



**California High School Exit Examination**  
**Technical Report**  
**February 2004 - May 2004 Administrations**  
**December 2004**  
**Educational Testing Service**



## Table of Contents

<b>Executive Summary .....</b>	<b>11</b>
<b>Chapter 1: Background.....</b>	<b>16</b>
TARGET POPULATION .....	16
TEST DESCRIPTION .....	16
SCORES FOR ANALYSIS AND REPORTING .....	17
WEIGHTING SCHEME FOR REPORTED SCORES.....	18
SUBSCORES.....	18
NEW CAHSEE PASSING SCORES .....	19
RAW SCORE TO SCALE SCORE CONVERSION.....	19
<b>Chapter 2: CAHSEE Test Development Procedures .....</b>	<b>20</b>
<b>Chapter 3: Equating Procedures.....</b>	<b>23</b>
ITEM CALIBRATION AND TRANSFORMATION.....	23
ESTABLISHING THE REPORTING SCALE FOR THE NEW CAHSEE: FEBRUARY 2004.....	24
TRUE-SCORE EQUATING MARCH AND MAY 2004.....	28
BRAILLE, LARGE PRINT, AND AUDIO CD FORM EQUATING .....	28
CONVERSION TABLES AND CONDITIONAL STANDARD ERRORS OF MEASUREMENT .....	29
APPENDIX 3.A: SCORING TABLES FOR OPERATIONAL & AUDIO CD CONVERSIONS—FEBRUARY 2004.....	30
APPENDIX 3.B: SCORING TABLES FOR OPERATIONAL AND AUDIO CD CONVERSIONS MARCH – 2004.....	32
APPENDIX 3.C: SCORING TABLES FOR OPERATIONAL AND AUDIO CD CONVERSIONS—MAY 2004.....	34
APPENDIX 3.D: SCORING TABLES FOR BRAILLE CONVERSIONS.....	36
APPENDIX 3.E: STANDARD ERRORS OF THETA BASED ON WEIGHTED RAW SCORES .....	38
<b>Chapter 4: Summary of Item-Level Analyses.....</b>	<b>39</b>
CLASSICAL ITEM ANALYSES .....	39
DIFFERENTIAL ITEM FUNCTIONING (DIF).....	40
IRT ANALYSES .....	43
SUMMARY OF ITEM-LEVEL STATISTICS .....	44
APPENDIX 4.A: CAHSEE FIELD-TEST ITEM REVIEW – DESCRIPTION AND EXAMPLES OF CLASSIFICATION CATEGORIES .....	46
APPENDIX 4.B: ITEM AND DIF STATISTICS FOR FEBRUARY 2004.....	48
APPENDIX 4.C: ITEM AND DIF STATISTICS FOR MARCH 2004.....	55
APPENDIX 4.D: ITEM AND DIF STATISTICS FOR MAY 2004.....	63
<b>Chapter 5: Summary of Test-Level Analyses .....</b>	<b>69</b>
DEMOGRAPHIC DISTRIBUTIONS .....	69
SCORE DISTRIBUTIONS AND SUMMARY STATISTICS .....	69
STRAND INTERCORRELATIONS, RELIABILITY AND STANDARD ERROR OF MEASUREMENT .....	69
CR ITEM AND RATER AGREEMENT SUMMARY.....	71
DECISION CLASSIFICATION ANALYSES .....	72
APPENDIX 5.A: DEMOGRAPHIC SUMMARIES, FREQUENCY DISTRIBUTIONS, INTERCORRELATIONS, AND DECISION CONSISTENCIES—FEBRUARY 2004 .....	73

APPENDIX 5.B: DEMOGRAPHIC SUMMARIES, FREQUENCY DISTRIBUTIONS, INTERCORRELATIONS, AND DECISION CONSISTENCIES—MARCH 2004 .....	86
APPENDIX 5.C: DEMOGRAPHIC SUMMARIES, FREQUENCY DISTRIBUTIONS, INTERCORRELATIONS, AND DECISION CONSISTENCIES—MAY 2004 .....	100
<b>Chapter 6: Special Analyses.....</b>	<b>114</b>
DISABILITY/ACCOMMODATIONS SUMMARY .....	114
GENERALIZABILITY ANALYSES.....	115
MARK DISCRIMINATION ANALYSES.....	116
Background.....	116
CAHSEE Answer Document Scanning.....	117
Purpose.....	117
Data Source.....	119
Data Analysis.....	119
APPENDIX 6.A: RESULTS OF ACCOMMODATIONS AND DISABILITY ANALYSES – FEBRUARY 2004 .....	120
APPENDIX 6.B: RESULTS OF ACCOMMODATIONS AND DISABILITY ANALYSES – MARCH 2004	132
Alg1.....	135
APPENDIX 6.C: RESULTS OF ACCOMMODATIONS AND DISABILITY ANALYSES - MAY 2004....	144
APPENDIX 6.D: MARK DISCRIMINATION ANALYSIS .....	154
<b>Chapter 7: Standard Setting.....</b>	<b>161</b>
OVERVIEW OF THE PROCESS .....	161
CHARACTERISTICS OF STANDARD SETTING PANELS .....	162
MATERIALS.....	162
Item Map.....	163
Ordered Item Booklet .....	163
Evaluation Forms .....	163
PROCEDURE .....	164
TRAINING.....	164
Table Leader Training.....	165
Practice.....	165
RATINGS .....	166
ANALYSIS .....	166
FEEDBACK AND DISCUSSION.....	166
RESULTS .....	167
ENGLISH LANGUAGE ARTS (ELA).....	168
Summary of ELA Panel Results .....	169
MATHEMATICS.....	177
Summary of Mathematics Panel Results .....	178
EVALUATION OF THE BOOKMARK PROCESS.....	185
Evaluation Results from the ELA panel .....	185
Evaluation Results from the Mathematics Panel .....	186
CONCLUSIONS.....	188
APPENDIX 7.A: CAHSEE BOOKMARK WORKSHOP AGENDA.....	189
APPENDIX 7.B: ITEM MAPS .....	191
APPENDIX 7.C: EVALUATION FORMS.....	195

APPENDIX 7.D: TECHNICAL NOTE ON PROCEDURES TO GENERATE PREDICTED CONSEQUENCES DATA FOR THE CAHSEE STANDARD SETTING.....	202
APPENDIX 7.E: TECHNICAL NOTE ON CONVERTING PANELIST BOOKMARKS TO PASSING SCORES ON THE CAHSEE.....	209
Definition of Terms.....	209
Overview of the Process .....	210
Creating the Ordered-Item Booklet .....	210
Providing Consequence Data.....	210
Calculating the Final Passing Score.....	211
Examples.....	211
REFERENCES .....	213

## Tables and Figures

Table E.1: Summary of Examinees Tested for Each Administration by Content .....	12
Table E.2: Summary of Passing Rates by Content Area and Test Administration.....	12
Table E.3: CAHSEE Summary Statistics—English Language Arts .....	13
Table E.4: CAHSEE Summary Statistics—Mathematics.....	14
Table 1.1: Items Included in the 2004 Administrations.....	17
Table 1.2: Mathematics and ELA Subscores.....	19
Table 2. 1: Difficulty (B) and Discrimination (Rbis) Specifications for ELA .....	21
Table 2. 2: Number of Items and Mean B-Values by Content Strand for ELA .....	21
Table 2. 3: Difficulty (B) and Discrimination (Rbis) Specifications for Mathematics .....	22
Table 2. 4: Number of Items and Mean B-Values by Content Strand for Mathematics.....	22
Table 3.1: Passing Scores on the February 2004 CAHSEE.....	24
Table 3.2: February 2004 CAHSEE ELA Reporting Scale.....	26
Table 3.3: February 2004 CAHSEE Mathematics Reporting Scale.....	27
Table 3.A.1: Operational and Audio CD Conversion – ELA February 2004.....	30
Table 3.A.2: Operational and Audio CD Conversion – Mathematics February 2004.....	31
Table 3.B.1: Operational and Audio CD Conversion – ELA March 2004.....	32
Table 3.B.2: Operational and Audio CD Conversion – Mathematics March 2004.....	33
Table 3.C.1: Operational and Audio CD Conversion – ELA May 2004.....	34
Table 3.C.2: Operational and Audio CD Conversion – Mathematics May 2004 .....	35
Table 3.D.1: Braille Conversion – ELA February, March, and May 2004 .....	36
Table 3.D.2: Braille Conversion – Mathematics February, March, and May 2004 .....	37
Table 4.1: Summary of Items and Forms by Administration .....	39
Table 4.2: DIF Categories.....	42
Table 4.3: Summary of IRT Model Data Fit Distribution and C-DIF Items Across Administrations: Operational Items.....	43
Table 4.4: Summary of IRT Model Data Fit Distribution and C-DIF Items Across Administrations: Field-Test Items .....	44
Table 4.5: Items Recommended for Future Forms .....	44
Table 4.6: Listing of Summary Tables .....	45
Table 4.B.1: Listing of Operational Items Exhibiting Significant Negative DIF February 2004	48
Table 4.B.2: Listing of Field-Test Items Exhibiting Significant Negative DIF .....	49
Table 4.B.3: Listing of CR Item Statistics - ELA.....	50
Table 4.B.4: Summary Univariate Operational Item Statistics - ELA .....	50
Table 4.B.6: Summary Univariate Field-Test Item Statistics - ELA.....	51
Table 4.B.7: Summary Univariate Field-Test Item Statistics - Mathematics .....	51
Table 4.B.8: IRT Model Data Fit Distribution of Operational Items - ELA .....	52
Table 4.B.9: IRT Model Data Fit Distribution of Operational Items - Mathematics .....	52
Table 4.B.10: IRT Model Data Fit Distribution of Field-Test Items - ELA.....	52
Table 4.B.11: IRT Model Data Fit Distribution of Field-Test Items - Mathematics.....	53
Table 4.B.12: Distribution of Operational Item DIF Classifications - ELA.....	53
Table 4.B.13: Distribution of Operational Item DIF Classifications - Mathematics.....	53
Table 4.B.14: Distribution of Field-Test Item DIF Classifications - ELA .....	54
Table 4.B.15: Distribution of Field-Test Item DIF Classifications - Mathematics .....	54
Table 4.C.1: Listing of Operational Items Exhibiting Significant Negative DIF March 2004 ...	55

Table 4.C.2: Listing of Field-Test Items Exhibiting Significant Negative DIF .....	55
Table 4.C.3: Listing of CR Item Statistics - ELA.....	58
Table 4.C.4: Summary of Univariate Operational Item Statistics - ELA .....	58
Table 4.C.5: Summary of Univariate Operational Item Statistics - Mathematics .....	59
Table 4.C.6: Summary of Univariate Field-test Item Statistics - ELA.....	59
Table 4.C.7: Summary of Univariate Field-test Item Statistics - Mathematics.....	59
Table 4.C.8: IRT Model Data Fit Distribution of Operational Items - ELA .....	60
Table 4.C.9: IRT Model Data Fit Distribution of Operational Items - Mathematics .....	60
Table 4.C.10: IRT Model Data Fit Distribution of Field-test Items - ELA .....	60
Table 4.C.11: IRT Model Data Fit Distribution of Field-test Items - Mathematics .....	60
Table 4.C.12: Distribution of Operational Item DIF Classifications - ELA.....	61
Table 4.C.13: Distribution of Operational Item DIF Classifications - Mathematics.....	61
Table 4.C.14: Distribution of Field-Test Item DIF Classifications - ELA .....	61
Table 4.C.15: Distribution of Field-Test Item DIF Classifications - Mathematics .....	62
Table 4.D.1: Listing of Operational Items Exhibiting Significant Negative DIF May 2004 .....	63
Table 4.D.2: Listing of Field-Test Items Exhibiting Significant Negative DIF .....	63
Table 4.D.3: Listing of CR Item Statistics - ELA.....	64
Table 4.D.4: Summary Univariate Operational Item Statistics - ELA .....	64
Table 4.D.5: Summary Univariate Operational Item Statistics - Mathematics .....	65
Table 4.D.6: Summary Univariate Field-Test Item Statistics - ELA.....	65
Table 4.D.7: Summary Univariate Field-Test Item Statistics - Mathematics.....	65
Table 4.D.8: IRT Model Data Fit Distribution of Operational Items - ELA .....	66
Table 4.D.9: IRT Model Data Fit Distribution of Operational Items - Mathematics .....	66
Table 4.D.10: IRT Model Data Fit Distribution of Field-Test Items - ELA .....	66
Table 4.D.11: IRT Model Data Fit Distribution of Field-Test Items - Mathematics.....	66
Table 4.D.12: Distribution of Operational Item DIF Classifications - ELA .....	67
Table 4.D.13: Distribution of Operational Item DIF Classifications - Mathematics.....	67
Table 4.D.14: Distribution of Field-Test Item DIF Classifications - ELA.....	67
Table 4.D.15: Distribution of Field-Test Item DIF Classifications - Mathematics.....	68
Table 5.A.1: Demographic Summary For All Examinees <sup>1</sup> ELA February 2004.....	73
Table 5.A.2: Demographic Summary For All Examinees <sup>1</sup> Mathematics February 2004.....	74
Table 5.A.3: NCLB Demographic Summary for All CAHSEE Examinees ELA February 2004	75
Table 5.A.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics February 2004.....	76
Table 5.A.5: Demographic Summary For All Examinees <sup>1</sup> ELA – February 2004.....	77
Table 5.A.6: Demographic Summary For All Examinees <sup>1</sup> Mathematics February 2004 .....	78
Table 5.A.7: Frequency Distributions ELA February 2004 .....	79
Table 5.A.8: Frequency Distributions Mathematics February 2004.....	79
Table 5.A.9: Frequency Distributions ELA February 2004 .....	80
Table 5.A.10: Frequency Distributions Mathematics February 2004.....	81
Table 5.A.11: Intercorrelations <sup>1</sup> and Reliability Estimates By Section February 2004 .....	82
Table 5.A.12: Agreement of First and Second Ratings on CR item.....	83
Table 5.A.13: Summary Statistics for CR item .....	83
Table 5.A.14: NCLB Reliability Classifications: February 2004.....	84
Table 5.A.15: Pass/Not Pass Classifications February 2004.....	85
Table 5.B.1: Demographic Summary For All Examinees <sup>1</sup> ELA March 2004.....	86

