8335	4	19372	90
420-429	9677	5	29049	85
410-419	10095	5	39144	80
400-409	15349	8	54493	72
390-399	14517	7	69010	65
380-389	13347	7	82357	58
370-379	15879	8	98236	50
360-369	17951	9	116187	40
350-359[1]	13654	7	129841	33
340-349	16214	8	146055	25
330-339	14670	8	160725	18
320-329	11753	6	172478	12
310-319	10438	5	182916	6
300-309	5830	3	188746	3
290-299	3444	2	192190	1
280-289	1673	1	193863	1
275-279	1228	1	195091	0

[1] Passing Score = 350

Table 8.E.2: Frequency Distributions, Mathematics—February 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9981	5	9981	95
440-449	3833	2	13814	93
430-439	8066	4	21880	89
420-429	7961	4	29841	85
410-419	11820	6	41661	79
400-409	11493	6	53154	73
390-399	10994	6	64148	67
380-389	17347	9	81495	59
370-379	13280	7	94775	52
360-369	19254	10	114029	42
350-359[1]	16980	9	131009	33
340-349	21278	11	152287	22
330-339	16246	8	168533	14
320-329	13318	7	181851	7
310-319	9284	5	191135	3
300-309	3716	2	194851	1
290-299	1155	1	196006	0
280-289	186	0	196192	0
275-279	195	0	196387	0

[1] Passing Score = 350

Table 8.E.3: Frequency Distributions, ELA for ESEA—February 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	7532	4	7532	96
440-449	3505	2	11037	94
430-439	8335	4	19372	90
420-429	9677	5	29049	85
410-419	10095	5	39144	80
403-409[1]	10345	5	49489	75
390-402	19521	10	69010	65
380-389[2]	13347	7	82357	58
370-379	15879	8	98236	50
360-369	17951	9	116187	40
350-359	13654	7	129841	33
340-349	16214	8	146055	25
330-339	14670	8	160725	18
320-329	11753	6	172478	12
310-319	10438	5	182916	6
300-309	5830	3	188746	3
290-299	3444	2	192190	1
280-289	1673	1	193863	1
275-279	1228	1	195091	0

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

Table 8.E.4: Frequency Distributions, Mathematics for ESEA—February 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	9981	5	9981	95
440-449	3833	2	13814	93
430-439	8066	4	21880	89
422-429[1]	7961	4	29841	85
410-421	11820	6	41661	79
400-409	11493	6	53154	73
390-399	10994	6	64148	67
380-389[2]	17347	9	81495	59
370-379	13280	7	94775	52
360-369	19254	10	114029	42
350-359	16980	9	131009	33
340-349	21278	11	152287	22
330-339	16246	8	168533	14
320-329	13318	7	181851	7
310-319	9284	5	191135	3
300-309	3716	2	194851	1
290-299	1155	1	196006	0
280-289	186	0	196192	0
275-279	195	0	196387	0

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

Table 8.E.5: Demographic Summary for All Examinees, ELA February 2009

	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	195091	129841	67	65250	33	371	74	73	72	65	71	2.4
Grade												
Tenth	146088	116875	80	29213	20	383	80	78	78	72	76	2.5
Eleventh	21240	6227	29	15013	71	337	58	58	56	45	55	2.0
Twelfth	24502	5444	22	19058	78	334	56	56	53	43	53	2.0
Adult Education	3261	1295	40	1966	60	348	64	64	62	52	59	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	103506	64437	62	39069	38	366	73	71	70	62	68	2.3
Female	91530	65380	71	26150	29	378	75	75	75	67	74	2.5
Unknown	55	24	44	31	56	350	65	62	62	52	59	2.1
Race/Ethnicity												
American Indian or Alaska Native	1936	1303	67	633	33	369	75	73	72	65	69	2.3
Asian	17323	13051	75	4272	25	385	77	77	78	72	77	2.6
Pacific Islander	1504	997	66	507	34	368	73	71	71	64	70	2.4
Filipino	5938	4931	83	1007	17	388	80	79	80	74	78	2.7
Hispanic or Latino	88820	47530	54	41290	46	357	67	67	66	57	65	2.2
African American	18786	10428	56	8358	44	358	70	68	66	57	63	2.3
White (not of Hispanic origin)	58360	49953	86	8407	14	393	84	82	82	77	80	2.6
Unknown	2424	1648	68	776	32	374	76	74	74	66	71	2.4
Language Fluency												
English-Only Students	110335	84204	76	26131	24	381	79	77	77	70	75	2.5
Initially Fluent English Proficient (IFEP)	13355	10956	82	2399	18	387	80	80	79	73	78	2.6
Reclassified Fluent English Proficient (RFEP)	24768	21600	87	3168	13	384	79	80	79	73	79	2.5
English-Learner Students	43464	11776	27	31688	73	336	56	57	55	44	55	2.0
Unknown	3169	1305	41	1864	59	348	64	64	62	52	59	2.1
Economically Disadvantaged												
No	77637	65470	84	12167	16	392	83	81	81	76	80	2.6
Yes	98490	52858	54	45632	46	356	67	67	66	57	65	2.2
Unknown	18964	11513	61	7451	39	366	72	71	70	62	68	2.3
Special Education Program Participation												
Students Receiving Services	22803	6158	27	16645	73	334	57	56	53	44	50	2.0
Students Not Receiving Services	172288	123683	72	48605	28	376	76	75	75	67	74	2.5

[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

Table 8.E.6: Demographic Summary for All Examinees, Mathematics—February 2009

	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	196387	131009	67	65378	33	373	70	68	68	61	54
Grade											
Tenth	145605	116338	80	29267	20	384	76	74	74	67	62
Eleventh	22059	6858	31	15201	69	341	53	51	50	45	34
Twelfth	25171	6539	26	18632	74	339	51	49	49	44	33
Adult Education	3552	1274	36	2278	64	346	56	52	53	51	34
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	99977	66746	67	33231	33	373	70	69	68	62	54
Female	96345	64241	67	32104	33	372	70	67	68	61	55
Unknown	65	22	34	43	66	344	54	53	50	47	35
Race/Ethnicity											
American Indian or Alaska Native	2008	1270	63	738	37	366	68	65	65	58	49
Asian	15363	13532	88	1831	12	404	80	83	82	77	75
Pacific Islander	1517	1031	68	486	32	370	68	67	67	60	55
Filipino	5892	4980	85	912	15	390	77	77	78	70	66
Hispanic or Latino	89475	49177	55	40298	45	360	63	61	61	55	46
African American	20543	10038	49	10505	51	354	60	59	58	50	43
White (not of Hispanic origin)	59099	49368	84	9731	16	390	79	77	77	70	64
Unknown	2490	1613	65	877	35	372	70	68	67	61	52
Language Fluency											
English-Only Students	115012	82824	72	32188	28	378	73	71	71	63	57
Initially Fluent English Proficient (IFEP)	14037	10977	78	3060	22	386	76	74	75	68	62
Reclassified Fluent English Proficient (RFEP)	26179	21805	83	4374	17	386	77	75	76	68	63
English-Learner Students	37711	14122	37	23589	63	346	52	54	53	48	39
Unknown	3448	1281	37	2167	63	347	57	53	54	51	34
Economically Disadvantaged											
No	78996	64817	82	14179	18	390	78	77	77	70	65
Yes	97688	54692	56	42996	44	361	63	62	62	55	47
Unknown	19703	11500	58	8203	42	366	66	64	64	58	49
Special Education Program Participation											
Students Receiving Services	20468	5737	28	14731	72	339	50	50	47	42	34
Students Not Receiving Services	175919	125272	71	50647	29	377	72	70	70	63	57