Table 5.B.2: Demographic Summary For All Examinees <sup>1</sup> Mathematics March 2004 .....	87
Table 5.B.3: NCLB Demographic Summary for All CAHSEE Examinees ELA March 2004....	88
Table 5.B.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics March 2004.....	89
Table 5.B.5: Demographic Summary For All Examinees <sup>1</sup> ELA March 2004.....	90
Table 5.B.6: Demographic Summary For All Examinees <sup>1</sup> Mathematics March 2004.....	91
Table 5.B.7: Frequency Distributions ELA March 2004.....	92
Table 5.B.8: Frequency Distributions Mathematics 2004 .....	93
Table 5.B.9: Frequency Distributions ELA March 2004.....	94
Table 5.B.10: Frequency Distributions Mathematics March 2004.....	95
Table 5.B.11: Intercorrelations <sup>1</sup> and Reliability Estimates By Section March 2004.....	96
Table 5.B.12: Agreement of First and Second Ratings on CR item .....	97
Table 5.B.13: Summary Statistics for CR item.....	97
Table 5.B.14: NCLB Classifications March 2004.....	98
Table 5.B.15: Pass/Not Pass Classifications March 2004 .....	99
Table 5.C.1: Demographic Summary For All Examinees <sup>1</sup> English Language Arts May 2004.	100
Table 5.C.2: Demographic Summary For All Examinees <sup>1</sup> Mathematics May 2004.....	101
Table 5.C.3: NCLB Demographic Summary for All CAHSEE Examinees English Language Arts May 2004 .....	102
Table 5.C.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics May 2004.....	103
Table 5.C.5: Demographic Summary For All Examinees <sup>1</sup> English Language Arts May 2004..	104
Table 5.C.6: Demographic Summary For All Examinees <sup>1</sup> Mathematics May 2004.....	105
Table 5.C.7: Frequency Distributions CAHSEE ELA May 2004 .....	106
Table 5.C.8: Frequency Distributions CAHSEE Mathematics May 2004 .....	107
Table 5.C.9: Frequency Distributions CAHSEE ELA May 2004 .....	108
Table 5.C.10: Frequency Distributions: CAHSEE Mathematics May 2004 .....	109
Table 5.C.11: Intercorrelations <sup>1</sup> and Reliability Estimates By Section May 2004.....	110
Table 5.C.12: Agreement of First and Second Ratings on CR item .....	111
Table 5.C.13: Summary Statistics for CR item.....	111
Table 5.C.14: NCLB Reliability Classifications May 2004 .....	112
Table 5.C.15: Pass/Not Pass Reliability Classifications May 2004.....	113
Table 6.1: Generalizability Results February 2004 .....	115
Table 6.2: Generalizability Results March 2004 .....	116
Table 6.3: Generalizability Results May 2004 .....	116
Table 6.A.1: Summary Statistics by Test Condition and Disability ELA February 2004.....	120
Table 6.A.2: Summary Statistics by Test Condition and Disability Mathematics February 2004 .....	121
Table 6.A.3: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type ELA February 2004 .....	122
Table 6.A.4: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type Mathematics February 2004 .....	123
Table 6.A.5: Percentiles of Scale Scores for Students with Accommodations* ELA February 2004.....	124
Table 6.A.6: Percentiles of Scale Scores for Students with Accommodations* Mathematics February 2004 .....	125

Table 6.A.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* ELA February 2004 .....	126
Table 6.A.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* Mathematics February 2004 .....	127
Table 6.A.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* ELA February 2004 .....	128
Table 6.A.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* Mathematics February 2004.....	129
Table 6.A.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* ELA February 2004 .....	130
Table 6.A.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* Mathematics February 2004.....	131
Table 6.B.1: Summary Statistics by Test Condition and Disability ELA .....	132
Table 6.B.2: Summary Statistics by Test Condition and Disability Mathematics.....	133
Table 6.B.3: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type ELA March 2004.....	134
Table 6.B.4: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type Mathematics March 2004.....	135
Table 6.B.5: Percentiles of Scale Scores for Students with Accommodations* ELA March 2004 .....	136
Table 6.B.6: Percentiles of Scale Scores for Students with Accommodations* Mathematics March 2004 .....	137
Table 6.B.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* ELA March 2004.....	138
Table 6.B.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* Mathematics March 2004 .....	139
Table 6.B.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* ELA March 2004 .....	140
Table 6.B.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* Mathematics March 2004.....	141
Table 6.B.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* ELA March 2004 .....	142
Table 6.B.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* Mathematics March 2004.....	143
Table 6.C.1: Summary Statistics by Test Condition and Disability ELA .....	144
Table 6.C.2: Summary Statistics by Test Condition and Disability Mathematics.....	145
Table 6.C.3: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type ELA May 2004.....	146
Table 6.C.4: Demographic Summary For All Examinees <sup>1</sup> by Accommodation Type Mathematics May 2004 .....	147
Table 6.C.5: Percentiles of Scale Scores for Students with Accommodations* ELA May 2004.....	148
Table 6.C.6: Percentiles of Scale Scores for Students with Accommodations* Mathematics May 2004.....	149
Table 6.C.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* ELA May 2004 .....	150

Table 6.C.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type* Mathematics May 2004 .....	150
Table 6.C.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* ELA May 2004.....	151
Table 6.C.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type* Mathematics May 2004.....	152
Table 6.C.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* ELA May 2004.....	153
Table 6.C.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type* Mathematics May 2004.....	153
Table 6.D.1: ELA February 2004 Frequency Distribution of Total Number of Erasures .....	154
Table 6.D.2: Mathematics February 2004 Frequency Distribution of Total Number of Erasures .....	155
Table 6.D.3: ELA March 2004 Frequency Distribution of Total Number Erasures .....	156
Table 6.D.4: Mathematics March 2004 Frequency Distribution of Total Number Erasures .....	157
Table 6.D.5: ELA May 2004 Frequency Distribution of Total Number Erasures .....	158
Table 6.D.6: Mathematics May 2004 Frequency Distribution of Total Number Erasures.....	159
Table 6.D.7: Number and Percentage* of Schools Meeting the Threshold Where Scores Changed from Not Pass to Pass as a Result of Erasures .....	160
Table 7.1: Characteristics of the Standard Setting Panelists.....	162
Table 7.2a ELA Educator Table Results by Round.....	168
Table 7.2b ELA Community Table Results by Round .....	169
Table 7.3a: Summary Median Passing Scores for ELA Shown as the Minimum Number of Points Out of 90 .....	169
Table 7.3b: Standard Errors of Judgment of Bookmarks Placements by Round.....	170
Table 7.4: ELA Consequence Data for Every Score Point.....	171
Table 7.5: ELA Consequence Data by Race/ethnicity and School Performance .....	174
Table 7.6: ELA Consequence Data by Gender and Special Populations .....	175
Figure 7.1: ELA Passing Scores by Race/ethnicity .....	176
Table 7.7a Mathematics Educator Table Results by Round.....	177
Table 7.7b Mathematics Community Table Results by Round .....	177
Table 7.8a: Summary Median Mathematics Passing scores by Round .....	178
Table 7.8b: Standard Errors of Judgment of Bookmark Placements by Round .....	178
Table 7.9: Mathematics Consequence Data for Every Score Point.....	179
Table 7.10: Mathematics Consequence Data by Race/ethnicity and School Performance .....	182
Table 7.11: Mathematics Consequence Data by Gender and Special Populations.....	183
Figure 7.2: Mathematics Passing Scores by Race/Ethnicity.....	184
Table 7.12: Percentage of ELA Panelists Reporting that Various Factors Were Very Influential in Placing the Bookmark.....	185
Table 7.13: Percentage of Mathematics Panelists Reporting that Various Factors Were Very Influential in Placing the Bookmark.....	187
Table 7.D.1: Quadrature Weights for March and May 2003 10 <sup>th</sup> Graders .....	204
Table 7.D.2: Quadrature Weights for Combined March and May 2003 10 <sup>th</sup> Graders by Subgroup - Mathematics.....	205
Table 7.D.3: Quadrature Weights for Combined March and May 2003 10 <sup>th</sup> Graders by Subgroup - ELA .....	206

Table 7.D.4: Predicted Passing Rates on the Standard Setting Forms Versus Actual Passing Rates (Actual Rates from March and May 2003).....	207
Figure 7.D.1: Simulated versus Actual Percentages of Males and Females at Various Raw Score Points—English Language Arts MC Items.....	208
Table 7.E.1: Sample Mathematics Conversion Table.....	212
Table 7.E.2: Sample ELA Conversion Table.....	212

## 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 and Arts (ELA) and mathematics set by the California State Board of Education (SBE). The CAHSEE is typically administered five times per year to allow several testing opportunities each year for those students who have not yet passed the examination. This report covers the three CAHSEE test administrations given in February, March and May 2004.

The test forms administered in 2004 were based on revised test blueprints for both the ELA and mathematics sections. In ELA, the number of constructed-response (CR) items was reduced from two to one and the number of multiple-choice (MC) items was reduced from 82 to 72 in order to reduce the administration burden from two days to one day. In addition, the relative numbers of items in other ELA strands were revised and the relative weight of the CR item in relation to the MC items was changed so that the weight of writing in the total score remained at about 50%. This resulted in a 90-point score comprised of 45 points from MC reading items, 27 points based on MC writing items, and 18 points based on a CR writing item.

For mathematics, a change in test difficulty specifications was made after research indicated that the California Content Standards could be validly assessed using items that were slightly less complex, particularly in the areas of algebra and measurement and geometry. Test difficulty specifications for the CHASEE mathematics test were also changed so that the new version of the test would be slightly easier than in the past.

In addition, new standards were set for both the ELA and mathematics sections based on the new test blueprint and difficulty specifications in September 2003. The new standards were subsequently applied and a new reporting scale established beginning with the February 2004 administration.

Each test form was constructed from items that had been previously administered and placed onto the operational scale. Each test form also included a set of anchor items that were used to maintain the operational (theta) scale. For February and March, approximately 30 items were common between the two forms. These items had been previously administered in an operational test form. The May forms included 20-25 anchor items, of which, approximately half were from the February 2004 administration and the other half were from the March 2004 administration. All items included on operational test forms had been evaluated by review panels for bias and sensitivity and to certify that each item was matched correctly with its intended content standard. In addition, each test form was reviewed and approved by the California Department of Education (CDE).

Table E.1 presents the administration dates and total number of examinees taking one or both CAHSEE content areas during the February, March and May 2004 administrations. The majority of examinees in February and March were first time examinees. Examinees taking only one content area were repeat examinees that did not pass that content area during a previous administration.

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

<b>Administration</b>	<b>Total Examinees</b>	<b>ELA &amp; Math</b>	<b>ELA Only</b>	<b>Math Only</b>
February 2004	151,436	140,726	5,737	4,973
March 2004	299,650	280,963	10,419	8,268
May 2004	16,375	7,018	4,998	4,359

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**

<b>Administration</b>	<b>English Language Arts</b>		<b>Mathematics</b>	
	<b>N Tested</b>	<b>N Passed (%)</b>	<b>N Tested</b>	<b>N Passed (%)</b>
February 2004	146,463	107,748 (74%)	145,699	104,322 (72%)
March 2004	291,382	221,494 (76%)	289,231	219,389 (76%)
May 2004	12,016	5,927 (49%)	11,377	5,190 (46%)

The Educational Testing Service (ETS) conducted a series of analyses to evaluate the items and test forms for each of the CAHSEE administrations. For each administration the following analyses were completed: classical item analyses, differential item functioning, item response theory (IRT) calibration, scaling and equating. In addition, scoring tables were created for each operational test form and reliability indices were calculated. In addition, this technical report includes results from the following studies: inter-rater agreement for the ELA CR item, generalizability, decision accuracy, and mark discrimination studies. Additional summary analyses conducted for students having special accommodation needs are included in this report, as well as the results from the September 2003 standard setting study.

All item analyses, including calibration, equating and scaling, were completed using GENASYS (ETS proprietary software) or commercially available software (e.g., SAS, SPSS, GENOVA). In all cases, analyses were conducted on complete records for each content area (i.e., students must have attempted at least 5 items on the test form). 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. 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 the February, March and May CAHSEE administrations included in this report are presented in tables E.3 and E.4 below. These statistics indicate satisfactory performance of the test form constructed for these examinations.

**Table E.3: CAHSEE Summary Statistics—English Language Arts**

<b>Administration</b>	<b>February 2004</b>	<b>March 2004</b>	<b>May 2004</b>
<b>Scale Score Information</b>			
Number of Examinees	146463	291382	12016
Mean	374	379	351
SD	38	40	41
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	375	381	348
<b>Test Information</b>			
Reliability	0.94	0.95	0.95
Raw Score Standard Error of Measurement	3.86	3.73	4.21
Mean Omits	0.40	0.34	1.69
SD Omits	3.59	3.26	8.30
% Responding all items	93	92	87
% responding all items – 1 item	98	98	94
% responding all items –2 items	99	99	95
% responding all items –3 items	99	99	96
% responding all items –4 items	99	99	96
% responding all items –5 items	99	99	96
<b>Item Information*</b>			
Number of Items	72	72	72
Mean Observed Average Item Score (AIS)	0.71	0.73	0.60
Equated Mean Rasch B-Value	-0.01	0.05	-0.04
Mean r-Biserial	0.53	0.55	0.58
SD r-Biserial	0.09	0.09	0.09
<b>Examinee Information</b>			
<b>ELA Only Examinees</b>			
Number of Examinees	5737	10419	4998
Mean Scale Score	343	344	349
Scale Score SD	36	39	42
Median Scale Score	339	340	346
<b>ELA and Mathematics Examinees</b>			
Number of Examinees	140726	280963	7018
Mean Scale Score	375	380	352
Scale Score SD	37	39	41
Median Scale Score	375	381	350

Note. \*Means and standard deviations for item information section are computed on 72 MC items only

**Table E.4: CAHSEE Summary Statistics—Mathematics**

<b>Administration</b>	<b>February 2004</b>	<b>March 2004</b>	<b>May 2004</b>
<b>Scale Score Information</b>			
Number of Examinees	145699	289231	11377
Mean	374	379	350
SD	37	38	34
Possible Range	275-450	275-450	275-450
Obtained Range	275-450	275-450	275-450
Median	371	376	345
<b>Test Information</b>			
Reliability	0.94	0.95	0.94
Raw Score Standard Error of Measurement	3.71	3.64	3.93
Mean Omits	0.28	0.23	.99
SD Omits	2.54	2.16	6.04
% Responding all items	91	91	86
% responding all items – 1 item	97	98	95
% responding all items –2 items	99	99	97
% responding all items –3 items	99	99	97
% responding all items –4 items	99	99	98
% responding all items –5 items	99	100	98
<b>Item Information</b>			
Number of Items	80	80	80
Mean Observed Average. Item Score (AIS)	0.66	0.67	0.52
Equated Mean Rasch B-Value	-0.24	-0.17	-0.22
Mean r-Biserial	0.54	0.55	0.52
SD r-Biserial	0.09	0.09	0.12
<b>Examinee Information</b>			
<b>Mathematics Only Examinees</b>			
Number of Examinees	4973	8268	4359
Mean Scale Score	349	351	350
Scale Score SD	32	32	34
Median Scale Score	345	347	345
<b>ELA and Mathematics Examinees</b>			
Number of Examinees	140726	280963	7018
Mean Scale Score	374	380	350
Scale Score SD	37	37	34
Median Scale Score	371	378	345

The chapters that follow describe the technical procedures applied to the CAHSEE for the three 2004 administrations, and the results of statistical analyses based on the data from these administrations. Chapter 1 provides background information and Chapter 2 summarizes the statistical specifications followed in test form assembly. The equating procedures and documentation of the three equating results are presented in Chapter 3. Chapter 4 summarizes item-level analyses of the three 2004 CAHSEE administrations, including classical item analyses, DIF, IRT analyses, and evaluations of model-data fit. Chapter 5 summarizes test-level analyses, including estimates of reliability, decision consistency, decision accuracy, and technical characteristics of the CR item scoring for ELA. Chapter 6 describes a number of special analyses applied to the three CAHSEE administrations, including summaries of test results for students with disabilities and special accommodations, generalizability analyses, and mark discrimination analyses on the CAHSEE mathematics and ELA tests by school. Chapter 7 describes the procedures used and results of the September 2003 standard setting study.

## **Chapter 1: 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 of 2001. Beginning October 1, 2001, CDE entered a contract with Educational Testing Service (ETS) and its subcontractors for the development and administration of the CAHSEE. In 2004, the CAHSEE was administered three times. This report covers the three administrations: February 2004, March 2004 and May 2004.

### **Target Population**

The target population for the CAHSEE is all California public high school students. The 2004 administrations consisted primarily of 10<sup>th</sup> grade students testing for the first time.

### **Test Description**

The CAHSEE offers an examination in mathematics and an examination in English-language arts (ELA). Students may take either one or both examinations in a single administration. Both examinations are right-scored (i.e., 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 three administrations, although it was not necessary to use this form.

The ELA examination measures reading and writing skills as defined by SBE standards through grade 10. The reading portion of the examination covers topics such as vocabulary and informational and literary reading. The writing portion of the examination covers topics such as writing strategies, applications, and conventions. The CAHSEE mathematics examination measures standards adopted by the State Board of Education (SBE) through Algebra I. It covers topics such as statistics, data analysis and probability, number sense, measurement and geometry, algebra, and mathematical reasoning.

Beginning with the February 2004 administration, several changes in the test blueprints were made. To reduce the administration burden from two days to a single day, the ELA portion of the CAHSEE was revised in several ways. The revisions included reducing the number of CR items from two to one and reducing the number of MC items from 82 to 72. In addition, the relative numbers of items in other ELA strands were revised and the relative weight of the CR item in relation to the MC items was changed so that the weight of writing in the total score remained at about 50 percent. This resulted in a 90-point score that is comprised of 45 points based on MC reading items, 27 points based on MC writing items, and 18 points based on a CR item writing item. The CR item is related either to a literature or informational passage or to a written response to a writing prompt.

For mathematics, a change in test difficulty specifications was made after research indicated that the California Content Standards could be validly assessed using items that were slightly less

complex, particular in areas such as algebra and measurement and geometry. Test difficulty specifications for the CAHSEE mathematics test were changed so that the new version of the test would be slightly easier than in the past. The mathematics examination consists of 80 MC items.

In addition to the operational items, the test forms also include field-test items. In order to maintain the item pool for construction of future forms, multiple forms were administered, each containing the same operational item set and a set of unique MC field-test items. In mathematics, the field-test section consisted of 12 MC items, and in ELA, the field-test section consisted of 7 MC items per form. ELA CR items are field-tested separately from the operational CAHSEE administrations. A summary of the items included in the three administrations is shown in Table 1.1.

**Table 1.1: Items Included in the 2004 Administrations**

<b>Administration</b>	<b>Operational Items</b>	<b>Field-Test Forms</b>	<b>Number of Field-Test Items*</b>
<b>February 2004</b>			
English-Language Arts	72 MC, 1 CR	25	174
Mathematics	80 MC	25	298
<b>March 2004</b>			
English-Language Arts	72 MC, 1 CR	94	655
Mathematics	80 MC	95	1139
<b>May 2004</b>			
English-Language Arts	72 MC, 1 CR	25	172
Mathematics	80 MC	25	300

\*Note. In some cases the same field-test items may have appeared in more than one form.

### **Scores for Analysis and Reporting**

The MC CAHSEE items are right-scored, each correct response is worth one point, and the points are summed to calculate the total MC score. The CR item included on the ELA test is scored on a rubric ranging from 1 to 4 points. Constructed-response scores in half-point intervals are possible because two raters score the writing exercise and the results are averaged 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 Online Scoring Network (OSN) is used for scoring CAHSEE CR items. This system was developed at ETS for the purpose of transmitting CR items electronically to readers at remote locations via the Internet. CR items are displayed on a PC screen to readers who enter their scores via simple mouse clicks. The structure of a typical scoring session, whether half-day or full day, is standardized. At the beginning of each scoring session, readers enter the OSN software and score a calibration set - a set of pre-scored responses to one topic arranged in an electronic folder. Each scoring leader can monitor reader performance via the OSN software. Depending on a reader's results (exact, adjacent, or discrepant scores), the scoring leader will either allow the reader to proceed to operational scoring of examinee responses or will require the reader to score an additional calibration set. If the reader cannot calibrate successfully, he or

she is not allowed to score operational CR items. OSN automates the distribution of sets of papers to the required number of readers. Two readers are provided for each response, with a third scorer providing resolution of non-adjacent scores. As with paper and pencil scoring, the third reading, when required, is provided by staff with more experience and authority, typically a scoring leader. OSN produces topic and reader summary reports for future reference by Test Development (topics) and OSN staff (readers).

### **Weighting Scheme for Reported Scores**

The ELA section consists of 45 MC items measuring reading, 27 MC items measuring writing, and one CR prompt. The weighting of these components as follows:

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% of the total points)

Total points CR = 18 points (20% of the total points)

Total points: Composite (MC + Weighted CR) = 90 points.

There is no special weighting scheme for the mathematics section – each item is worth one score point.

In all cases, total raw scores on the CAHSEE are transformed to a reported scale that ranges from 275 to 450, with the minimum passing score set at 350.

### **Subscores**

In addition to total scores, student performance is reported for a number of content strands associated with each test. Table 1.2 below summarizes the subscores reported for the CAHSEE. In addition, a CR item score is reported. Number correct and percentage correct scores are reported for each content strand. In mathematics, a separate score based on items that measure Mathematical Reasoning (8 items total) in addition to the primary strand classification is calculated but not reported to students.

**Table 1.2: Mathematics and ELA Subscores**

Mathematics Content Standard	Number of Items	ELA Content Standard	Number of Items
Probability & Statistics (PS)	13	Word Analysis (RV)	7
Number Sense (NS)	17	Reading Comprehension (RC)	18
Algebra & Functions (AF)	20	Literary Responses & Analysis (RL)	20
Measurement & Geometry (MG)	18	Writing Strategies (WS)	12
Algebra 1 (A1)	12	Writing Conventions (WC)	15
Mathematical Reasoning (MR) <sup>1</sup>	8	Writing Applications-CR (WA)	1

Note. <sup>1</sup>Items in this category are also classified under one of the other identified strands.

### **New CAHSEE Passing Scores**

In September 2003, a standard setting study was carried out using CAHSEE test forms that were constructed to meet the new blueprints and difficulty specifications (see Chapter 7). In November 2003, after reviewing the results of this study, the California Board of Education set new passing scores on the CAHSEE corresponding to specific score levels on the test forms used in the standard settings: 55 percent correct on the mathematics portion (44 out of 80) and 60 percent of the points on the ELA portion (54 points out of 90). As a result, a new reporting scale was established beginning with the February 2004 administration. Additional details are presented in Chapter 3.

### **Raw Score to Scale Score Conversion**

Students taking the CAHSEE have multiple opportunities to take the exam 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, number correct or raw scores are converted to scale scores. 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 250 to 450, with the minimum passing score set at 350. As described in Chapter 3 of this report, a new scale was set for CAHSEE in February 2004.

## Chapter 2: CAHSEE Test Development Procedures

The CAHSEE Test Development procedures in preparation for the 2004 administrations included new item development, item review, and test form assembly.

Test forms were constructed to specific specifications. As described earlier, several changes in the test blueprints were made prior to the February 2004 testing. The ELA portion of the CAHSEE was revised by reducing the number of CR and MC items. In addition, the relative numbers of items in the ELA strands were revised and the relative weight of the CR item in relation to the MC items was changed so that the weight of writing in the total score remained at about 50 percent. This resulted in a 90-point score that is comprised of 45 points based on MC reading items, 27 points based on MC writing items, and 18 points based on a CR item writing item. For mathematics, the test difficulty specifications for the CAHSEE mathematics test were changed so that the new version of the test would be slightly easier.

To begin the test construction process, test developers first constructed anchor item sets that were representative of the test form with regard to content and item difficulty. The numbers of anchor items varied. Approximately 30 items were identified as anchor items for the February and March administrations; 20-25 items were identified as anchor items for the May 2004 administration. Approximately one-half of the anchor items in the May form were from the February and March administrations in order to minimize item exposure for students retaking the test in another testing cycle. Initial anchor item selections were reviewed and approved prior to the selection of the remaining operational items.

Tables 2.1 to 2.4 list the recommended statistical specifications for CAHSEE test assembly, articulated in terms of equated Rasch item difficulty values (b-values) and item biserial correlations for the total test. All CAHSEE items have been calibrated and scaled to the CAHSEE item bank. In general, test developers are asked to match the statistical characteristics as closely as possible. The distributions of items in each of the intervals are used as guidelines by the test developers to match the mean equated Rasch difficulty for each form. The guidelines for item biserial correlations are lower priority because the characteristics of CAHSEE student samples differ considerably from administration to administration, and these differences affect the magnitudes of the biserial correlations. Tables 2.2 and 2.4 provide the statistical guidelines for strands in each of the test forms.

In addition to difficulty specifications, information about model-data fit is taken into account during test assembly (see Chapter 4 for a description of the procedures used for evaluating model-data fit). Test developers were instructed to exclude items with fit classifications F, as well as items that had been flagged for severe (C-) DIF.

Once constructed, the forms are reviewed and approved by the CDE.

**Table 2. 1: Difficulty (B) and Discrimination (Rbis) Specifications for ELA**

<b>Low B</b>	<b>High B</b>	<b>Specifications</b>
-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
<b># MC Items</b>		<b>72</b>
<b>MC Mean</b>		<b>-0.10 - 0.10</b>
<b>MC SD</b>		<b>0.55 - 0.70</b>
<b>Low Rbis</b>	<b>High Rbis</b>	<b>Specifications</b>
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
<b># MC Items</b>		<b>72</b>
<b>MC Mean</b>		<b>0.44 - 0.54</b>
<b>MC SD</b>		<b>0.10 - 0.15</b>

**Table 2. 2: Number of Items and Mean B-Values by Content Strand for ELA**

<b>Content Class</b>	<b># items</b>	<b>Mean B</b>
Reading Comprehension	18	-0.05 - -0.20
Reading Literary Analysis	20	-0.20 - -0.00
Reading Vocabulary	7	-0.40 - -0.15
Writing Conventions	15	-0.03 - -0.25
Writing Strategies	12	0.07 - -0.50
<b>Total</b>	<b>72</b>	

**Table 2. 3: Difficulty (B) and Discrimination (Rbis) Specifications for Mathematics**

<b>Low B</b>	<b>High B</b>	<b>Specifications</b>
-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
<b># Items</b>		<b>80</b>
<b>Mean</b>		<b>-0.30 - -0.20</b>
<b>SD</b>		<b>0.65 - 0.80</b>
<b>Low Rbis</b>		
<b>Low Rbis</b>	<b>High Rbis</b>	<b>Specifications</b>
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
<b># Items</b>		<b>80</b>
<b>Mean</b>		<b>0.44 - 0.54</b>
<b>SD</b>		<b>0.10 - 0.15</b>

**Table 2. 4: Number of Items and Mean B-Values by Content Strand for Mathematics**

<b>Content Class</b>	<b># Items</b>	<b>Mean B</b>
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
<b>Total</b>	<b>80</b>	

## Chapter 3: Equating Procedures

### Item Calibration and Transformation

Equating the CAHSEE involves three steps: item calibration, item parameter scaling, and true-score equating. The samples used for item calibration, scaling, and equating include scanned and scored student records provided in statistical file extracts. Typically, over 95 percent of the data are available for equating, although in the March administration a slightly lower percentage of the ELA records were available because of the large numbers of CR items scored. Students taking special test forms (Large Print, CD-ROM 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 score equating and production of scoring tables for score reporting, calibration, scaling and equating analyses are conducted including only the operational items. Field-test items are analyzed and calibrated separately.

A proprietary version of the PARSCALE computer program (Muraki & Bock, 1995) was used for all item calibration work. This program estimates parameters for a 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 (polytomously scored) items, as required by the CAHSEE program.

The PARSCALE program is run in two stages. In the first stage, the estimation imposes normal constraints on the updated prior distribution. 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 (E-M) 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 scaled to the bank estimates 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 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 and the linking constants re-estimated. The differences are calculated as follows:

$$WRMSD = \sqrt{\sum_{j=1}^{61} w_j [P_n(\theta_j) - P_r(\theta_j)]^2}, \quad (1)$$

where  $\theta_j$  ranges from  $-3.0$  to  $3.0$  by  $0.1$ ,  $w_j$  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  $j$ , and  $P_r(\theta_j)$  is the probability of correct response for the reference form (i.e., the 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 calibration work and has been used satisfactorily across a number of CAHSEE administrations.

For each administration, plots of new transformed difficulty estimates against the original bank estimates are then produced and inspected. Results across administrations have consistently indicated high correlations between the new and original bank difficulty estimates, and typically no more than two or three items are eliminated from the linkings 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.

### **Establishing the Reporting Scale for the New CAHSEE: February 2004**

Following the February 2004 administration, new cut-scores were established. Although the majority of the items included in the test forms for this administration were the same items that were used in the September 2003 standard setting, some of the MC items on the two tests were different, and the CR item used for the ELA portion of the CAHSEE was different. Because of these differences, it was necessary to translate the cut-scores of 55 percent-correct for mathematics and the 60 percent-correct for ELA, established by the SBE, into matching performance levels on the February tests.

Once 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. First, 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 established in the standard setting (.5356 for ELA and -.0701 for mathematics, see Table 3.1). The raw scores associated with passing on the February 2004 form were 54 for ELA and 43 for mathematics<sup>1</sup>.

**Table 3.1: 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

<sup>1</sup> 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 cutpoints.