[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 2009

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		195091	112734	58	32868	17	49489	25	82357	42
Grade	Tenth	146088	66680	46	31133	21	48275	33	79408	54
	Eleventh	21240	19814	93	871	4	555	3	1426	7
	Twelfth	24502	23469	96	602	2	431	2	1033	4
	Adult Education	3261	2771	85	262	8	228	7	490	15
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	103506	65102	63	16740	16	21664	21	38404	37
	Female	91530	47589	52	16124	18	27817	30	43941	48
	Unknown	55	43	78	4	7	8	15	12	22
Race/Ethnicity	American Indian or Alaskan Native	1936	1175	61	358	18	403	21	761	39
	Asian	17323	7609	44	2836	16	6878	40	9714	56
	Pacific Islander	1504	930	62	277	18	297	20	574	38
	Filipino	5938	2359	40	1362	23	2217	37	3579	60
	Hispanic or Latino	88820	65747	74	12231	14	10842	12	23073	26
	African American	18786	13680	73	2671	14	2435	13	5106	27
	White (not of Hispanic origin)	58360	19946	34	12716	22	25698	44	38414	66
	Unknown	2424	1288	53	417	17	719	30	1136	47
Language	English-Only Students	110335	51659	47	21597	20	37079	34	58676	53
	Initially Fluent English Proficient	13355	5623	42	2856	21	4876	37	7732	58
	Reclassified Fluent English	24768	11420	46	6538	26	6810	27	13348	54
	English-Learner Students	43464	41365	95	1614	4	485	1	2099	5
	Unknown	3169	2667	84	263	8	239	8	502	16
Economically Disadvantaged	No	77637	27750	36	16303	21	33584	43	49887	64
	Yes	98490	73049	74	13727	14	11714	12	25441	26
	Unknown	18964	11935	63	2838	15	4191	22	7029	37
Special Program	Receiving Services	22803	20977	92	1168	5	658	3	1826	8
	Not Receiving Services	172288	91757	53	31700	18	48831	28	80531	47

[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 2009

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		196387	114892	59	51654	26	29841	15	81495	41
Grade	Tenth	145605	66587	46	49618	34	29400	20	79018	54
	Eleventh	22059	20856	95	990	4	213	1	1203	5
	Twelfth	25171	24198	96	788	3	185	1	973	4
	Adult Education	3552	3251	92	258	7	43	1	301	8
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	99977	57281	57	26504	27	16192	16	42696	43
	Female	96345	57552	60	25147	26	13646	14	38793	40
	Unknown	65	59	91	3	5	3	5	6	9
Race/Ethnicity	American Indian or Alaskan Native	2008	1308	65	510	25	190	9	700	35
	Asian	15363	4187	27	4758	31	6418	42	11176	73
	Pacific Islander	1517	946	62	405	27	166	11	571	38
	Filipino	5892	2247	38	2245	38	1400	24	3645	62
	Hispanic or Latino	89475	65871	74	18038	20	5566	6	23604	26
	African American	20543	16187	79	3480	17	876	4	4356	21
	White (not of Hispanic origin)	59099	22665	38	21590	37	14844	25	36434	62
	Unknown	2490	1481	59	628	25	381	15	1009	41
Language Fluency	English-Only Students	115012	60169	52	34099	30	20744	18	54843	48
	Initially Fluent English Proficient (IFEP)	14037	6401	46	4293	31	3343	24	7636	54
	Reclassified Fluent English Proficient	26179	11813	45	9379	36	4987	19	14366	55
	English-Learner Students	37711	33411	89	3592	10	708	2	4300	11
	Unknown	3448	3098	90	291	8	59	2	350	10
Economically Disadvantaged	No	78996	31151	39	27281	35	20564	26	47845	61
	Yes	97688	70614	72	20160	21	6914	7	27074	28
	Unknown	19703	13127	67	4213	21	2363	12	6576	33
Special Education Program	Receiving Services	20468	18594	91	1515	7	359	2	1874	9
	Not Receiving Services	175919	96298	55	50139	29	29482	17	79621	45

[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 2009

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	285	306	339	371	404	445	450	371	41	195091
Grade[4]										
Tenth	293	316	357	384	412	450	450	383	38	146088
Eleventh	275	293	319	337	353	387	421	337	28	21240
Twelfth	275	295	318	333	347	376	416	334	26	24502
Adult Education	289	306	327	343	364	412	445	348	31	3261
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	281	300	335	364	396	438	450	366	41	103506
Female	293	312	345	379	407	450	450	378	41	91530
Unknown	275	295	319	341	366	438	450	350	42	55
Race/Ethnicity										
American Indian or Alaska Native	287	304	341	369	396	438	450	369	39	1936
Asian	289	310	351	387	421	450	450	385	44	17323
Pacific Islander	287	306	339	366	393	432	450	368	38	1504
Filipino	300	321	362	390	416	450	450	388	37	5938
Hispanic or Latino	283	302	331	353	381	421	445	357	36	88820
African American	277	299	331	355	384	421	450	358	37	18786
White (not of Hispanic origin)	293	321	369	396	421	450	450	393	38	58360
Unknown	275	300	341	376	407	450	450	374	44	2424
Language Fluency										
English-Only Students	287	310	351	384	412	450	450	381	41	110335
Initially Fluent English Proficient (IFEP)	295	319	360	387	416	450	450	387	38	13355
Reclassified Fluent English Proficient (RFEP)	308	335	364	384	404	438	450	384	31	24768
English-Learner Students	279	295	319	335	351	379	404	336	26	43464
Unknown	279	304	325	343	366	412	445	348	33	3169
Economically Disadvantaged										
No	297	321	366	396	421	450	450	392	39	77637
Yes	281	300	331	353	381	421	445	356	36	98490
Unknown	279	302	335	362	396	438	450	366	42	18964
Special Education Program Participation										
Students Receiving Services	275	287	312	331	351	390	421	334	31	22803
Students Not Receiving Services	291	312	347	376	407	445	450	376	40	172288

[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 10 students can only take 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 2009

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	301	315	342	367	403	450	450	373	40	196387
Grade[4]										
Tenth	306	321	356	384	414	450	450	384	38	145605
Eleventh	296	308	325	340	354	380	418	341	23	22059
Twelfth	296	308	325	338	351	373	414	339	21	25171
Adult Education	304	315	331	343	356	389	427	346	23	3552
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	299	313	342	369	403	450	450	373	41	99977
Female	304	317	342	367	400	447	450	372	38	96345
Unknown	299	308	321	340	354	400	450	344	31	65
Race/Ethnicity										
American Indian or Alaska Native	299	313	338	364	392	439	450	366	37	2008
Asian	311	331	376	410	439	450	450	404	40	15363
Pacific Islander	301	315	343	367	394	439	450	370	36	1517
Filipino	313	329	362	392	418	450	450	390	37	5892
Hispanic or Latino	301	313	334	354	380	427	450	360	34	89475
African American	296	308	331	349	373	418	447	354	33	20543
White (not of Hispanic origin)	308	325	362	392	422	450	450	390	38	59099
Unknown	296	315	340	365	400	450	450	372	40	2490
Language Fluency										
English-Only Students	301	317	347	376	410	450	450	378	40	115012
Initially Fluent English Proficient (IFEP)	306	323	354	384	418	450	450	386	40	14037
Reclassified Fluent English Proficient (RFEP)	315	333	360	384	410	450	450	386	35	26179
English-Learner Students	299	308	327	342	360	397	433	346	27	37711
Unknown	301	313	331	343	358	394	433	347	25	3448
Economically Disadvantaged										
No	308	325	360	392	422	450	450	390	39	78996
Yes	299	313	334	354	382	427	450	361	35	97688
Unknown	299	313	336	358	392	447	450	366	39	19703
Special Education Program Participation										
Students Receiving Services	290	304	319	334	352	394	433	339	29	20468
Students Not Receiving Services	306	319	345	373	406	450	450	377	39	175919

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.F: Frequency Distributions and Demographic Summaries—March 2009

Table 8.F.1: Frequency Distributions, ELA—March 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	13087	4	13087	96
440-449	5998	2	19085	95
430-439	15260	4	34345	91
420-429	9057	2	43402	88
410-419	29937	8	73339	80
400-409	21199	6	94538	75
390-399	31916	9	126454	66
380-389	40500	11	166954	55
370-379	36498	10	203452	45
360-369	32263	9	235715	37
350-359[1]	34885	9	270600	27
340-349	29126	8	299726	20
330-339	23128	6	322854	13
320-329	16789	5	339643	9
310-319	14345	4	353988	5
300-309	8379	2	362367	3
290-299	5609	2	367976	1
280-289	2813	1	370789	1
275-279	1918	1	372707	0

[1]Passing Score = 350

Table 8.F.2: Frequency Distributions, Mathematics—March 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	25088	7	25088	93
440-449	15389	4	40477	89
430-439	7868	2	48345	87
420-429	15392	4	63737	83
410-419	24046	6	87783	76
400-409	24184	7	111967	70
390-399	24349	7	136316	63
380-389	38479	10	174795	53
370-379	28657	8	203452	45
360-369	39875	11	243327	35
350-359[1]	31088	8	274415	26
340-349	33834	9	308249	17
330-339	23294	6	331543	11
320-329	18599	5	350142	6
310-319	13416	4	363558	2
300-309	5714	2	369272	1
290-299	1834	0	371106	0
280-289	324	0	371430	0
275-279	274	0	371704	0

[1] Passing Score = 350

Table 8.F.3: Frequency Distributions, ELA for ESEA—March 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	13087	4	13087	96
440-449	5998	2	19085	95
430-439	15260	4	34345	91
420-429	9057	2	43402	88
410-419	29937	8	73339	80
403-409[1]	10609	3	83948	77
390-402	42506	11	126454	66
380-389[2]	40500	11	166954	55
370-379	36498	10	203452	45
360-369	32263	9	235715	37
350-359	34885	9	270600	27
340-349	29126	8	299726	20
330-339	23128	6	322854	13
320-329	16789	5	339643	9
310-319	14345	4	353988	5
300-309	8379	2	362367	3
290-299	5609	2	367976	1
280-289	2813	1	370789	1
275-279	1918	1	372707	0

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

Table 8.F.4: Frequency Distributions, Mathematics for ESEA—March 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	25088	7	25088	93
440-449	15389	4	40477	89
430-439	7868	2	48345	87
422-429[1]	15392	4	63737	83
410-421	24046	6	87783	76
400-409	24184	7	111967	70
390-399	24349	7	136316	63
380-389[2]	38479	10	174795	53
370-379	28657	8	203452	45
360-369	39875	11	243327	35
350-359	31088	8	274415	26
340-349	33834	9	308249	17
330-339	23294	6	331543	11
320-329	18599	5	350142	6
310-319	13416	4	363558	2
300-309	5714	2	369272	1
290-299	1834	0	371106	0
280-289	324	0	371430	0
275-279	274	0	371704	0

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

Table 8.F.5: Demographic Summary for All Examinees, ELA—March 2009

	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	372707	270600	73	102107	27	374	73	73	77	66	72	2.5
Grade												
Tenth	320483	255024	80	65459	20	379	76	76	80	69	75	2.5
Eleventh	29616	9498	32	20118	68	339	59	55	61	48	55	2.0
Twelfth	18628	4310	23	14318	77	333	56	52	56	45	52	2.0
Adult Education	3979	1767	44	2212	56	350	63	62	67	55	60	2.1
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	193283	132808	69	60475	31	369	73	72	75	64	70	2.4
Female	179357	137766	77	41591	23	379	74	74	80	69	75	2.6
Unknown	67	26	39	41	61	340	60	55	60	50	52	2.0
Race/Ethnicity												
American Indian or Alaska Native	2678	1958	73	720	27	372	74	73	77	65	70	2.4
Asian	33088	27503	83	5585	17	392	78	80	83	77	80	2.7
Pacific Islander	2440	1746	72	694	28	371	71	71	76	65	71	2.5
Filipino	9929	8378	84	1551	16	386	76	78	82	74	78	2.7
Hispanic or Latino	195559	125635	64	69924	36	362	69	68	73	61	68	2.3
African American	29699	18166	61	11533	39	360	69	67	73	58	64	2.3
White (not of Hispanic origin)	96584	85249	88	11335	12	393	81	82	85	75	81	2.7
Unknown	2730	1965	72	765	28	374	74	73	77	66	71	2.4
Language Fluency												
English-Only Students	190220	152787	80	37433	20	382	77	77	81	70	76	2.6
Initially Fluent English Proficient (IFEP)	30322	26570	88	3752	12	390	80	80	84	74	80	2.7
Reclassified Fluent English Proficient (RFEP)	67972	61249	90	6723	10	383	78	78	83	73	79	2.6
English-Learner Students	80030	28065	35	51965	65	340	59	56	61	50	57	2.0
Unknown	4163	1929	46	2234	54	351	63	62	68	56	61	2.1
Economically Disadvantaged												
No	131977	114768	87	17209	13	391	80	81	84	75	80	2.7
Yes	199408	125159	63	74249	37	361	69	67	72	61	67	2.3
Unknown	41322	30673	74	10649	26	377	75	74	79	68	73	2.5
Special Education Program Participation												
Students Receiving Services	37217	11460	31	25757	69	335	58	53	59	46	52	2.0
Students Not Receiving Services	335490	259140	77	76350	23	378	75	75	79	69	75	2.5

[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 2009

	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	371704	274415	74	97289	26	378	71	71	70	65	58
Grade											
Tenth	318571	256746	81	61825	19	384	74	75	74	69	62
Eleventh	29817	10515	35	19302	65	343	54	53	50	47	37
Twelfth	19138	5547	29	13591	71	340	53	51	46	45	35
Adult Education	4176	1606	38	2570	62	347	57	56	51	52	35
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	187704	138523	74	49181	26	379	72	72	70	66	58
Female	183915	135852	74	48063	26	377	70	71	70	65	59
Unknown	85	40	47	45	53	349	56	56	54	52	40
Race/Ethnicity											
American Indian or Alaska Native	2704	1888	70	816	30	373	69	70	67	62	54
Asian	30970	28778	93	2192	7	414	84	86	86	83	80
Pacific Islander	2457	1841	75	616	25	377	70	71	70	64	59
Filipino	9777	8607	88	1170	12	395	78	79	79	74	69
Hispanic or Latino	194943	129485	66	65458	34	367	66	66	65	60	52
African American	31437	17470	56	13967	44	359	62	62	59	54	47
White (not of Hispanic origin)	96617	84421	87	12196	13	395	80	79	79	74	66
Unknown	2799	1925	69	874	31	374	69	70	67	64	55
Language Fluency											
English-Only Students	193585	150492	78	43093	22	383	74	74	73	67	60
Initially Fluent English Proficient (IFEP)	30829	26326	85	4503	15	394	78	78	78	73	67
Reclassified Fluent English Proficient (RFEP)	69923	61291	88	8632	12	389	77	77	77	71	64
English-Learner Students	73001	34475	47	38526	53	351	56	57	55	52	44
Unknown	4366	1831	42	2535	58	349	58	57	52	53	37
Economically Disadvantaged											
No	132845	114137	86	18708	14	394	79	79	78	73	67
Yes	197032	129861	66	67171	34	368	66	66	65	60	52
Unknown	41827	30417	73	11410	27	380	72	72	70	66	58
Special Education Program Participation											
Students Receiving Services	33165	10910	33	22255	67	341	51	51	48	45	37
Students Not Receiving Services	338539	263505	78	75034	22	382	73	73	73	67	60

[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 2009

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		372707	205753	55	83006	22	83948	23	166954	45
Grade	Tenth	320483	156883	49	80747	25	82853	26	163600	51
	Eleventh	29616	27641	93	1374	5	601	2	1975	7
	Twelfth	18628	17902	96	482	3	244	1	726	4
	Adult Education	3979	3326	84	403	10	250	6	653	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	193283	115661	60	41259	21	36363	19	77622	40
	Female	179357	90036	50	41739	23	47582	27	89321	50
	Unknown	67	56	84	8	12	3	4	11	16
Race/Ethnicity	American Indian or Alaskan Native	2678	1486	55	642	24	550	21	1192	45
	Asian	33088	11840	36	7164	22	14084	43	21248	64
	Pacific Islander	2440	1422	58	583	24	435	18	1018	42
	Filipino	9929	4008	40	2714	27	3207	32	5921	60
	Hispanic or Latino	195559	133744	68	38654	20	23161	12	61815	32
	African American	29699	20670	70	5547	19	3482	12	9029	30
	White (not of Hispanic origin)	96584	31147	32	27093	28	38344	40	65437	68
	Unknown	2730	1436	53	609	22	685	25	1294	47
Language	English-Only Students	190220	84537	44	48287	25	57396	30	105683	56
	Initially Fluent English Proficient	30322	11490	38	8261	27	10571	35	18832	62
	Reclassified Fluent English	67972	31276	46	21829	32	14867	22	36696	54
	English-Learner Students	80030	75037	94	4167	5	826	1	4993	6
	Unknown	4163	3413	82	462	11	288	7	750	18
Economically Disadvantaged	No	131977	46298	35	35491	27	50188	38	85679	65
	Yes	199408	138745	70	38142	19	22521	11	60663	30
	Unknown	41322	20710	50	9373	23	11239	27	20612	50
Special Education Program	Receiving Services	37217	33894	91	2398	6	925	2	3323	9
	Not Receiving Services	335490	171859	51	80608	24	83023	25	163631	49