Second, the new reporting scales for the ELA and mathematics portion 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 (see Tables 3.2 and 3.3). The transformation constants associated with these scoring tables are as follows:

$$\text{ELA: Scale Score} = \text{Theta} * 33.7230 + 332.1605$$

$$\text{Mathematics: Scale Score} = \text{Theta} * 32.2900 + 352.2119$$

The resulting scale has several notable characteristics:

1. The passing raw scores (54 and 43, respectively) are set to Scale scores of 350.
2. The “Proficient” cut scores to be used for NCLB accountability purposes have 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 communication of these important NCLB cut-scores to a wide audience, 380 was used for both ELA and mathematics.
3. The “Advanced” cut scores to be used for NCLB accountability purposes were at 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, respectively based on the following reasoning. Ideally, the minimum scale score on both ELA and mathematics would be set at the raw score level that a student might be expected to achieve by random guessing. Because the 350 level on the ELA and mathematics scales are at different points on the scales, it is not possible to place a common minimum scale score at the chance performance level on both tests. Given that CAHSEE MC items have four choices, chance-level performance begins at approximately 25 percent correct and 20 percent correct is arguably a conservative standard to characterize performance below chance. The highest raw score corresponding to the recommended minimum scale score of 275 is close to 20 percent-correct performance for ELA and is below the 20 percent-correct performance level for mathematics. Alternatively, a minimum scale score of 300 could be considered, which would result in the highest raw score corresponding to a minimum scale score of 300 being 36 percent-correct for ELA and 22.5 percent-correct for mathematics. A minimum scale score of 300 would be reasonable for mathematics but not for ELA. Thus, a minimum scale score of 275 was judged to provide the best compromise for setting a common minimum reported scale score for the two tests.

**Table 3.2: February 2004 CAHSEE ELA Reporting Scale**

Min SS: 275      Max SS: 450  
 RS at Cut: 54      SS at Cut: 350      Slope: 33.7230  
 RS at Y1: 68      SS at Y1: 380      Intercept: 332.1605

RS	Theta	Scaled Score	Rounded SS	RS	Theta	Scaled Score	Rounded SS
0	-39.0000	275.0000	275	46	0.0894	335.1754	335
1	-4.7633	275.0000	275	47	0.1430	336.9829	337
2	-4.0426	275.0000	275	48	0.1968	338.7972	339
3	-3.6103	275.0000	275	49	0.2511	340.6284	341
4	-3.2967	275.0000	275	50	0.3057	342.4696	342
5	-3.0486	275.0000	275	51	0.3608	344.3278	344
6	-2.8420	275.0000	275	52	0.4163	346.1994	346
7	-2.6643	275.0000	275	53	0.4724	348.0913	348
8	-2.5079	275.0000	275	54	0.5290	350.0000	350
9	-2.3679	275.0000	275	55	0.5863	351.9323	352
10	-2.2410	275.0000	275	56	0.6442	353.8849	354
11	-2.1246	275.0000	275	57	0.7028	355.8611	356
12	-2.0169	275.0000	275	58	0.7622	357.8642	358
13	-1.9167	275.0000	275	59	0.8224	359.8943	360
14	-1.8228	275.0000	275	60	0.8835	361.9548	362
15	-1.7344	275.0000	275	61	0.9457	364.0524	364
16	-1.6507	276.4939	276	62	1.0089	366.1837	366
17	-1.5711	279.1783	279	63	1.0734	368.3588	368
18	-1.4952	281.7379	282	64	1.1391	370.5744	371
19	-1.4224	284.1929	284	65	1.2064	372.8440	373
20	-1.3525	286.5501	287	66	1.2753	375.1675	375
21	-1.2851	288.8231	289	67	1.3460	377.5517	378
22	-1.2200	291.0184	291	68	1.4186	380.0000	380
23	-1.1568	293.1497	293	69	1.4936	382.5292	383
24	-1.0953	295.2237	295	70	1.5710	385.1394	385
25	-1.0354	297.2437	297	71	1.6512	387.8440	388
26	-0.9768	299.2199	299	72	1.7345	390.6531	391
27	-0.9194	301.1556	301	73	1.8214	393.5836	394
28	-0.8631	303.0542	303	74	1.9123	396.6491	397
29	-0.8076	304.9258	305	75	2.0077	399.8662	400
30	-0.7529	306.7705	307	76	2.1082	403.2554	403
31	-0.6988	308.5949	309	77	2.2145	406.8402	407
32	-0.6453	310.3991	310	78	2.3276	410.6542	411
33	-0.5922	312.1897	312	79	2.4482	414.7212	415
34	-0.5395	313.9670	314	80	2.5775	419.0816	419
35	-0.4870	315.7374	316	81	2.7170	423.7860	424
36	-0.4348	317.4978	317	82	2.8685	428.8950	429
37	-0.3826	319.2581	319	83	3.0345	434.4930	434
38	-0.3306	321.0117	321	84	3.2189	440.7116	441
39	-0.2785	322.7687	323	85	3.4277	447.7529	448
40	-0.2264	324.5256	325	86	3.6718	450.0000	450
41	-0.1742	326.2860	326	87	3.9723	450.0000	450
42	-0.1219	328.0497	328	88	4.3797	450.0000	450
43	-0.0694	329.8201	330	89	5.0577	450.0000	450
44	-0.0167	331.5973	332	90	20.0000	450.0000	450
45	0.0362	333.3813	333				

**Table 3.3: February 2004 CAHSEE Mathematics Reporting Scale**

Min SS            275 Max SS            450  
 RS at Cut        43 SS at Cut:            350 Slope:        32.2900  
 RS at Y1        58 SS at Y1            379.5001 Intercept: 352.2119

RS	Theta	Scaled Score	Rounded SS	RS	Theta	Scaled Score	Rounded SS
0	-39.0000	275.0000	275	41	-0.1813	346.3577	346
1	-4.8768	275.0000	275	42	-0.1249	348.1788	348
2	-4.1635	275.0000	275	43	-0.0685	350.0000	350
3	-3.7374	275.0000	275	44	-0.0118	351.8308	352
4	-3.4290	275.0000	275	45	0.0451	353.6681	354
5	-3.1850	275.0000	275	46	0.1023	355.5151	356
6	-2.9817	275.0000	275	47	0.1599	357.3750	357
7	-2.8064	275.0000	275	48	0.2180	359.2511	359
8	-2.6515	275.0000	275	49	0.2766	361.1433	361
9	-2.5123	275.0000	275	50	0.3359	363.0581	363
10	-2.3854	275.1874	275	51	0.3958	364.9922	365
11	-2.2683	278.9686	279	52	0.4566	366.9555	367
12	-2.1594	282.4849	282	53	0.5184	368.9510	369
13	-2.0573	285.7817	286	54	0.5811	370.9756	371
14	-1.9610	288.8913	289	55	0.6450	373.0389	373
15	-1.8696	291.8426	292	56	0.7102	375.1442	375
16	-1.7825	294.6550	295	57	0.7768	377.2947	377
17	-1.6991	297.3480	297	58	0.8451	379.5001	380
18	-1.6189	299.9377	300	59	0.9151	381.7604	382
19	-1.5417	302.4304	302	60	0.9872	384.0885	384
20	-1.4669	304.8457	305	61	1.0615	386.4876	386
21	-1.3945	307.1835	307	62	1.1384	388.9707	389
22	-1.3241	309.4567	309	63	1.2181	391.5443	392
23	-1.2555	311.6718	312	64	1.3011	394.2243	394
24	-1.1885	313.8353	314	65	1.3879	397.0271	397
25	-1.1229	315.9535	316	66	1.4789	399.9655	400
26	-1.0587	318.0265	318	67	1.5748	403.0621	403
27	-0.9956	320.0640	320	68	1.6765	406.3460	406
28	-0.9336	322.0660	322	69	1.7851	409.8527	410
29	-0.8726	324.0356	324	70	1.9018	413.6209	414
30	-0.8123	325.9827	326	71	2.0285	417.7120	418
31	-0.7528	327.9040	328	72	2.1674	422.1971	422
32	-0.6940	329.8026	330	73	2.3220	427.1891	427
33	-0.6357	331.6851	332	74	2.4971	432.8431	433
34	-0.5780	333.5483	334	75	2.7003	439.4044	439
35	-0.5206	335.4017	335	76	2.9442	447.2799	447
36	-0.4636	337.2422	337	77	3.2526	450.0000	450
37	-0.4068	339.0763	339	78	3.6785	450.0000	450
38	-0.3503	340.9007	341	79	4.3920	450.0000	450
39	-0.2939	342.7218	343	80	20.0000	450.0000	450
40	-0.2376	344.5398	345				

## True-Score Equating March and May 2004

For the March and May 2004 administration, once the items 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 (February 2004). The base forms consist of item parameter estimates from PARSCALE calibrations of data from the February 2004 administrations that have been linked to the new CAHSEE scale. 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):

$$\xi(\theta) = \sum_{i=1}^n P_i(\theta), \quad (2)$$

where  $P_i(\theta)$  is the probability of a correct response to item  $i$  at ability level  $\theta$  (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:

$$\xi(\theta) = w_{mc} \sum_{i=1}^{nmc} P_i(\theta) + w_{cr} \sum_{j=1}^{ncr} \sum_{x=1}^m s_x P_{xj}(\theta), \quad (3)$$

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

For each integer score  $\xi_n$  on the new form, the true-score equating procedure first solves for the corresponding ability level using equations 2 (for mathematics) or 3 (for ELA). Next, the procedure uses that ability level ( $\theta$ ) to find the corresponding score,  $\xi_b$ , on the base form. Finally, each score  $\xi_b$  is transformed to the CAHSEE scale-score scale using the raw-score-to-scale-score conversion table developed for the February 2004 administration and linear interpolation.

## Braille, Large Print, and Audio CD Form Equating

Separate IRT true score equating can be carried out for versions of the CAHSEE test forms in situations where these test forms do not consist of all of the items that are administered in the operational tests. In the past, this has occurred only for mathematics in cases where, in the judgment of test development experts, it was not possible to translate all of the items to Braille without compromising the validity of the items. In these cases, the Braille form equating uses the operational item difficulty estimates for only the items that were included in a particular form. The February 2004 form was the version used for the Braille form for all three administrations. This form was identical to the standard form, so no special equating analyses were required. Similarly, the large print and audio CD versions of the test forms were identical to the standard form administered in February, March and April.

## **Conversion Tables and Conditional Standard Errors of Measurement**

Following the equating analyses, scale score conversion tables and conditional standard errors of measurement (CSEMs) were produced. CSEMs for CAHSEE scale scores are based on item response theory (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 CSEMs 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 CSEMs estimated in GENASYS are characterized by minor irregularities that are smoothed in a subsequent step. Operational and audio CD score conversions and the smoothed CSEMs at score points for the ELA and mathematics tests are presented in Tables 3.A.1 to 3.C.2 and from 3.D.1 through 3.D.2 for the Braille forms. Appendix 3.E presents equations for calculating the standard errors of theta based on weighted raw scores and using the Rasch and Rasch partial credit model.

**Appendix 3.A: Scoring Tables for Operational & Audio CD Conversions—February 2004**

**Table 3.A.1: Operational and Audio CD Conversion – ELA February 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM	Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5856	450		44	331.5973	332	9
89	502.7214	450		43	329.8201	330	9
88	479.8572	450		42	328.0497	328	9
87	466.1185	450		41	326.2860	326	9
86	455.9847	450	23	40	324.5256	325	9
85	447.7529	448	21	39	322.7687	323	9
84	440.7116	441	17	38	321.0117	321	9
83	434.4930	434	17	37	319.2581	319	9
82	428.8950	429	17	36	317.4978	317	9
81	423.7860	424	15	35	315.7374	316	9
80	419.0816	419	14	34	313.9670	314	9
79	414.7212	415	14	33	312.1897	312	9
78	410.6542	411	13	32	310.3991	310	9
77	406.8402	407	13	31	308.5949	309	9
76	403.2554	403	12	30	306.7705	307	9
75	399.8662	400	12	29	304.9258	305	9
74	396.6491	397	12	28	303.0542	303	10
73	393.5836	394	11	27	301.1556	301	10
72	390.6531	391	11	26	299.2199	299	10
71	387.8440	388	11	25	297.2437	297	10
70	385.1394	385	10	24	295.2237	295	10
69	382.5292	383	10	23	293.1497	293	10
68	380.0000	380	10	22	291.0184	291	10
67	377.5517	378	10	21	288.8231	289	11
66	375.1675	375	10	20	286.5501	287	11
65	372.8440	373	10	19	284.1929	284	11
64	370.5744	371	9	18	281.7379	282	11
63	368.3588	368	9	17	279.1783	279	12
62	366.1837	366	9	16	276.4939	276	12
61	364.0524	364	9	15	273.6713	275	12
60	361.9548	362	9	14	270.6902	275	
59	359.8943	360	9	13	267.5236	275	
58	357.8642	358	9	12	264.1446	275	
57	355.8611	356	9	11	260.5126	275	
56	353.8849	354	9	10	256.5872	275	
55	351.9323	352	9	9	252.3078	275	
54	350.0000	350	9	8	247.5866	275	
53	348.0913	348	9	7	242.3123	275	
52	346.1994	346	9	6	236.3197	275	
51	344.3278	344	9	5	229.3525	275	
50	342.4696	342	9	4	220.9858	275	
49	340.6284	341	9	3	210.4103	275	
48	338.7972	339	9	2	195.8318	275	
47	336.9829	337	9	1	171.5277	275	
46	335.1754	335	9	0	147.2240	275	
45	333.3813	333	9				

**Table 3.A.2: Operational and Audio CD Conversion – Mathematics February 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	
79	494.0284	450	
78	470.9871	450	
77	457.2381	450	19
76	447.2799	447	17
75	439.4034	439	16
74	432.843	433	14
73	427.1891	427	13
72	422.1971	422	12
71	417.712	418	12
70	413.6199	414	11
69	409.8524	410	11
68	406.3459	406	10
67	403.0621	403	10
66	399.9655	400	10
65	397.0271	397	10
64	394.2243	394	9
63	391.5435	392	9
62	388.9706	389	9
61	386.4876	386	9
60	384.0885	384	9
59	381.7604	382	9
58	379.5001	380	8
57	377.2947	377	8
56	375.1442	375	8
55	373.0389	373	8
54	370.9756	371	8
53	368.951	369	8
52	366.9548	367	8
51	364.9918	365	8
50	363.058	363	8
49	361.1432	361	8
48	359.2511	359	8
47	357.375	357	8
46	355.5151	356	8
45	353.6681	354	8
44	351.8306	352	8
43	350	350	8
42	348.1788	348	8
41	346.3577	346	8
40	344.5398	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.7218	343	8
38	340.9007	341	8
37	339.0763	339	8
36	337.2422	337	8
35	335.4017	335	8
34	333.5483	334	8
33	331.6851	332	8
32	329.8026	330	8
31	327.904	328	8
30	325.9828	326	8
29	324.0359	324	8
28	322.066	322	8
27	320.064	320	8
26	318.0265	318	8
25	315.9535	316	8
24	313.8354	314	8
23	311.6718	312	8
22	309.4567	309	9
21	307.1835	307	9
20	304.8457	305	9
19	302.4304	302	9
18	299.9377	300	9
17	297.348	297	9
16	294.655	295	9
15	291.8427	292	10
14	288.8916	289	10
13	285.7817	286	10
12	282.4849	282	11
11	278.9686	279	11
10	275.1874	275	11
9	271.0898	275	
8	266.5953	275	
7	261.5933	275	
6	255.9329	275	
5	249.3684	275	
4	241.4897	275	
3	231.5314	275	
2	217.7727	275	
1	194.7402	275	
0	171.7079	275	