[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 2009

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		371704	196909	53	111058	30	63737	17	174795	47
Grade	Tenth	318571	146722	46	108672	34	63177	20	171849	54
	Eleventh	29817	28056	94	1433	5	328	1	1761	6
	Twelfth	19138	18352	96	638	3	148	1	786	4
	Adult Education	4176	3777	90	315	8	84	2	399	10
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	187704	96697	52	56052	30	34955	19	91007	48
	Female	183915	100141	54	54993	30	28781	16	83774	46
	Unknown	85	71	84	13	15	1	1	14	16
Race/Ethnicity	American Indian or Alaskan Native	2704	1593	59	778	29	333	12	1111	41
	Asian	30970	5742	19	9315	30	15913	51	25228	81
	Pacific Islander	2457	1345	55	773	31	339	14	1112	45
	Filipino	9777	3270	33	3881	40	2626	27	6507	67
	Hispanic or Latino	194943	127466	65	51514	26	15963	8	67477	35
	African American	31437	23311	74	6378	20	1748	6	8126	26
	White (not of Hispanic origin)	96617	32613	34	37637	39	26367	27	64004	66
	Unknown	2799	1569	56	782	28	448	16	1230	44
Language	English-Only Students	193585	91116	47	63735	33	38734	20	102469	53
	Initially Fluent English Proficient	30829	11647	38	10229	33	8953	29	19182	62
	Reclassified Fluent English	69923	29108	42	26959	39	13856	20	40815	58
	English-Learner Students	73001	61200	84	9719	13	2082	3	11801	16
	Unknown	4366	3838	88	416	10	112	3	528	12
Economically Disadvantaged	No	132845	47496	36	47845	36	37504	28	85349	64
	Yes	197032	127952	65	51090	26	17990	9	69080	35
	Unknown	41827	21461	51	12123	29	8243	20	20366	49
Special Program	Receiving Services	33165	29432	89	2999	9	734	2	3733	11
	Not Receiving Services	338539	167477	49	108059	32	63003	19	171062	51

[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 2009

	Percentiles[1]							Mean Scale Score	SD[2]	N Tested[3]
	1	5	25	50	75	95	99			
Total Examinees	287	308	346	373	402	443	450	374	39	372707
Grade[4]										
Tenth	291	315	355	380	406	443	450	379	37	320483
Eleventh	277	295	321	338	355	386	419	339	27	29616
Twelfth	275	291	317	333	348	375	410	333	26	18628
Adult Education	289	306	329	346	368	410	436	350	31	3979
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	284	304	342	370	395	436	450	369	39	193283
Female	295	315	352	378	406	443	450	379	38	179357
Unknown	275	287	308	338	366	399	436	340	38	67
Race/Ethnicity										
American Indian or Alaska Native	284	302	346	375	399	436	450	372	38	2678
Asian	293	319	363	395	424	450	450	392	41	33088
Pacific Islander	291	312	346	370	395	430	450	371	36	2440
Filipino	302	325	363	389	410	443	450	386	35	9929
Hispanic or Latino	285	304	338	361	386	419	443	362	35	195559
African American	279	301	334	359	386	419	450	360	36	29699
White (not of Hispanic origin)	295	327	370	395	419	450	450	393	36	96584
Unknown	275	299	344	378	406	443	450	374	42	2730
Language Fluency										
English-Only Students	289	314	357	383	410	443	450	382	38	190220
Initially Fluent English Proficient (IFEP)	301	331	366	389	414	450	450	390	35	30322
Reclassified Fluent English Proficient (RFEP)	315	340	366	383	402	436	450	383	28	67972
English-Learner Students	279	297	323	340	357	383	406	340	26	80030
Unknown	281	304	329	346	368	410	436	351	32	4163
Economically Disadvantaged										
No	299	327	368	392	419	450	450	391	36	131977
Yes	285	304	338	361	386	419	450	361	35	199408
Unknown	284	308	348	378	406	443	450	377	40	41322
Special Education Program Participation										
Students Receiving Services	275	287	312	334	355	389	424	335	32	37217
Students Not Receiving Services	293	317	352	378	402	443	450	378	37	335490

[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 10 students can only take 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 2009

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	303	319	349	376	407	450	450	378	39	371704
Grade[4]										
Tenth	306	323	356	382	410	450	450	384	38	318571
Eleventh	295	308	328	341	356	382	423	343	24	29817
Twelfth	293	306	327	340	352	376	414	340	22	19138
Adult Education	301	314	330	343	358	395	440	347	25	4176
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	301	317	349	378	410	450	450	379	41	187704
Female	306	321	349	374	404	450	450	377	38	183915
Unknown	275	293	327	349	367	414	448	349	34	85
Race/Ethnicity										
American Indian or Alaska Native	301	314	345	369	397	448	450	373	38	2704
Asian	317	341	389	423	450	450	450	414	37	30970
Pacific Islander	303	321	349	374	400	448	450	377	37	2457
Filipino	312	332	369	395	423	450	450	395	36	9777
Hispanic or Latino	301	314	341	364	389	433	450	367	35	194943
African American	295	310	334	354	380	423	450	359	34	31437
White (not of Hispanic origin)	308	330	369	395	423	450	450	395	37	96617
Unknown	298	312	341	371	404	450	450	374	41	2799
Language Fluency										
English-Only Students	303	319	352	382	414	450	450	383	40	193585
Initially Fluent English Proficient (IFEP)	310	330	364	392	428	450	450	394	39	30829
Reclassified Fluent English Proficient (RFEP)	319	336	364	387	414	450	450	389	34	69923
English-Learner Students	298	310	330	347	367	407	448	351	30	73001
Unknown	301	314	330	345	362	404	448	349	27	4366
Economically Disadvantaged										
No	310	328	366	395	428	450	450	394	38	132845
Yes	301	314	341	364	389	440	450	368	36	197032
Unknown	301	317	347	378	410	450	450	380	41	41827
Special Education Program Participation										
Students Receiving Services	290	301	319	336	358	400	440	341	31	33165
Students Not Receiving Services	308	323	352	380	410	450	450	382	38	338539

[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 10 students can only take CAHSEE one time in the spring during the February, March or May administration.

Appendix 8.G: Frequency Distributions and Demographic Summaries—May 2009

Table 8.G.1: Frequency Distributions, ELA—May 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	254	1	254	99
440-449	117	0	371	99
430-439	306	1	677	98
420-429	424	1	1101	97
410-419	545	1	1646	96
400-409	911	2	2557	94
390-399	1014	2	3571	92
380-389	1240	3	4811	89
370-379	1957	5	6768	84
360-369	3503	8	10271	76
350-359[1]	3778	9	14049	67
340-349	5721	13	19770	54
330-339	6965	16	26735	38
320-329	4889	11	31624	26
310-319	3985	9	35609	17
300-309	3008	7	38617	10
290-299	2169	5	40786	5
280-289	1161	3	41947	2
275-279	834	2	42781	0

[1]Passing Score = 350

Table 8.G.2: Frequency Distributions, Mathematics—May 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	292	1	292	99
440-449	142	0	434	99
430-439	308	1	742	98
420-429	294	1	1036	98
410-419	392	1	1428	97
400-409	594	1	2022	95
390-399	975	2	2997	93
380-389	1096	3	4093	90
370-379	1745	4	5838	86
360-369	2488	6	8326	80
350-359[1]	4907	12	13233	69
340-349	7358	17	20591	52
330-339	7070	17	27661	35
320-329	6705	16	34366	19
310-319	5129	12	39495	7
300-309	2371	6	41866	2
290-299	425	1	42291	1
280-289	127	0	42418	0
275-279	128	0	42546	0

[1] Passing Score = 350

Table 8.G.3: Frequency Distributions, ELA for ESEA—May 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	254	1	254	99
440-449	117	0	371	99
430-439	306	1	677	98
420-429	424	1	1101	97
410-419	545	1	1646	96
403-409[1]	608	1	2254	95
390-402	1317	3	3571	92
380-389[2]	1240	3	4811	89
370-379	1957	5	6768	84
360-369	3503	8	10271	76
350-359	3778	9	14049	67
340-349	5721	13	19770	54
330-339	6965	16	26735	38
320-329	4889	11	31624	26
310-319	3985	9	35609	17
300-309	3008	7	38617	10
290-299	2169	5	40786	5
280-289	1161	3	41947	2
275-279	834	2	42781	0

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

Table 8.G.4: Frequency Distributions, Mathematics for ESEA—May 2009

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	292	1	292	99
440-449	142	0	434	99
430-439	308	1	742	98
422-429[1]	157	0	899	98
410-421	529	1	1428	97
400-409	594	1	2022	95
390-399	975	2	2997	93
380-389[2]	1096	3	4093	90
370-379	1745	4	5838	86
360-369	2488	6	8326	80
350-359	4907	12	13233	69
340-349	7358	17	20591	52
330-339	7070	17	27661	35
320-329	6705	16	34366	19
310-319	5129	12	39495	7
300-309	2371	6	41866	2
290-299	425	1	42291	1
280-289	127	0	42418	0
275-279	128	0	42546	0

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

Table 8.G.5: Demographic Summary for All Examinees, ELA—May 2009

	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	42781	14049	33	28732	67	340	67	57	60	46	56	2.1
Grade												
Tenth	10198	5738	56	4460	44	357	76	65	68	56	64	2.3
Eleventh	13493	4096	30	9397	70	337	66	55	59	44	55	2.1
Twelfth	15926	2904	18	13022	82	330	61	51	54	40	51	2.0
Adult Education	3164	1311	41	1853	59	348	72	64	64	50	60	2.2
Unknown	-	-	-	-	-	-	-	-	-	-	-	-
Gender												
Male	24495	7453	30	17042	70	337	68	55	58	44	54	2.1
Female	18250	6586	36	11664	64	344	66	59	62	49	59	2.2
Unknown	36	10	28	26	72	332	66	52	57	43	48	1.8
Race/Ethnicity												
American Indian or Alaska Native	436	177	41	259	59	344	71	59	62	48	58	2.1
Asian	2251	623	28	1628	72	338	59	55	57	45	58	2.1
Pacific Islander	340	107	31	233	69	340	65	55	60	44	57	2.2
Filipino	651	238	37	413	63	347	67	59	62	49	61	2.3
Hispanic or Latino	26340	7220	27	19120	73	336	65	54	58	44	54	2.1
African American	4755	1440	30	3315	70	336	67	55	58	43	53	2.1
White (not of Hispanic origin)	7318	3967	54	3351	46	358	76	65	68	56	64	2.3
Unknown	690	277	40	413	60	345	70	59	62	50	57	2.1
Language Fluency												
English-Only Students	18057	7664	42	10393	58	347	72	60	63	49	59	2.2
Initially Fluent English Proficient (IFEP)	1661	750	45	911	55	349	73	61	64	51	61	2.3
Reclassified Fluent English Proficient (RFEP)	2213	1256	57	957	43	355	77	64	68	55	65	2.3
English-Learner Students	17959	3179	18	14780	82	329	59	51	54	41	51	2.0
Unknown	2891	1200	42	1691	58	348	71	63	64	50	60	2.2
Economically Disadvantaged												
No	9157	4458	49	4699	51	354	74	63	66	54	62	2.3
Yes	26814	6980	26	19834	74	335	64	53	57	43	53	2.1
Unknown	6810	2611	38	4199	62	344	69	60	62	48	58	2.1
Special Education Program Participation												
Students Receiving Services	9339	1660	18	7679	82	326	59	49	51	38	47	1.9
Students Not Receiving Services	33442	12389	37	21053	63	344	69	59	62	48	58	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 2009

	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	42546	13233	31	29313	69	342	54	48	50	47	38
Grade											
Tenth	10045	5269	52	4776	48	357	62	58	58	55	47
Eleventh	13543	3906	29	9637	71	340	52	46	48	46	38
Twelfth	15337	2869	19	12468	81	335	50	42	45	43	34
Adult Education	3621	1189	33	2432	67	343	55	47	52	50	37
Unknown	-	-	-	-	-	-	-	-	-	-	-
Gender											
Male	21426	6842	32	14584	68	342	53	49	48	47	37
Female	21066	6376	30	14690	70	343	55	46	51	47	39
Unknown	54	15	28	39	72	338	49	45	48	46	36
Race/Ethnicity											
American Indian or Alaska Native	486	146	30	340	70	341	53	48	48	46	36
Asian	1204	552	46	652	54	357	58	56	57	56	49
Pacific Islander	323	92	28	231	72	341	52	48	48	46	37
Filipino	614	248	40	366	60	351	57	52	56	53	44
Hispanic or Latino	25342	6727	27	18615	73	339	52	45	48	45	36
African American	5932	1336	23	4596	77	336	50	43	46	42	35
White (not of Hispanic origin)	7856	3839	49	4017	51	356	63	57	57	56	45
Unknown	789	293	37	496	63	345	56	50	51	49	39
Language Fluency											
English-Only Students	20938	7462	36	13476	64	346	56	50	51	49	39
Initially Fluent English Proficient (IFEP)	1974	720	36	1254	64	347	58	50	53	50	40
Reclassified Fluent English Proficient (RFEP)	2916	1270	44	1646	56	351	61	52	56	52	43
English-Learner Students	13442	2689	20	10753	80	335	48	42	45	43	36
Unknown	3276	1092	33	2184	67	343	55	47	52	50	37
Economically Disadvantaged											
No	10007	4349	43	5658	57	352	60	54	55	53	43
Yes	25122	6393	25	18729	75	338	51	45	47	45	36
Unknown	7417	2491	34	4926	66	344	55	48	51	49	38
Special Education Program Participation											
Students Receiving Services	7856	1284	16	6572	84	331	46	42	41	40	32
Students Not Receiving Services	34690	11949	34	22741	66	345	56	49	52	49	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 2009

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		42781	37970	89	2557	6	2254	5	4811	11
Grade	Tenth	10198	7170	70	1474	14	1554	15	3028	30
	Eleventh	13493	12624	94	529	4	340	3	869	6
	Twelfth	15926	15504	97	267	2	155	1	422	3
	Adult Education	3164	2672	84	287	9	205	6	492	16
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	24495	22106	90	1323	5	1066	4	2389	10
	Female	18250	15831	87	1233	7	1186	6	2419	13
	Unknown	36	33	92	1	3	2	6	3	8
Race/Ethnicity	American Indian or Alaskan Native	436	371	85	42	10	23	5	65	15
	Asian	2251	2009	89	92	4	150	7	242	11
	Pacific Islander	340	313	92	16	5	11	3	27	8
	Filipino	651	544	84	42	6	65	10	107	16
	Hispanic or Latino	26340	24667	94	1122	4	551	2	1673	6
	African American	4755	4370	92	233	5	152	3	385	8
	White (not of Hispanic origin)	7318	5130	70	943	13	1245	17	2188	30
	Unknown	690	566	82	67	10	57	8	124	18
Language	English-Only Students	18057	14724	82	1630	9	1703	9	3333	18
	Initially Fluent English Proficient	1661	1366	82	150	9	145	9	295	18
	Reclassified Fluent English	2213	1778	80	279	13	156	7	435	20
	English-Learner Students	17959	17672	98	228	1	59	0	287	2
	Unknown	2891	2430	84	270	9	191	7	461	16
Economically Disadvantaged	No	9157	6924	76	989	11	1244	14	2233	24
	Yes	26814	25285	94	988	4	541	2	1529	6
	Unknown	6810	5761	85	580	9	469	7	1049	15
Special Education Program	Receiving Services	9339	9020	97	215	2	104	1	319	3
	Not Receiving Services	33442	28950	87	2342	7	2150	6	4492	13

[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 2009

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		42546	38453	90	3194	8	899	2	4093	10
Grade	Tenth	10045	7345	73	2017	20	683	7	2700	27
	Eleventh	13543	12785	94	626	5	132	1	758	6
	Twelfth	15337	14973	98	314	2	50	0	364	2
	Adult Education	3621	3350	93	237	7	34	1	271	7
	Unknown	-	-	-	-	-	-	-	-	-
Gender	Male	21426	19159	89	1740	8	527	2	2267	11
	Female	21066	19242	91	1452	7	372	2	1824	9
	Unknown	54	52	96	2	4	-	-	2	4
Race/Ethnicity	American Indian or	486	446	92	33	7	7	1	40	8
	Asian	1204	916	76	171	14	117	10	288	24
	Pacific Islander	323	299	93	20	6	4	1	24	7
	Filipino	614	507	83	76	12	31	5	107	17
	Hispanic or Latino	25342	23945	94	1203	5	194	1	1397	6
	African American	5932	5665	95	239	4	28	0	267	5
	White (not of Hispanic	7856	5987	76	1371	17	498	6	1869	24
	Unknown	789	688	87	81	10	20	3	101	13
Language	English-Only Students	20938	18115	87	2147	10	676	3	2823	13
	Initially Fluent English	1974	1732	88	182	9	60	3	242	12
	Reclassified Fluent English	2916	2507	86	329	11	80	3	409	14
	English-Learner Students	13442	13085	97	312	2	45	0	357	3
	Unknown	3276	3014	92	224	7	38	1	262	8
Economically Disadvantaged	No	10007	8070	81	1390	14	547	5	1937	19
	Yes	25122	23760	95	1182	5	180	1	1362	5
	Unknown	7417	6623	89	622	8	172	2	794	11
Special Education Program	Receiving Services	7856	7569	96	252	3	35	0	287	4
	Not Receiving Services	34690	30884	89	2942	8	864	2	3806	11