**Appendix 3.B: Scoring Tables for Operational and Audio CD Conversions  
March – 2004**

**Table 3.B.1: Operational and Audio CD Conversion – ELA March 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM	Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450		44	333.9864	334	9
89	505.7707	450		43	332.1596	332	9
88	485.5542	450		42	330.3355	330	9
87	470.9048	450		41	328.5123	329	9
86	460.3805	450		40	326.6897	327	9
85	451.9462	450	21	39	324.8643	325	9
84	444.7841	445	17	38	323.0351	323	9
83	438.4952	438	17	37	321.1987	321	9
82	432.8511	433	17	36	319.3574	319	9
81	427.7177	428	15	35	317.5016	318	9
80	422.9950	423	14	34	315.6368	316	9
79	418.6171	419	14	33	313.7533	314	9
78	414.5340	415	13	32	311.8535	312	9
77	410.7057	411	13	31	309.9319	310	9
76	407.1016	407	12	30	307.9861	308	9
75	403.6926	404	12	29	306.0126	306	9
74	400.4482	400	12	28	304.0089	304	10
73	397.3490	397	11	27	301.9719	302	10
72	394.3782	394	11	26	299.8966	300	10
71	391.5221	392	11	25	297.7773	298	10
70	388.7698	389	10	24	295.6118	296	10
69	386.1086	386	10	23	293.3925	293	10
68	383.5291	384	10	22	291.1159	291	10
67	381.0216	381	10	21	288.7746	289	11
66	378.5829	379	10	20	286.3550	286	11
65	376.2039	376	10	19	283.8540	284	11
64	373.8799	374	9	18	281.2607	281	11
63	371.6052	372	9	17	278.5675	279	12
62	369.3780	369	9	16	275.7575	276	12
61	367.1924	367	9	15	272.8179	275	12
60	365.0448	365	9	14	269.7289	275	
59	362.9322	363	9	13	266.4660	275	
58	360.8523	361	9	12	263.0023	275	
57	358.8036	359	9	11	259.3015	275	
56	356.7816	357	9	10	255.3226	275	
55	354.7849	355	9	9	251.0046	275	
54	352.8110	353	9	8	246.2634	275	
53	350.8587	351	9	7	240.9882	275	
52	348.9270	349	9	6	235.0148	275	
51	347.0136	347	9	5	228.0875	275	
50	345.1172	345	9	4	219.7784	275	
49	343.2352	343	9	3	209.2723	275	
48	341.3664	341	9	2	194.7368	275	
47	339.5078	340	9	1	171.0927	275	

46	337.6603	338	9
45	335.8199	336	9

0	147.2238	275	
---	----------	-----	--

**Table 3.B.2: Operational and Audio CD Conversion – Mathematics March 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	
79	496.6641	450	
78	476.1014	450	
77	461.6918	450	
76	451.4598	450	17
75	443.4198	443	16
74	436.7510	437	14
73	431.0128	431	13
72	425.9530	426	12
71	421.4086	421	12
70	417.2679	417	11
69	413.4533	413	11
68	409.9116	410	10
67	406.5985	407	10
66	403.4722	403	10
65	400.5053	401	10
64	397.6748	398	9
63	394.9623	395	9
62	392.3554	392	9
61	389.8446	390	9
60	387.4144	387	9
59	385.0574	385	9
58	382.7650	383	8
57	380.5314	381	8
56	378.3492	378	8
55	376.2134	376	8
54	374.1197	374	8
53	372.0625	372	8
52	370.0382	370	8
51	368.0432	368	8
50	366.0726	366	8
49	364.1271	364	8
48	362.2010	362	8
47	360.2905	360	8
46	358.3947	358	8
45	356.5102	357	8
44	354.6357	355	8
43	352.7676	353	8
42	350.9034	351	8
41	349.0425	349	8
40	347.1815	347	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	345.3177	345	8
38	343.4505	343	8
37	341.5766	342	8
36	339.6947	340	8
35	337.8002	338	8
34	335.8932	336	8
33	333.9698	334	8
32	332.0307	332	8
31	330.0685	330	8
30	328.0835	328	8
29	326.0709	326	8
28	324.0261	324	8
27	321.9509	322	8
26	319.8358	320	8
25	317.6780	318	8
24	315.4753	315	8
23	313.2196	313	8
22	310.9080	311	9
21	308.5343	309	9
20	306.0910	306	9
19	303.5674	304	9
18	300.9568	301	9
17	298.2474	298	9
16	295.4266	295	9
15	292.4809	292	10
14	289.3906	289	10
13	286.1346	286	10
12	282.6854	283	11
11	279.0099	279	11
10	275.0521	275	11
9	270.7637	275	
8	266.0654	275	
7	260.8451	275	
6	254.9478	275	
5	248.1210	275	
4	239.9406	275	
3	229.6026	275	
2	215.2287	275	
1	193.2560	275	
0	171.7079	275	

**Appendix 3.C: Scoring Tables for Operational and Audio CD Conversions—May 2004**

**Table 3.C.1: Operational and Audio CD Conversion – ELA May 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM	Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5854	450		44	328.3331	328	8
89	495.5762	450		43	326.5791	327	8
88	471.1133	450		42	324.8281	325	9
87	456.8118	450	26	41	323.0810	323	9
86	446.7395	447	23	40	321.3360	321	9
85	438.7300	439	21	39	319.5968	320	9
84	431.9539	432	17	38	317.8573	318	9
83	426.0464	426	17	37	316.1226	316	9
82	420.7700	421	17	36	314.3879	314	9
81	415.9829	416	14	35	312.6554	313	9
80	411.5861	412	14	34	310.9229	311	9
79	407.5085	408	13	33	309.1901	309	9
78	403.7040	404	13	32	307.4546	307	9
77	400.1360	400	12	31	305.7156	306	9
76	396.7685	397	12	30	303.9711	304	9
75	393.5771	394	11	29	302.2194	302	9
74	390.5460	391	11	28	300.4576	300	9
73	387.6540	388	11	27	298.6800	299	10
72	384.8828	385	11	26	296.8847	297	10
71	382.2207	382	10	25	295.0687	295	10
70	379.6544	380	10	24	293.2220	293	10
69	377.1803	377	10	23	291.3386	291	10
68	374.7811	375	10	22	289.4157	289	10
67	372.4548	372	10	21	287.4434	287	10
66	370.1920	370	9	20	285.4139	285	11
65	367.9896	368	9	19	283.3173	283	11
64	365.8360	366	9	18	281.1429	281	11
63	363.7311	364	9	17	278.8801	279	11
62	361.6659	362	9	16	276.5094	277	11
61	359.6412	360	9	15	274.0000	275	12
60	357.6500	358	9	14	271.3439	275	
59	355.6881	356	9	13	268.5126	275	
58	353.7543	354	9	12	265.4769	275	
57	351.8438	352	9	11	262.1962	275	
56	349.9547	350	9	10	258.6233	275	
55	348.0881	348	9	9	254.6975	275	
54	346.2369	346	9	8	250.3321	275	
53	344.4038	344	8	7	245.4076	275	
52	342.5816	343	8	6	239.7558	275	
51	340.7732	341	8	5	233.1140	275	
50	338.9719	339	8	4	225.0433	275	
49	337.1833	337	8	3	214.7114	275	
48	335.3990	335	8	2	200.2429	275	
47	333.6238	334	8	1	175.6309	275	
46	331.8558	332	8	0	147.2238	275	
45	330.0923	330	8				

**Table 3.C.2: Operational and Audio CD Conversion – Mathematics May 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	
79	494.5100	450	
78	471.8718	450	
77	457.9717	450	19
76	447.9354	448	17
75	440.0040	440	16
74	433.4013	433	14
73	427.7116	428	13
72	422.6887	423	12
71	418.1762	418	12
70	414.0599	414	11
69	410.2704	410	11
68	406.7444	407	10
67	403.4424	403	10
66	400.3296	400	10
65	397.3768	397	10
64	394.5613	395	9
63	391.8684	392	9
62	389.2853	389	9
61	386.7938	387	9
60	384.3870	384	9
59	382.0529	382	9
58	379.7873	380	8
57	377.5787	378	8
56	375.4254	375	8
55	373.3191	373	8
54	371.2559	371	8
53	369.2324	369	8
52	367.2392	367	8
51	365.2799	365	8
50	363.3508	363	8
49	361.4432	361	8
48	359.5584	360	8
47	357.6916	358	8
46	355.8418	356	8
45	354.0064	354	8
44	352.1821	352	8
43	350.3658	350	8
42	348.5594	349	8
41	346.7560	347	8
40	344.9560	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	343.1574	343	8
38	341.3574	341	8
37	339.5550	340	8
36	337.7453	338	8
35	335.9292	336	8
34	334.1027	334	8
33	332.2662	332	8
32	330.4140	330	8
31	328.5452	329	8
30	326.6565	327	8
29	324.7436	325	8
28	322.8073	323	8
27	320.8421	321	8
26	318.8429	319	8
25	316.8081	317	8
24	314.7318	315	8
23	312.6097	313	8
22	310.4373	310	9
21	308.2087	308	9
20	305.9167	306	9
19	303.5509	304	9
18	301.1048	301	9
17	298.5672	299	9
16	295.9247	296	9
15	293.1647	293	10
14	290.2679	290	10
13	287.2122	287	10
12	283.9715	284	11
11	280.5113	281	11
10	276.7874	277	11
9	272.7447	275	12
8	268.3059	275	
7	263.3585	275	
6	257.7470	275	
5	251.2254	275	
4	243.3779	275	
3	233.4264	275	
2	219.6205	275	
1	196.3774	275	
0	171.7079	275	

### Appendix 3.D: Scoring Tables for Braille Conversions

**Table 3.D.1: Braille Conversion – ELA February, March, and May 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM	Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
90	525.5856	450		44	331.5973	332	9
89	502.7214	450		43	329.8201	330	9
88	479.8572	450		42	328.0497	328	9
87	466.1185	450		41	326.2860	326	9
86	455.9847	450	23	40	324.5256	325	9
85	447.7529	448	21	39	322.7687	323	9
84	440.7116	441	17	38	321.0117	321	9
83	434.4930	434	17	37	319.2581	319	9
82	428.8950	429	17	36	317.4978	317	9
81	423.7860	424	15	35	315.7374	316	9
80	419.0816	419	14	34	313.9670	314	9
79	414.7212	415	14	33	312.1897	312	9
78	410.6542	411	13	32	310.3991	310	9
77	406.8402	407	13	31	308.5949	309	9
76	403.2554	403	12	30	306.7705	307	9
75	399.8662	400	12	29	304.9258	305	9
74	396.6491	397	12	28	303.0542	303	10
73	393.5836	394	11	27	301.1556	301	10
72	390.6531	391	11	26	299.2199	299	10
71	387.8440	388	11	25	297.2437	297	10
70	385.1394	385	10	24	295.2237	295	10
69	382.5292	383	10	23	293.1497	293	10
68	380.0000	380	10	22	291.0184	291	10
67	377.5517	378	10	21	288.8231	289	11
66	375.1675	375	10	20	286.5501	287	11
65	372.8440	373	10	19	284.1929	284	11
64	370.5744	371	9	18	281.7379	282	11
63	368.3588	368	9	17	279.1783	279	12
62	366.1837	366	9	16	276.4939	276	12
61	364.0524	364	9	15	273.6713	275	12
60	361.9548	362	9	14	270.6902	275	
59	359.8943	360	9	13	267.5236	275	
58	357.8642	358	9	12	264.1446	275	
57	355.8611	356	9	11	260.5126	275	
56	353.8849	354	9	10	256.5872	275	
55	351.9323	352	9	9	252.3078	275	
54	350.0000	350	9	8	247.5866	275	
53	348.0913	348	9	7	242.3123	275	
52	346.1994	346	9	6	236.3197	275	
51	344.3278	344	9	5	229.3525	275	
50	342.4696	342	9	4	220.9858	275	
49	340.6284	341	9	3	210.4103	275	
48	338.7972	339	9	2	195.8318	275	
47	336.9829	337	9	1	171.5277	275	
46	335.1754	335	9	0	147.2240	275	
45	333.3813	333	9				

**Table 3.D.2: Braille Conversion – Mathematics February, March, and May 2004**

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
80	517.0695	450	
79	494.0284	450	
78	470.9871	450	
77	457.2381	450	19
76	447.2799	447	17
75	439.4034	439	16
74	432.843	433	14
73	427.1891	427	13
72	422.1971	422	12
71	417.712	418	12
70	413.6199	414	11
69	409.8524	410	11
68	406.3459	406	10
67	403.0621	403	10
66	399.9655	400	10
65	397.0271	397	10
64	394.2243	394	9
63	391.5435	392	9
62	388.9706	389	9
61	386.4876	386	9
60	384.0885	384	9
59	381.7604	382	9
58	379.5001	380	8
57	377.2947	377	8
56	375.1442	375	8
55	373.0389	373	8
54	370.9756	371	8
53	368.951	369	8
52	366.9548	367	8
51	364.9918	365	8
50	363.058	363	8
49	361.1432	361	8
48	359.2511	359	8
47	357.375	357	8
46	355.5151	356	8
45	353.6681	354	8
44	351.8306	352	8
43	350	350	8
42	348.1788	348	8
41	346.3577	346	8
40	344.5398	345	8

Raw Score	Unrounded Scale Score	Rounded Scale Score	CSEM
39	342.7218	343	8
38	340.9007	341	8
37	339.0763	339	8
36	337.2422	337	8
35	335.4017	335	8
34	333.5483	334	8
33	331.6851	332	8
32	329.8026	330	8
31	327.904	328	8
30	325.9828	326	8
29	324.0359	324	8
28	322.066	322	8
27	320.064	320	8
26	318.0265	318	8
25	315.9535	316	8
24	313.8354	314	8
23	311.6718	312	8
22	309.4567	309	9
21	307.1835	307	9
20	304.8457	305	9
19	302.4304	302	9
18	299.9377	300	9
17	297.348	297	9
16	294.655	295	9
15	291.8427	292	10
14	288.8916	289	10
13	285.7817	286	10
12	282.4849	282	11
11	278.9686	279	11
10	275.1874	275	11
9	271.0898	275	
8	266.5953	275	
7	261.5933	275	
6	255.9329	275	
5	249.3684	275	
4	241.4897	275	
3	231.5314	275	
2	217.7727	275	
1	194.7402	275	
0	171.7079	275	