[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 2009

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	275	291	318	337	358	404	439	340	33	42781
Grade[4]										
Tenth	275	289	326	358	387	427	450	357	42	10198
Eleventh	275	291	318	337	353	384	422	337	29	13493
Twelfth	275	289	314	331	345	369	400	330	25	15926
Adult Education	285	303	328	343	364	408	439	348	31	3164
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	275	287	314	335	355	397	433	337	33	24495
Female	279	297	322	341	360	408	446	344	33	18250
Unknown	277	281	311	326	351	408	422	332	33	36
Race/Ethnicity										
American Indian or Alaska Native	275	289	320	342	364	404	433	344	34	436
Asian	275	289	314	333	353	412	450	338	35	2251
Pacific Islander	289	293	322	337	355	389	422	340	28	340
Filipino	277	299	322	339	364	427	450	347	37	651
Hispanic or Latino	275	291	316	335	351	384	417	336	28	26340
African American	275	287	314	335	355	390	422	336	31	4755
White (not of Hispanic origin)	275	291	328	353	387	433	450	358	42	7318
Unknown	275	287	318	339	369	417	450	345	38	690
Language Fluency										
English-Only Students	275	289	320	343	369	417	450	347	38	18057
Initially Fluent English Proficient (IFEP)	275	295	326	347	369	417	450	349	35	1661
Reclassified Fluent English Proficient (RFEP)	281	305	335	355	374	408	439	355	31	2213
English-Learner Students	275	289	314	330	345	366	384	329	23	17959
Unknown	285	301	326	343	366	408	439	348	32	2891
Economically Disadvantaged										
No	275	293	326	349	379	427	450	354	40	9157
Yes	275	291	316	333	351	382	412	335	28	26814
Unknown	275	291	320	341	364	412	446	344	35	6810
Special Education Program Participation										
Students Receiving Services	275	283	305	326	343	374	404	326	28	9339
Students Not Receiving Services	277	295	322	341	362	408	446	344	33	33442

[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 10 students can only take 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 2009

	Percentiles[1]									
	1	5	25	50	75	95	99	Mean Scale Score	SD[2]	N Tested[3]
Total Examinees	294	307	323	338	354	398	445	342	28	42546
Grade[4]										
Tenth	291	307	327	352	383	430	450	357	38	10045
Eleventh	294	309	325	338	352	383	420	340	24	13543
Twelfth	294	307	321	334	345	366	398	335	20	15337
Adult Education	300	311	327	340	355	388	420	343	24	3621
Unknown	-	-	-	-	-	-	-	-	-	-
Gender										
Male	291	304	321	338	355	401	445	342	30	21426
Female	297	309	325	338	354	395	437	343	26	21066
Unknown	291	304	323	336	350	378	420	338	24	54
Race/Ethnicity										
American Indian or Alaska Native	294	307	323	338	354	393	425	341	27	486
Asian	291	309	329	345	376	445	450	357	40	1204
Pacific Islander	297	311	323	338	354	390	425	341	26	323
Filipino	302	311	329	343	366	425	450	351	33	614
Hispanic or Latino	294	307	323	336	350	381	415	339	23	25342
African American	288	304	321	333	347	376	408	336	23	5932
White (not of Hispanic origin)	297	309	331	348	378	430	450	356	36	7856
Unknown	275	302	325	341	359	404	445	345	31	789
Language Fluency										
English-Only Students	291	307	325	340	359	411	450	346	31	20938
Initially Fluent English Proficient (IFEP)	297	309	327	341	361	408	450	347	30	1974
Reclassified Fluent English Proficient (RFEP)	302	315	334	347	362	408	445	351	27	2916
English-Learner Students	294	307	321	333	345	370	398	335	20	13442
Unknown	300	311	327	340	355	390	425	343	25	3276
Economically Disadvantaged										
No	297	309	329	345	368	425	450	352	34	10007
Yes	294	307	323	336	350	381	415	338	24	25122
Unknown	291	307	325	340	357	401	445	344	29	7417
Special Education Program Participation										
Students Receiving Services	285	300	315	327	341	372	408	331	23	7856
Students Not Receiving Services	297	309	327	340	357	401	445	345	28	34690

[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 10 students can only take 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 an Office of Testing Integrity (OTI) that resides in the ETS legal department. The OTI provides quality assurance services for all testing programs administered by ETS. The Quality Assurance division of ETS also publishes and maintains the *ETS Standards for Quality and Fairness,* which 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.

In addition, each department at ETS that is involved in the testing cycle designs and implements an independent set of procedures to ensure the quality of their products. 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 ETS content staff and approved by the CDE. 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's written guidelines for clarity, style, accuracy, and appropriateness and with approved item specifications. The artwork and graphics for the items are created during the internal content review period so assessment specialists can evaluate the correctness and appropriateness of the art early in the item development process. ETS selects visual stimuli that are relevant to the item content and that are easily understood.

Editorial Review

Another step in the ETS internal review process involves a team of specially trained editors who check questions for clarity, correctness and grade-level appropriateness of language, adherence to style guidelines, and conformity to item-writing best practices. The editorial review also includes cycles of copyediting and proofreading.

Bias and Sensitivity Review

One of the final steps in the ETS internal review process is to have all items and stimuli reviewed for bias and sensitivity. Only ETS 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 will read 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 a single achievement test to be used statewide for the testing of students in California public schools. The SPAR panel representatives ensure that the test items conform to the requirements of *Education Code* 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 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.

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, at ETS. 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 secure file transfer protocol (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 offsite, to prevent loss from a system breakdown or a natural disaster.
- The offsite 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 retuned 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 notifies school districts that do not return their materials in a timely manner. The CAHSEE Support Center contacts the district CAHSEE coordinators 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 as part of the development of the scannable form and then 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.” Early return answer documents, numbering 300 to 500, are scanned and a quality control (QC) 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 items on the diagnostic sheets, the scanner is recalibrated before being allowed to continue processing student documents.

Documents received in poor condition (torn, folded, or water-stained) that could not be fed through the high-speed scanners are either scanned using a flat-bed 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, and

- After student results are posted by district CAHSEE 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 Online Scoring Network. 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 scan 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. A photocopy is made of each identified answer document 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 multiple–choice 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, after which 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. CDE also performs a replication of the equating results. 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

The Human Resources Research Organization (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

For the quality control of various CAHSEE student and summary reports, four general areas are evaluated, including the following:

- Comparing report formats to input sources from the CDE-approved samples
- Validating and verifying the report data by querying the appropriate student data
- Evaluating the production print execution performance by comparing the number of report copies, sequence of report order, and offset characteristics to the CDE's requirements
- Proofreading of reports at the CDE, ETS, and Pearson prior to any school district mailings

All reports are required to include a single, accurate CDS code, a charter school number (if applicable), a school district name, and a school name. All elements conform to the CDE's official CDS code and naming records. From the start of processing through scoring and reporting, the CDS Master File is used to verify and confirm accurate codes and names. The CDE Master File is provided by the CDE throughout the year as updates are available.

For students who use more than one answer document, the matching process, as described earlier in this chapter, provides for the creation of individual student records from which reports are created.

After the reports are validated against CDE requirements, a set of reports for pilot districts are provided to the CDE and ETS for review and approval. Pearson sends paper reports on the actual report forms, folded as they are expected to look in production. The CDE and ETS sign off on the report package after a thorough review.

Once the CDE grants approval of the reports generated from the pilot test, Pearson proceeds with the first production batch test. The first production batch is selected to validate a subset of districts that contains examples of key reporting characteristics representative of the state. The first production batch test incorporates client-selected school districts and provides the last check prior to generating all reports and mailing them to the districts.