### Appendix 3.E: 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

$$S = \text{rawscore} = \sum_{j=1}^{n_1} w_{j1} U_j + \sum_{j=1}^{n_2} w_{j2} Y_j$$

$\hat{\theta}$  = MLE of  $\theta$

$$\text{SEM}(\hat{\theta}) \approx [I(\theta, S)]^{-1/2}$$

$$I(\theta, S) = \frac{\left[ \frac{d E(S | \theta)}{d\theta} \right]^2}{\sigma^2(S | \theta)} \quad \text{see Lord (1980), pp. 67 and 73.}$$

$$E(S | \theta) = \sum_{j=1}^{n_1} w_{j1} P_{j1}(\theta) + \sum_{j=1}^{n_2} w_{j2} \sum_{k=1}^{m_j} (k-1) P_{j2k}(\theta)$$

$$\frac{dE(S | \theta)}{d\theta} = \sum_{j=1}^{n_1} w_{j1} P'_{j1}(\theta) + \sum_{j=1}^{n_2} w_{j2} \sum_{k=1}^{m_j} (k-1) P'_{j2k}(\theta)$$

where

$$P'_{j1}(\theta) = P_{j1}(\theta)(1 - P_{j1}(\theta))$$

$$P'_{j2k}(\theta) = P_{j2k}(\theta) \left[ k - \sum_{r=1}^{m_j} r P_{j2r}(\theta) \right]$$

$$\sigma^2(S | \theta) = \sum_{j=1}^{n_1} (w_{j1})^2 P_{j1}(\theta)(1 - P_{j1}(\theta)) + \sum_{j=1}^{n_2} (w_{j2})^2 \sigma^2(y_j | \theta)$$

$$\sigma^2(Y_j | \theta) = \left[ \sum_{k=1}^{m_j} (k-1)^2 P_{j2k}(\theta) \right] - \left[ \sum_{k=1}^{m_j} (k-1) P_{j2k}(\theta) \right]^2$$

Note also Lord (1980), Eq. (5-23) and Eq. (6-6) for transforming the standard error to other metrics.

## Chapter 4: Summary of Item-Level Analyses

This section summarizes item-level statistics obtained for the CAHSEE for classical item analyses, differential item functioning (DIF), and IRT calibrations. The numbers of items included in the analyses are listed in Table 4.1. As described earlier, items contributing to operational forms are analyzed without the field-test items for the creation of scoring tables. Field-test items are analyzed separately. Therefore, data are summarized separately for operational items and field-test items for each administration. A summary of the items included in the 2004 administrations is shown in Table 4.1 below.

**Table 4.1: Summary of Items and Forms by Administration**

Administration	Operational Items	Field-Test Forms	Number of Field-Test Items*
<b>February 2004</b>			
English-Language Arts	72 MC, 1 CR	25	174
Mathematics	80 MC	25	298
<b>March 2004</b>			
English-Language Arts	72 MC, 1 CR	94	655
Mathematics	80 MC	95	1139
<b>May 2004</b>			
English-Language Arts	72 MC, 1 CR	25	172
Mathematics	80 MC	25	300

\*Note. In some cases the same field-test items may have appeared in more than one form.

### Classical Item Analyses

For each administration, classical item analyses are completed prior to DIF and item calibration, scaling and equating. These analyses involve computing, for every item in each form, a set of statistics based on classical test theory. Each statistic is designed to provide some key information about the quality of each item from an empirical perspective. 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 (“r-biserial”)<sup>2</sup>:

This statistic is measured by the polyserial correlation between the item score and the test criterion score and describes the relationship between performance on the specific item and performance on the entire form. The higher the value, the better the task of separating the examinees. Items with negative correlations can indicate serious problems with the item content (e.g., multiple correct answers or unusually complex content), or can indicate that students have not been taught the content. For ELA, the test criterion score was the number-correct score on the MC items, plus the weighted CR item score. For mathematics, the test criterion score was the number-correct score.

The proportion of students choosing each response option:

These statistics indicate the percent of examinees that select each of the available answer options and the percent of examinees that omitted the item.

Distracter analyses for MC items.

The Genasys system provides graphical displays of the each option, which are reviewed.

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, we would expect that if students have an adequate amount of testing time, 95% of students should attempt to answer each question. When a pattern of omit percentages exceeds 5% for a series of items at the end of a timed section, this may indicate that there was insufficient time for students to complete all items. Alternatively, if the omit percentage is greater than 5% for a single item, this could be an indication of an item/test layout problem. For example, students might accidentally skip an item that follows a lengthy stem.

Although all of these analyses are completed for the CAHSEE, results are summarized here for the item discrimination value (r-biserial) for operational and field-test items. The remaining statistics are included in summary files that are used to populate the item bank.

### **Differential Item Functioning (DIF)**

Following the classical item analyses, DIF studies were completed. 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. DIF statistics are used to identify those items that identifiable groups of students (e.g. females, African Americans,

---

<sup>2</sup> The estimated polyserial correlation between scores on the item and on the criterion is computed by the formula:

$$r_{polyreg} = \frac{\beta_i \sigma_x}{\sqrt{\beta_i^2 \sigma_x^2 + 1}}$$

where the  $\beta_i$  are a series of parameters estimated by maximum likelihood from the item analysis data (Dragow, 1988; Lewis & Thayer, 1996).

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 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 are required to determine the source and meaning of any differences that are seen.

ETS used two DIF detection methods: the Mantel-Haenszel and standardization approaches. As part of the Mantel-Haenszel procedure, the statistic described by Holland & Thayer (1988), known as MH D-DIF, was used<sup>3</sup>. This statistic is expressed as the differences between the focal and reference group 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 12 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 significantly 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 4.2). Category A contains negligible DIF, Category B items

---

<sup>3</sup> 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, 1985):

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

exhibit slight or moderate DIF, and Category C items 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 that, conditional on the matching variable, the reference group has lower mean item score than the focal group. For constructed-response items the MH D-DIF is not calculated, but analogous flagged rules based on the chi-square statistic have been developed resulting in classification into A, B, or C DIF categories.

**Table 4.2: DIF Categories**

DIF Category	Definition
A (negligible)	MH D-DIF not significantly different from zero, or has an absolute value less than one.
B (slight to moderate)	MH D-DIF is significantly different from zero, and is either a) less than 1.5, or b) not significantly different from one.
C (moderate to large)	MH D-DIF is significantly different from one, and has an absolute value greater than 1.5.

All DIF analyses were performed according to the procedures specified in the document “CAHSEE DIF Procedures” dated February 22, 2002. According to those procedures, DIF analyses were performed for English Learners for mathematics but not for ELA. Operational items flagged for negative C (C-)DIF are reviewed by an expert committee DIF review panel consisting of CAHSEE item development staff, CDE staff responsible for the CAHSEE, external educators identified by 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 produce final test scores.

Five ELA and four operational mathematics items were classified as C- items based on the final DIF analyses in February 2004. In March 2004 two ELA items and four mathematics items were flagged for C- DIF. One ELA item and five mathematics items in May 2004 were flagged for C- DIF. These items were reviewed by a panel of content experts, as specified in the CAHSEE DIF procedures, and found to measure essential test content. One of the items was flagged for DIF in both the February and March 2004 administrations. Because the item passed the review process by the CAHSEE DIF Review Panel the item was not reviewed a second time. All items were deemed appropriate for scoring for all three administrations.

For the field-test in the February 2004 administration, 11 ELA items and 19 mathematics items were classified in the C- DIF category. For the March 2004 administration, 37 ELA field-test items and 44 mathematics field-test items were classified as C- DIF. For the May 2004 administration, no ELA field-test items and only two mathematics field-test items were classified as having C- DIF. These items will not be used in future operational administrations without revision and additional field-testing unless their inclusion is essential to meeting test specifications. In this case, the items will be reviewed by another bias and sensitivity committee prior to use.

## IRT Analyses

The item calibration studies (see Chapter 3 for details) estimate item difficulty, as measured by the Rasch difficulty parameter estimate (b-value) for both the operational items and the field-test items. In addition, statistics describing the fit of the Rasch model to the data, which is reported in letter categories of A, B, C, D, and F (IRT flag) are also produced. A description and examples of this model-data fit rating scheme are provided in Appendix 4.A. All items receiving a rating of “F” were also rated as “do not use.” Table 4.3 summarizes the average b-values across each administration for the ELA and mathematics operational items. This table also summarizes the IRT fit flags across each administration.

**Table 4.3: Summary of IRT Model Data Fit Distribution and C-DIF Items Across Administrations: Operational Items**

Subject/ Admin.	Number Items	Rasch Item Difficulties			Fit = A		Fit = B		Fit = C		Fit = D		Fit = F	
		Median B	Mean B	SD B	#	%	#	%	#	%	#	%	#	%
ELA/ February	73	-0.045	-0.006	0.800	33	45.2	23	31.5	13	17.8	4	5.5	0	0.0
Math/ February	80	-0.190	-0.240	0.753	26	32.5	31	38.8	22	27.5	1	1.3	0	0.0
ELA/ March	73	-0.016	0.049	0.814	41	56.6	7	9.6	19	26.0	5	6.9	0	0.0
Math/ March	80	-0.040	-0.175	0.868	41	51.3	15	18.8	17	21.3	4	5.0	3	3.8
ELA/ May	73	-0.016	-0.035	0.642	12	16.4	20	27.4	38	52.1	2	2.7	0	0.0
Math/ May	80	-0.256	-0.216	0.716	20	25.0	13	16.3	43	53.8	2	2.5	2	2.5

Table 4.4 summarizes the average b-values across each administration for the ELA & mathematics field-test items. These tables also summarize the IRT fit flags across each administration. On average the field-test items for both content areas in each administration are more difficult than the operational items as measured by both the mean and median b-values. Across administrations, the percentage of field-test items receiving an IRT fit value of F for ELA ranges from 4.6% to 13.9% and for mathematics ranges from 10.5% to 28.3%. For both ELA and mathematics, the highest percentage of field-test items receiving a rating of F were for the May 2004 administration.

**Table 4.4: Summary of IRT Model Data Fit Distribution and C-DIF Items Across Administrations: Field-Test Items**

Subject/ Admin.	Number Items	Rasch Item Difficulties			Fit = A		Fit = B		Fit = C		Fit = D		Fit = F	
		Median B	Mean B	SD B	#	%	#	%	#	%	#	%	#	%
ELA/ February	174	0.490	0.577	0.921	42	24.1	41	23.6	76	43.7	7	4.0	8	4.6
Math/ February	298	0.596	0.569	1.010	46	15.4	52	17.5	143	48.0	17	5.7	40	13.4
ELA/ March	655	0.617	0.715	0.966	158	24.1	158	24.1	258	39.9	34	5.9	47	7.8
Math/ March	1139	0.617	0.625	1.030	188	16.5	318	27.9	464	40.7	50	4.4	119	10.5
ELA/ May	172	0.542	0.581	0.746	6	3.5	17	9.9	105	61.1	20	11.6	24	13.9
Math/ May	300	0.273	0.309	0.757	6	2.0	22	7.3	17	55.3	21	7.00	85	28.3

The number and percentages of field-test ELA and mathematics items that received model-data fit ratings of A, B, C or D and were recommended for use in future operational tests are summarized in table 4.5.

**Table 4.5: Items Recommended for Future Forms**

Administration Date	Subject	Number of Items (%)
February 2004	ELA	166 (95%)
	Mathematics	258 (87%)
March 2004	ELA	608 (93%)
	Mathematics	1020 (90%)
May 2004	ELA	148 (86%)
	Mathematics	215 (72%)

### Summary of Item-Level Statistics

To simplify the presentation of this data, all tables for this section are located in Appendix 4.B to Appendix 4.D. Appendix 4.B contains all of the output related to the February 2004 administration, Appendix 4.C contains all of the output from the March 2004 administration and

Appendix 4.D contains all of the output related to the May 2004 administration. The same table numbering system applies from 4.B through 4.D and is listed below (see table 4.6).

**Table 4.6: Listing of Summary Tables**

Table <sup>1</sup>	Content	Label
4.x.1	Significant Negative DIF: Operational Items	Listing of Operational Items Exhibiting Significant Negative DIF
4.x.2	Significant Negative DIF: Field-Test Items	Listing of Field-Test Items Exhibiting Significant Negative DIF
4.x.3	Summary of the ELA CR item, including the IRT b-value and step parameters, r-biserial correlation, and DIF results.	Listing of CR Item Response Item Statistics – ELA
4.x.4	Univariate Statistics for ELA Operational items: IRT b-value and r-Biserial for all operational items and by strand	Summary Univariate Operational Item Statistics – ELA
4.x.5	Univariate Statistics for mathematics Operational items: IRT b-value and r-Biserial for all operational items and by strand	Summary Univariate Operational Item Statistics – Mathematics
4.x.6	Univariate Statistics for ELA Field-Test items: IRT b-value and r-Biserial for all operational items and by strand	Summary Univariate Field-Test Item Statistics – ELA
4.x.7	Univariate Statistics for mathematics Field-Test items: IRT b-value and r-Biserial for all operational items and by strand	Summary Univariate Field-Test Item Statistics – Mathematics
4.x.8	IRT model data fit ELA operational items	IRT Model Data Fit Distribution of Operational Items – ELA
4.x.9	IRT model data fit mathematics operational items	IRT Model Data Fit Distribution of Operational Items – Mathematics
4.x.10	IRT model data fit ELA field-test items	IRT Model Data Fit Distribution of Field-Test Items – ELA
4.x.11	IRT model data fit mathematics field-test items	IRT Model Data Fit Distribution of Field-Test Items – Mathematics
4.x.12	Operational DIF classifications – ELA	Distribution of Operational DIF Classifications - ELA
4.x.13	Operational DIF classifications – mathematics	Distribution of Operational DIF Classifications – Mathematics
4.x.14	Field-test DIF classifications – ELA	Distribution of Field-Test DIF Classifications - ELA
4.x.15	Field-test DIF classifications – mathematics	Distribution of Field-Test DIF Classifications – Mathematics

Note. <sup>1</sup>x = Administration, where B=February, C=March, D=May.