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—2006–07, 2007–08, and 2008–09. The indicators of examinee performance include the mean and standard deviation of scale scores, the percentage of examinees classified into 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 English-Language Arts (ELA) and mathematics tests are the number of examinees assessed and the means and standard deviations of examinees' scale scores in 2006–07, 2007–08, and 2008–09.

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 State Board of Education 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 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[11, 12] 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 standard errors of measurement expressed in raw score units appear in Table10.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.

[11] These statistics are based on the equating samples.
[12] 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 2006–07, 2007–08, and 2008–09

Subject	Admin	Number of Students (with valid scores)			Scale Score Mean and Standard Deviation (SD)					
					2006–07		2007--08		2008--09	
		2006–07	2007--08	2008--09	Mean	SD	Mean	SD	Mean	SD
English–Language Arts	July	9,481	12,959	13,373	331	27	338	25	337	26
	October	46,931	44,351	43,955	336	32	342	31	342	29
	November	109,115	116,439	108,722	340	30	342	30	343	33
	December	3,894	3,631	3,544	338	29	341	28	338	29
	February	214,778	195,737	195,091	368	41	371	41	371	41
	March	359,058	374,545	372,707	370	38	372	39	374	39
	May	43,182	47,136	42,781	336	33	341	33	340	33
Mathematics	July	9,795	12,413	13,237	338	21	342	22	340	22
	October	48,586	44,121	44,958	341	26	341	26	343	26
	November	113,018	117,242	109,445	342	26	345	27	345	27
	December	3,815	3,832	3,544	341	24	342	23	345	22
	February	215,633	196,937	196,387	371	40	372	40	373	40
	March	361,140	374,364	371,704	374	41	377	40	378	39
	May	43,362	47,084	42,546	341	28	341	28	342	28

Table 10.A.2: Percentage of Students Passing Each Content Area across 2006–07, 2007–08, and 2008–09

Admin	Percent Passing					
	English–Language Arts			Mathematics		
	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
July	22	30	30	22	31	29
October	31	39	37	31	30	33
November	35	40	40	32	39	38
December	30	35	30	32	31	36
February	66	67	67	67	67	67
March	70	72	73	68	71	74
May	29	36	33	32	32	31

Table 10.A.3: Percentage of Below Proficient, Proficient, and Advanced across 2006–07, 2007–08, and 2008–09

Subject	Admin	Below Proficient			Proficient			Advanced		
		2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
English–Language Arts	July	95	95	95	3	3	3	2	2	2
	October	92	90	91	5	6	5	4	4	4
	November	92	90	88	5	6	7	3	3	5
	December	92	92	93	5	8	4	3	3	3
	February	60	59	58	19	17	17	21	24	25
	March	59	54	55	21	23	22	20	23	23
	May	90	88	89	6	7	6	4	5	5
Mathematics	July	96	96	96	3	4	4	1	1	1
	October	94	93	93	5	5	5	1	1	2
	November	93	92	91	5	8	7	2	2	2
	December	94	94	94	5	6	5	1	1	1
	February	61	60	59	26	27	26	14	14	15
	March	57	56	53	26	28	30	17	17	17
	May	91	91	90	7	7	8	2	2	2

Table 10.A.4: Table 10.A.4: Scale Score Distributions across 2006–07, 2007–08, and 2008–09 for ELA (July to December)

Observed Score Distribution	July			October			November			December		
	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
450	13	12	15	236	176	161	197	323	637	6	6	8
440 – 449	12	18	10	121	94	180	314	427	695	5	8	9
430 – 439	7	34	35	129	205	109	221	319	440	7	13	18
420 – 429	30	40	47	337	284	300	585	683	1,002	24	19	20
410 – 419	42	62	52	413	372	571	1,245	1,460	1,954	34	27	26
400 – 409	65	99	96	658	673	420	1,704	1,215	1,577	66	53	32
390 – 399	118	135	198	846	830	1,070	2,049	2,278	2,744	82	66	55
380 – 389	213	231	214	1,125	1,703	1,048	2,646	4,601	3,614	91	110	90
370 – 379	270	409	470	2,150	2,601	2,776	4,992	7,458	6,817	165	171	158
360 – 369	393	1,006	1,055	3,207	3,985	3,632	10,037	10,955	9,430	309	255	241
350 – 359	879	1,800	1,761	5,479	6,270	6,102	14,566	17,204	14,656	390	554	416
340 – 349	1,311	2,281	2,218	6,477	6,601	6,733	16,305	17,510	15,210	579	609	552
330 – 339	1,459	2,157	2,241	6,259	5707	6,096	14,756	14,625	13,541	541	582	559
320 – 329	1,649	1,912	1,769	5,241	4,576	5,684	14,035	11,140	10,920	609	446	482
310 – 319	1,052	1,114	1,344	5,038	3,424	3,447	8,658	10,016	8,496	371	326	336
300 – 309	925	832	1,074	3,387	3,197	2,866	7,964	6,389	6,813	284	175	247
290 – 299	474	458	451	2,679	1,844	1,660	4,522	5,640	5,682	216	119	164
280 – 289	305	225	169	1,831	1,031	699	2,572	2,170	2,583	67	55	78
275 – 279	264	134	154	1,318	778	401	1,747	2,026	1,911	48	37	53

Table 10.A.5: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for ELA (February to May)

Observed Score Distribution	February			March			May		
	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
450	5,751	7,536	7,532	6,710	9,728	13,087	151	265	254
440 – 449	6,112	3,238	3,505	8,363	4,810	5,998	88	120	117
430 – 439	3,851	7,414	8,335	5,689	13,298	15,260	260	340	306
420 – 429	9,002	8,765	9,677	14,278	16,886	9,057	386	408	424
410 – 419	9,910	9,805	10,095	26,247	19,469	29,937	455	499	545
400 – 409	15,926	15,317	15,349	18,992	31,946	21,199	560	899	911
390 – 399	16,213	14,917	14,517	38,429	32,953	31,916	970	1,101	1,014
380 – 389	20,000	14,107	13,347	27,445	42,264	40,500	1,502	1,833	1,240
370 – 379	18,132	17,368	15,879	34,020	28,261	36,498	1,843	2,326	1,957
360 – 369	16,923	15,512	17,951	37,356	40,840	32,263	2,416	3,448	3,503
350 – 359	19,561	18,120	13,654	32,592	27,846	34,885	4,098	5,779	3,778
340 – 349	17,363	16,883	16,214	28,133	28,792	29,126	5,079	6,873	5,721
330 – 339	14,634	13,774	14,670	23,242	22,273	23,128	6,568	6,303	6,965
320 – 329	14,081	10,599	11,753	21,427	17,030	16,789	5,026	5,083	4,889
310 – 319	9,080	9,130	10,438	13,640	15,163	14,345	4,321	3,930	3,985
300 – 309	8,621	5,296	5,830	10,376	9,672	8,379	4,428	3,527	3,008
290 – 299	4,993	3,756	3,444	7,638	6,936	5,609	2,613	2,068	2,169
280 – 289	2,770	2,370	1,673	2,845	3,816	2,813	1,406	1,309	1,161
275 – 279	1,855	1,830	1,228	1,636	2,562	1,918	1,012	1,025	834

Table 10.A.6: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for Mathematics (July to December)

Observed Score Distribution	July			October			November			December		
	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
450	13	21	28	212	195	231	797	731	680	10	12	14
440 – 449	8	22	14	80	101	166	295	275	319	6	7	5
430 – 439	25	9	22	184	196	95	268	643	681	9	14	13
420 – 429	16	35	31	244	230	240	670	720	807	16	15	8
410 – 419	43	38	46	282	306	391	1,066	792	1,285	19	18	17
400 – 409	40	78	81	404	415	454	1,226	1,282	1,392	27	39	42
390 – 399	71	130	121	688	722	477	1,382	2,169	2,111	53	34	41
380 – 389	131	191	216	961	834	1,133	2,201	2,913	2,832	80	85	61
370 – 379	245	443	443	1,989	1,734	1,503	4,517	6,037	5,416	99	117	129
360 – 369	513	1,130	834	3,905	3,202	4,284	7,905	10,830	9,208	358	324	379
350 – 359	1,074	1,791	1,978	6,314	5,121	5,973	16,329	19,619	16,540	532	507	550
340 – 349	2,114	2,751	2,746	9,548	8,233	9,158	21,618	18,488	16,689	772	832	831
330 – 339	2,302	2,171	2,366	7,614	7,293	7,384	18,415	19,986	20,525	642	693	675
320 – 329	1,600	1,748	1,999	7,761	7,686	6,136	15,841	14,009	14,690	531	555	458
310 – 319	1,012	1,244	1,419	4,125	4,871	4,544	11,937	11,030	10,617	389	397	242
300 – 309	403	458	614	2,768	2,122	1,921	6,639	5,441	4,208	186	140	62
290 – 299	133	112	225	1,130	544	589	1,263	1,786	965	67	25	11
280 – 289	34	15	32	216	218	159	390	293	282	13	12	5
275 – 279	18	26	22	161	98	120	259	198	198	6	6	1