## Appendix 4.A: CAHSEE Field-Test Item Review – Description and Examples of Classification Categories

The categories used by ETS statisticians to classify newly field-tested 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 that are independently assigned by two statisticians. A third rater resolves discrepant ratings. 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, starting from the left-hand side upper corner. These six items illustrate the span of the rating categories. The item number in the calibration and 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 (ELA  $\theta = \text{approx. } 0.43$ , Mathematics  $\theta = \text{approx. } 0.20$ )
- 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
- This category often applies to items that appear to be functioning well, but that are not well fit by the Rasch model

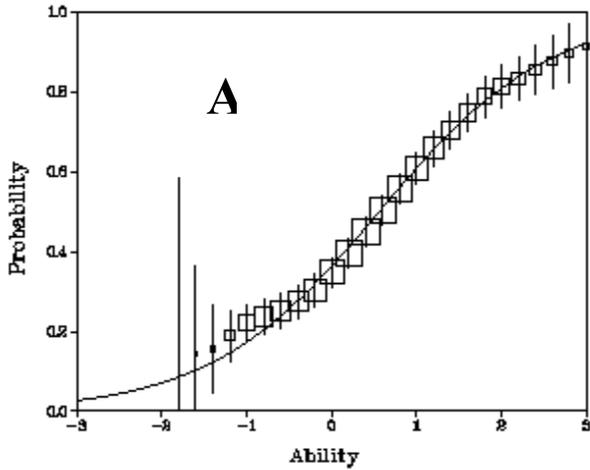
### **Flag D** (Item 94, FM002620)

- 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
- In addition to inadequate Rasch model-data fit, the item appears to be poorly suited for use (may be poorly discriminating overall or too difficult to provide any discrimination at the passing score)

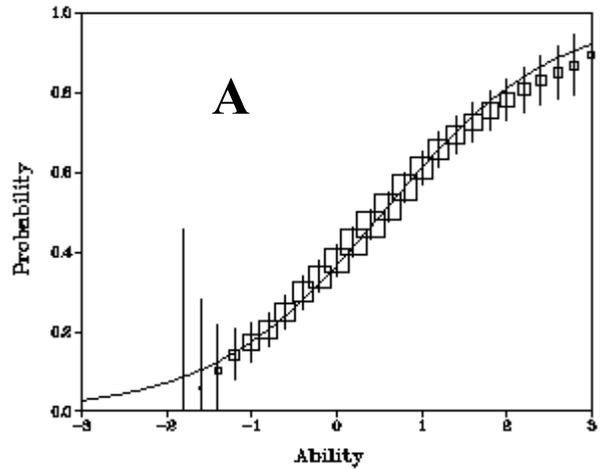
### **Flag F** (Item 98, FM002768)

- 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

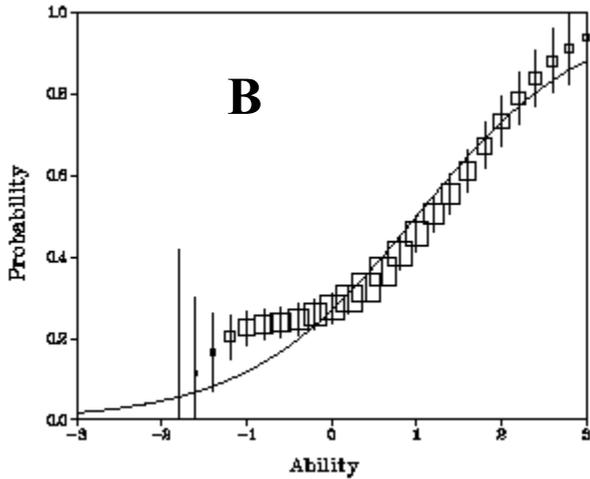
## Samples of CAHSEE Item-Fit Rating Categories



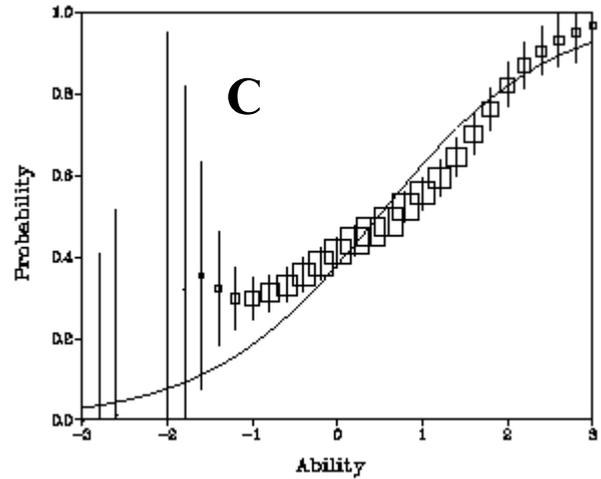
#93 FM002619901 4 Choice P+=0.494  
a=0.588 F,b= 0.561 ,c= 0.000 F ,CHI= 11.85



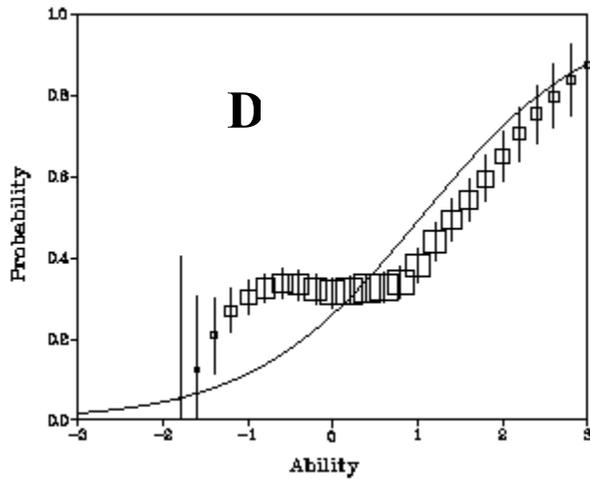
#95 FM002640901 4 Choice P+=0.496  
a=0.588 F,b= 0.550 ,c= 0.000 F ,CHI= 10.50



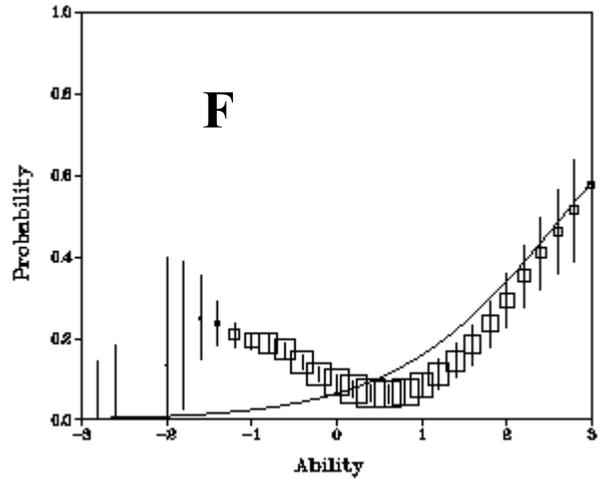
#96 FM002641901 4 Choice P+=0.402  
a=0.588 F,b= 1.003 ,c= 0.000 F ,CHI=116.63



#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 4.B: Item and DIF Statistics for February 2004

**Table 4.B.1: Listing of Operational Items Exhibiting Significant Negative DIF February 2004**

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	FM003994	L170A003	0	1	A	A	C-	A	B-	B-	A	A	<i>N/A for ELA</i>
ELA	FM004003	L170A014	0	5	B-	A	B-	C-	B-	B-	B-	B-	
ELA	FM005550	L238D005	0	33	C-	A	A	B-	B-	A	A	A	
ELA	VC016593	L00SA052	0	69	A	A	C-	B-	C-	C-	A	A	
ELA	FM004936	L10SA290	0	76	A	A	C-	A	B-	C-	A	A	
Math	FM006754	M13888	0	16	A	A	C-	B-	B-	C-	B-	B-	‡
Math	FM006611	M13053	0	26	C-	A	A	A	A	A	A	A	‡
Math	FM006937	M21748	0	27	A	A	C-	B-	A	C-	A	A	‡
Math	FM006355	M11861	0	30	A	A	C-	A	B-	C-	A	A	‡

‡ Minimum sample size for DIF analysis not met.

**Table 4.B.2. Listing of Field-Test Items Exhibiting Significant Negative 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	VC026114	L140A002	8	46	A	‡	B-	‡	‡	B-	C-	A	<i>N/A for</i>
ELA	VC026123	L140A017	8	47	A	‡	C-	‡	‡	C-	B-	B-	
ELA	VC026173	L174A002	11	46	A	‡	C-	‡	‡	B-	B-	A	
ELA	VC026158	L166B002	13	46	C-	‡	A	‡	‡	A	A	A	
ELA	VC017337	L10SA310	13	75	A	‡	C-	‡	‡	C-	A	B-	
ELA	VC026161	L166B005	14	47	A	‡	C-	‡	‡	C-	B-	A	
ELA	VC026135	L160A003	18	47	B-	‡	C-	‡	‡	C-	C-	B-	
ELA	VC026142	L160A011	18	50	A	‡	C-	‡	‡	C-	C-	B-	
ELA	VC017516	L10SA537	18	75	A	‡	C-	‡	‡	B-	A	B-	
ELA	VC026195	L176A001	20	46	A	‡	C-	‡	‡	C-	B-	B-	
ELA	VC019890	L271C003	22	46	B-	‡	C-	‡	‡	B-	A	A	
Math	FM007771	M21015	1	60	A	‡	C-	‡	‡	B-	A	A	‡
Math	FM007773	M21143	2	18	A	‡	A	‡	‡	A	B-	C-	‡
Math	FM007803	M23180	2	35	A	‡	C-	‡	‡	C-	A	A	‡
Math	FM007752	M20502	3	20	A	‡	B-	‡	‡	C-	A	A	‡
Math	VC032221	M25193	3	35	A	‡	C-	‡	‡	C-	B-	B-	‡
Math	FM007793	M21963	4	35	A	‡	C-	‡	‡	B-	B-	B-	‡
Math	VC024692	M21609	7	19	A	‡	C-	‡	‡	C-	A	A	‡
Math	VC022610	M12163	8	35	A	‡	C-	‡	‡	C-	A	A	‡
Math	VC023455	M20068	9	20	C-	‡	A	‡	‡	A	A	A	‡
Math	VC024099	M20864	13	21	A	‡	C-	‡	‡	C-	A	A	‡
Math	VC025081	M22107	14	36	A	‡	C-	‡	‡	C-	C-	A	‡
Math	VC024694	M21613	16	18	A	‡	C-	‡	‡	B-	B-	A	‡
Math	VC024765	M21696	16	20	B-	‡	C-	‡	‡	C-	B-	A	‡
Math	VC023665	M20332	18	19	C-	‡	A	‡	‡	B-	A	A	‡
Math	VC025172	M23083	20	18	A	‡	C-	‡	‡	C-	B-	B-	‡
Math	VC024801	M21751	20	21	A	‡	C-	‡	‡	C-	B-	A	‡
Math	VC026301	M10759	20	36	B-	‡	C-	‡	‡	B-	B-	A	‡
Math	VC022810	M12584	23	18	A	‡	A	‡	‡	A	C-	A	‡
Math	VC024975	M21966	23	21	A	‡	C-	‡	‡	B-	A	B-	‡

‡ Minimum sample size for DIF analysis not met.

**Table 4.B.3: Listing of CR Item Statistics - ELA**

Accession number	<b>FM004509</b>
CAHSEE ID	<b>L1000054</b>
Polyserial correlation	0.81
IRT b Value	0.65
Step category 1*	2.007
Step category 2	1.171
Step category 3	2.184
Step category 4	-0.382
Step category 5	-0.296
Step category 6	-2.349
Step category 7	-2.334
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	B+
DIF category, White-Filipino	B+
DIF category, White-Combined Asian	B+
DIF category, White-Hispanic	A
DIF category, White-African American	A
Least favorable DIF category among all focal groups**	B

Notes. \* Step categories refer to the parameters describing each item category in the CR item calibrations.

\*\* This refers to the most disadvantageous DIF category found among all focal groups for which a comparison was made. Positive DIF categories disadvantage the reference group and negative DIF categories disadvantage the focal group.

**Table 4.B.4: Summary Univariate Operational Item Statistics - ELA**

Content Area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	-0.01	0.80	-2.53	1.56	0.53	0.10	0.26	0.81
RC	18	0.01	0.52	-0.84	1.23	0.52	0.09	0.33	0.67
RL	20	-0.23	1.11	-2.53	1.56	0.51	0.11	0.28	0.68
RW	7	-0.47	0.93	-2.19	0.86	0.51	0.14	0.26	0.64
WA	1	0.65	.	0.65	0.65	0.81	.	0.81	0.81
WC	15	0.02	0.56	-0.90	1.27	0.54	0.07	0.43	0.65
WS	12	0.51	0.40	-0.11	1.37	0.55	0.05	0.45	0.61

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WA=Writing Applications, WC=Writing Conventions, WS=Writing Strategies.

**Table 4.B.5: Summary Univariate Operational Item Statistics - Mathematics**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.24	0.75	-1.97	1.50	0.55	0.09	0.30	0.69
A1	12	0.26	0.49	-0.67	1.19	0.55	0.08	0.41	0.69
AF	20	-0.33	0.77	-1.50	1.36	0.56	0.09	0.37	0.69
MG	18	-0.32	0.83	-1.72	1.50	0.56	0.09	0.30	0.69
MR	8	-0.28	0.47	-1.05	0.31	0.56	0.07	0.43	0.64
NS	17	-0.28	0.68	-1.44	1.36	0.52	0.09	0.30	0.65
PS	13	-0.40	0.82	-1.97	1.13	0.55	0.10	0.34	0.69

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.B.6: Summary Univariate Field-Test Item Statistics - ELA**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	174	0.58	0.92	-1.40	3.12	0.49	0.14	-0.06	0.68
RC	56	0.56	0.85	-0.79	3.12	0.50	0.12	0.05	0.68
RL	49	0.82	0.93	-1.06	2.98	0.46	0.18	-0.06	0.66
RW	44	0.37	0.88	-1.40	2.50	0.51	0.11	0.23	0.68
WC	25	0.51	1.07	-1.27	2.73	0.48	0.14	0.06	0.64

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WC=Writing Conventions, WS=Writing Strategies.