Table 10.A.7: Observed Score Distributions of CAHSEE across 2006–07, 2007–08, and 2008–09 for Mathematics (February to May)

Observed Score Distribution	February			March			May		
	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
450	10,150	8,381	9,981	22,113	23,201	25,088	287	272	292
440 – 449	7,345	6,965	3,833	7,612	16,158	15,389	120	122	142
430 – 439	3,833	3,875	8,066	15,344	8,118	7,868	258	302	308
420 – 429	7,808	7,545	7,961	15,024	15,531	15,392	296	356	294
410 – 419	12,013	11,715	11,820	21,887	23,050	24,046	353	376	392
400 – 409	12,138	11,732	11,493	21,346	22,720	24,184	600	706	594
390 – 399	16,147	15,053	10,994	20,506	29,423	24,349	890	981	975
380 – 389	15,598	13,911	17,347	31,946	28,270	38,479	1,059	1,214	1,096
370 – 379	18,826	16,140	13,280	30,243	33,195	28,657	1,819	1,915	1,745
360 – 369	18,547	15,719	19,254	29,528	31,468	39,875	2,832	2,830	2,488
350 – 359	22,778	20,035	16,980	29,699	36,497	31,088	5,252	5,802	4,907
340 – 349	18,528	16,896	21,278	34,964	29,274	33,834	6,062	6,677	7,358
330 – 339	19,850	18,870	16,246	26,867	30,496	23,294	7,761	8,620	7,070
320 – 329	13,895	13,462	13,318	23,644	20,984	18,599	6,263	6,743	6,705
310 – 319	9,118	10,660	9,284	18,747	13,486	13,416	5,609	5,814	5,129
300 – 309	6,841	4,243	3,716	8,234	9,737	5,714	2,599	2,860	2,371
290 – 299	1,473	1,107	1,155	2,636	1,930	1,834	921	1,024	425
280 – 289	432	343	186	485	553	324	201	236	127
275 – 279	313	285	195	315	273	274	180	234	128

Appendix 10.B—Historical Comparisons on Test Characteristics

Table 10.B.1: Average Proportion Correct of Operational Test Items across 2006–07, 2007–08, and 2008–09

Subject	Admin	Average *p*-value		
		2006–07	2007–08	2008–09
English–Language Arts	July	0.54	0.53	0.53
	October	0.57	0.58	0.56
	November	0.56	0.59	0.59
	December	0.55	0.56	0.56
	February	0.69	0.72	0.71
	March	0.70	0.72	0.73
	May	0.54	0.56	0.56
Mathematics	July	0.47	0.46	0.46
	October	0.48	0.47	0.48
	November	0.49	0.49	0.48
	December	0.48	0.46	0.47
	February	0.63	0.63	0.64
	March	0.65	0.65	0.67
	May	0.47	0.47	0.47

Table 10.B.2: Average IRT b-values of Operational Test Items across 2006–07, 2007–08, and 2008–09

Subject	Admin	Average IRT b-value		
		2006–07	2007–08	2008–09
English–Language Arts	July	-0.25	0.02	-0.01
	October	-0.25	-0.12	-0.05
	November	-0.12	-0.16	-0.19
	December	-0.10	-0.06	-0.18
	February	-0.03	-0.17	-0.08
	March	-0.03	-0.10	-0.09
	May	-0.10	-0.08	-0.12
Mathematics	July	-0.32	-0.17	-0.23
	October	-0.26	-0.26	-0.23
	November	-0.29	-0.23	-0.19
	December	-0.28	-0.16	-0.11
	February	-0.19	-0.18	-0.22
	March	-0.21	-0.18	-0.22
	May	-0.24	-0.24	-0.23

Table 10.B.3 Average Point-Biserial Correlation of Operational Test Items across 2006–07, 2007–08, and 2008–09

Subject	Admin	Average Point-Biserial Correlation		
		2006–07	2007–08	2008–09
English–Language Arts	July	0.34	0.32	0.32
	October	0.38	0.37	0.35
	November	0.36	0.36	0.38
	December	0.35	0.33	0.35
	February	0.45	0.46	0.45
	March	0.43	0.43	0.42
	May	0.39	0.38	0.38
Mathematics	July	0.28	0.28	0.29
	October	0.32	0.32	0.32
	November	0.33	0.34	0.34
	December	0.30	0.30	0.28
	February	0.46	0.46	0.46
	March	0.47	0.47	0.46
	May	0.35	0.35	0.35

Table 10.B.4 Reliabilities and Standard Error of Measurement (SEM) of Operational Test Forms across 2006–07, 2007–08, and 2008–09

Subject	Admin	Reliability			SEM		
		2006–07	2007–08	2008–09	2006–07	2007–08	2008–09
English–Language Arts	July	0.88	0.87	0.87	4.67	4.48	4.40
	October	0.91	0.90	0.89	4.48	4.41	4.63
	November	0.90	0.90	0.87	4.52	4.43	4.40
	December	0.90	0.88	0.89	4.40	4.27	4.33
	February	0.94	0.94	0.94	4.11	4.08	3.99
	March	0.93	0.93	0.93	4.10	3.98	3.98
	May	0.92	0.91	0.91	4.43	4.42	4.38
Mathematics	July	0.85	0.86	0.87	4.12	4.13	4.12
	October	0.89	0.89	0.89	4.05	4.09	4.04
	November	0.90	0.90	0.87	4.09	4.07	4.12
	December	0.88	0.87	0.85	4.09	4.07	4.13
	February	0.95	0.95	0.95	3.73	3.75	3.71
	March	0.95	0.95	0.95	3.68	3.67	3.63
	May	0.91	0.91	0.91	4.05	4.07	4.08

References

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

California Department of Education. (2006). *Mathematics Framework for California Public Schools*. Sacramento, CA: California Department of Education.

California Department of Education. (2008a). *California High School Exit Examination District and Test Site Coordinator's Manual (July 2008–May 2009)*. Sacramento, CA: California Department of Education.

California Department of Education. (2008b). *California High School Exit Examination Directions for Administration (July 2008–May 2009).* Sacramento, CA; California Department of Education.

California Department of Education. (2009). *California Content Standards*. http://www.cde.ca.gov/be/st/ss/.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.). *Educational Measurement* (2nd ed.). Washington, DC: American Council on Education.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Drasgow, F. (1988). Polychoric and polyserial correlations. In Kotz L., Johnson N. L. (Eds.), *Encyclopedia of statistical sciences (Vol. 7).* New York: Wiley.

Educational Testing Service. (2002). *ETS Standards for Quality and Fairness*. Office of Testing Integrity, Princeton, NJ: Educational Testing Service.

Feldt, L.S., & Brennan, R. L. (1989). Reliability. in R.L. Linn (Ed.), *Educational Measurement* (3rd ed). Phoenix, AZ: Oryx Press.

Gaffney, T., & Perryman, C. (July, 2009). *A longitudinal look at the factor structure of educational achievement tests.* Paper presented at the International Meeting of the Psychometric Society. University of Cambridge: Cambridge, UK.

Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational Measurement* (4th ed.). Washington, DC: American Council on Education and National Council on Measurement in Education.

Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (1998, April). *The bookmark standard setting procedure: Methodology and recent implementations.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.

Lewis. C., & Thayer, D. T. (1996). Unpublished Seminar Notes.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal of Educational Measurement, 32*, 179-197.

Lord, F. M. (1980). *Application of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.) (pp.13–103). New York: Macmillan.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives.* Mahwah, NJ: Lawrence Erlbaum Associates.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *17*, 159-176.

Muraki, E. and Bock, R. D. (1995). *PARSCALE: Parameter scaling of rating data* (Version 2.2). Chicago, IL: Scientific Software, Inc.

Patrick, R. & Way, D. (March, 2008). *Field testing and equating designs for state educational assessments*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, *7*, 201-210.

Taylor, L., Johnstone, C., 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.

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., & 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.