**Table 4.B.7: Summary Univariate Field-Test Item Statistics - Mathematics**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	298	0.57	1.01	-2.22	4.46	0.45	0.16	-0.37	0.73
A1	45	1.02	0.65	-0.55	2.61	0.43	0.13	0.05	0.63
AF	60	0.45	0.75	-1.13	2.44	0.46	0.15	-0.01	0.69
MG	75	0.84	1.16	-1.73	4.46	0.41	0.21	-0.37	0.70
MR	22	0.24	1.01	-1.85	2.22	0.47	0.14	0.08	0.66
NS	73	0.41	1.05	-1.99	2.99	0.51	0.13	0.05	0.73
PS	45	0.09	1.02	-2.22	2.42	0.44	0.16	-0.09	0.66

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.B.8: IRT Model Data Fit Distribution of Operational Items - ELA**

<b>IRT review Category</b>	<b>Status</b>	<b>Item count</b>	<b>Percent</b>
A	Use	33	45
B	Use	23	32
C	Use	13	18
D	Review	4	5
F	Do not use	0	0
<b>Total</b>		73	100

**Table 4.B.9: IRT Model Data Fit Distribution of Operational Items - Mathematics**

<b>IRT review Category</b>	<b>Status</b>	<b>Item count</b>	<b>Percent</b>
A	Use	26	33
B	Use	31	39
C	Use	22	28
D	Review	1	1
F	Do not use	0	0
<b>Total</b>		80	100

**Table 4.B.10: IRT Model Data Fit Distribution of Field-Test Items - ELA**

<b>IRT review Category</b>	<b>Status</b>	<b>Item count</b>	<b>Percent</b>
A	Use	42	24
B	Use	41	24
C	Use	76	44
D	Review	7	4
F	Do not use	8	5
<b>Total</b>		174	100

**Table 4.B.11: IRT Model Data Fit Distribution of Field-Test Items - Mathematics**

IRT review Category	Status	Item count	Percent
A	Use	46	15
B	Use	52	17
C	Use	143	48
D	Review	17	6
F	Do not use	40	13
<b>Total</b>		298	100

**Table 4.B.12: 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		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-	1	1	0	0	3	4	1	1	1	1	2	3	0	0	0	0	5	7
B-	4	5	0	0	4	5	4	5	8	11	4	5	2	3	2	3		
A	67	92	73	100	62	85	67	92	63	86	64	88	71	97	71	97		
B+	1	1	0	0	3	4	1	1	1	1	3	4	0	0	0	0		
C+	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0		
Small N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
TOTAL	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100		

**Table 4.B.13: 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		Native English-English Learner		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
C-	1	1	0	0	3	4	0	0	0	0	3	4	0	0	0	0	0	0	4	5
B-	1	1	0	0	8	10	3	4	4	5	7	9	1	1	1	1	0	0		
A	76	95	80	100	61	76	76	95	75	94	66	83	78	98	79	99	80	100		
B+	2	3	0	0	7	9	1	1	1	1	4	5	1	1	0	0	0	0		
C+	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0		
Small N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
TOTAL	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100		

**Table 4.B.14: Distribution of Field-Test Item DIF Classifications - ELA**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	1	1	0	0	9	5	0	0	0	0	6	3	3	2	0	0	11
B-	4	2	0	0	12	7	0	0	0	0	12	7	12	7	17	10		
A	162	93	0	0	147	84	0	0	1	1	149	86	158	91	156	90		
B+	6	3	0	0	4	2	0	0	0	0	5	3	1	1	1	1		
C+	1	1	0	0	2	1	0	0	0	0	2	1	0	0	0	0		
Small N	0	0	174	100	0	0	174	100	173	99	0	0	0	0	0	0		
TOTAL	174	100	174	100	174	100	174	100	174	100	174	100	174	100	174	100		

**Table 4.B.15: Distribution of Field-Test Item DIF Classifications - Mathematics**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Native English-English Learner		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	2	1	0	0	14	5	0	0	0	0	10	3	2	1	1	0	0	0	19
B-	9	3	0	0	17	6	0	0	0	0	18	6	14	5	15	5	1	0		
A	286	96	0	0	243	82	0	0	2	1	259	87	281	94	282	95	294	99		
B+	1	0	0	0	23	8	0	0	0	0	10	3	1	0	0	0	3	1		
C+	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0		
Small N	0	0	298	100	0	0	298	100	296	99	0	0	0	0	0	0	0	0		
TOTAL	298	100	298	100	298	100	298	100	298	100	298	100	298	100	298	100	298	100		

### Appendix 4.C: Item and DIF Statistics for March 2004

**Table 4.C.1: Listing of Operational Items Exhibiting Significant Negative DIF March 2004**

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	FM005160	L159A001	0	22	A	A	B-	A	A	B-	C-	A	<i>N/A for ELA</i>
ELA	FM004751	L10SA528	0	78	A	A	C-	C-	C-	C-	A	A	
Math	FM006918	M21563	0	16	A	A	C-	A	B-	B-	A	A	A
Math	FM006296	M11393	0	23	A	A	C-	A	B-	B-	B-	A	A
Math	FM006611	M13053	0	26	C-	A	A	A	A	A	A	A	A
Math	FM004357	M03147	0	62	A	A	C-	A	A	B-	A	A	A

**Table 4.C.2: Listing of Field-Test Items Exhibiting Significant Negative 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	FM007623	L163A013	6	47	B	+	C-	+	+	C-	A	A	<i>N/A for</i>
ELA	VC026266	L239E011	13	51	A	+	C-	+	+	C-	A	B-	
ELA	VC017845	L164A002	15	51	A	+	C-	+	+	B-	B-	A	
ELA	VC017851	L164A007	16	46	C-	+	A	+	+	A	A	B-	
ELA	VC017852	L164A008	16	48	A	+	C-	+	+	B-	A	A	
ELA	VC026222	L188B012	19	47	C-	+	B-	+	+	B-	B-	A	
ELA	VC019421	L251D001	21	46	A	+	B-	+	+	C-	C-	B-	
ELA	VC019719	L265E001	23	46	A	+	B-	+	+	B-	B-	C-	
ELA	VC017349	L10SA322	23	75	A	+	A	+	+	A	C-	B-	
ELA	VC019724	L265E006	24	46	A	+	C-	+	+	C-	B-	A	
ELA	VC019727	L265E009	24	50	A	+	C-	+	+	C-	C-	C-	
ELA	VC017254	L10SA171	24	75	A	+	C-	+	+	B-	A	A	
ELA	VC019196	L244C004	25	46	A	+	C-	+	+	B-	B-	A	
ELA	VC017350	L10SA323	26	75	A	+	C-	+	+	C-	A	A	
ELA	VC019599	L256E015	27	49	A	+	B-	+	+	B-	C-	A	
ELA	VC019590	L256E006	27	51	A	+	A	+	+	A	B-	C-	
ELA	VC019586	L256E002	28	46	C-	+	B-	+	+	A	A	A	
ELA	VC017266	L10SA191	28	75	A	+	C-	+	+	C-	C-	A	
ELA	VC017268	L10SA193	29	75	A	+	C-	+	+	C-	A	C-	
ELA	VC019704	L264E005	31	46	C-	+	A	+	+	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	VC019706	L264E007	32	46	C-	+	A	+	+	A	B-	C-	
ELA	VC019708	L264E009	32	50	B-	+	C-	+	+	C-	C-	C-	
ELA	VC026294	L30SA139	35	75	A	+	C-	+	+	C-	A	A	
ELA	VC017640	L120A008	36	49	A	+	C-	+	+	C-	A	A	
ELA	VC026075	L117A007	37	48	B-	+	A	+	+	A	C-	B-	
ELA	VC026156	L165A008	40	48	A	+	B-	+	+	C-	A	A	
ELA	VC017324	L10SA284	41	75	A	+	C-	+	+	C-	A	A	
ELA	VC026283	L299C010	44	47	A	+	C-	+	+	C-	C-	B-	
ELA	VC026112	L135A008	47	50	A	+	C-	+	+	C-	A	B-	
ELA	VC018347	L213D001	56	46	A	+	A	+	+	A	C-	A	
ELA	VC019455	L252C001	57	46	A	+	C-	+	+	C-	A	A	
ELA	VC017211	L10SA061	57	75	A	+	C-	+	+	C-	A	A	
ELA	VC018706	L229D001	62	47	A	+	C-	+	+	C-	C-	A	
ELA	VC017171	L101C002	65	47	A	+	C-	+	+	C-	C-	A	
ELA	VC017296	L10SA229	69	75	A	+	C-	+	+	C-	A	B-	
ELA	VC018845	L234E004	82	46	A	+	A	+	+	A	C-	A	
ELA	VC019972	L276C002	91	46	C-	+	B-	+	+	B-	B-	B-	
Math	VC022511	M11819	1	80	B-	+	B-	+	+	C-	B-	B-	B-
Math	VC022235	M11046	3	21	A	+	C-	+	+	C-	C-	A	A
Math	VC025323	M25073	9	21	A	+	C-	+	+	C-	A	A	A
Math	VC025065	M22080	10	21	A	+	C-	+	+	C-	B-	C-	A
Math	VC024442	M21301	12	21	A	+	C-	+	+	B-	A	C-	A
Math	VC022328	M11292	13	57	A	+	C-	+	+	B-	A	A	A
Math	VC024101	M20866	19	21	A	+	C-	+	+	A	C-	A	A
Math	VC023106	M13477	22	19	C-	+	A	+	+	A	A	A	A
Math	VC024104	M20869	23	21	A	+	B-	+	+	B-	B-	C-	A
Math	VC023801	M20499	29	19	C-	+	A	+	+	A	A	B	A
Math	VC024920	M21905	29	21	A	+	C-	+	+	B-	A	A	A
Math	VC024131	M20901	30	57	A	+	C-	+	+	C-	C-	A	B-
Math	VC023246	M13837	38	21	A	+	C-	+	+	B-	B-	A	A
Math	VC023502	M20128	38	59	A	+	C-	+	+	B-	A	A	A
Math	VC025224	M23147	44	20	C-	+	A	+	+	A	A	A	A
Math	VC024019	M20754	44	35	A	+	C-	+	+	C-	C-	A	A
Math	VC026394	M25157	47	21	A	+	C-	+	+	C-	B-	A	A
Math	VC026377	M25034	48	19	A	+	C-	+	+	C-	B-	A	A
Math	VC023454	M20067	54	19	C-	+	B	+	+	A	A	A	A
Math	VC024298	M21106	54	35	A	+	B-	+	+	C-	B-	A	A
Math	VC023867	M20575	55	18	A	+	B-	+	+	C-	B-	B-	A
Math	VC022732	M12438	55	21	A	+	C-	+	+	C-	C-	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	VC025281	M25018	58	18	A	+	C-	+	+	B-	A	A	A
Math	VC023235	M13820	58	58	A	+	C-	+	+	B-	C-	B-	A
Math	VC026390	M25151	60	21	A	+	C-	+	+	B-	A	A	A
Math	VC023777	M20472	62	19	C-	+	A	+	+	A	A	B-	B-
Math	VC025385	M25139	62	35	A	+	B-	+	+	C-	B-	A	A
Math	VC025359	M25113	66	20	A	+	B-	+	+	C-	B-	A	A
Math	VC026445	M25303	67	35	A	+	C-	+	+	C-	C-	B-	A
Math	VC024317	M21132	68	19	B-	+	C-	+	+	C-	B-	B-	A
Math	VC025273	M25010	69	19	A	+	A	+	+	A	C-	C-	A
Math	VC024890	M21867	69	21	A	+	C-	+	+	B-	B-	B-	A
Math	VC024689	M21605	70	18	A	+	C-	+	+	B-	A	A	A
Math	VC023093	M13462	74	20	C-	+	A	+	+	A	A	A	A
Math	VC025336	M25086	75	35	A	+	C-	+	+	B-	A	A	A
Math	VC023719	M20398	75	59	A	+	C-	+	+	B-	A	A	A
Math	VC024096	M20856	77	21	A	+	C-	+	+	B-	A	A	A
Math	VC023690	M20360	86	19	C-	+	A	+	+	A	A	B-	A
Math	VC023114	M13488	86	21	C-	+	A	+	+	A	A	A	A
Math	VC023691	M20361	87	20	C-	+	A	+	+	A	A	A	A
Math	VC022412	M11510	87	60	A	+	B-	+	+	B-	A	C-	A
Math	VC024954	M21944	88	58	A	+	C-	+	+	B-	A	A	A
Math	VC023894	M20605	91	19	A	+	B-	+	+	C-	B-	A	A
Math	VC023806	M20505	95	20	C-	+	A	+	+	A	A	A	A

**Table 4.C.3: Listing of CR Item Statistics - ELA**

Accession number	<b>FM004559</b>
CAHSEE ID	L233D018
Polyserial correlation	<b>0.84</b>
IRT b-value	1.06
Step category 1*	2.82
Step category 2	1.16
Step category 3	1.73
Step category 4	-0.42
Step category 5	-0.59
Step category 6	-2.25
Step category 7	-2.45
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**	A

Notes. \* Step categories refer to the parameters describing each item category in the CR item calibrations.

\*\* This refers to the most disadvantageous DIF category found among all focal groups for which a comparison was made. Positive DIF categories disadvantage the reference group and negative DIF categories disadvantage the focal group.

**Table 4.C.4: Summary of Univariate Operational Item Statistics - ELA**

Content area	Number of items	IRT b-value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	0.05	0.81	-1.87	1.59	0.55	0.10	0.26	0.84
RC	18	-0.12	0.78	-1.50	1.59	0.52	0.09	0.35	0.66
RL	20	-0.18	0.75	-1.52	1.09	0.58	0.08	0.40	0.71
RW	7	-0.41	0.74	-1.56	0.83	0.54	0.14	0.26	0.69
WA	1	1.06	.	1.06	1.06	0.84		0.84	0.84
WC	15	0.17	0.82	-1.87	1.05	0.55	0.08	0.35	0.66
WS	12	0.71	0.62	-0.47	1.56	0.56	0.09	0.40	0.66

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WA=Writing Applications, WC=Writing Conventions, WS=Writing Strategies.

**Table 4.C.5: Summary of Univariate Operational Item Statistics - Mathematics**

Content area	Number of items	IRT b-value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.17	0.87	-2.61	1.80	0.55	0.10	0.31	0.74
A1	12	0.34	0.63	-0.68	1.48	0.53	0.12	0.34	0.68
AF	20	-0.34	1.03	-2.61	1.57	0.54	0.10	0.31	0.66
MG	18	-0.04	0.95	-1.96	1.80	0.60	0.08	0.46	0.74
MR	8	0.00	0.94	-1.97	1.05	0.49	0.07	0.40	0.61
NS	17	-0.33	0.58	-1.52	0.49	0.54	0.09	0.40	0.73
PS	13	-0.39	0.87	-1.97	1.26	0.54	0.09	0.38	0.68

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.C.6: Summary of Univariate Field-test Item Statistics - ELA**

Content area	Number of items	IRT b-value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	655	0.71	0.97	-1.60	5.01	0.47	0.15	-0.50	0.73
RC	164	0.78	0.88	-1.18	3.16	0.48	0.14	-0.11	0.73
RL	124	0.85	0.92	-0.97	3.29	0.45	0.16	-0.04	0.67
RW	90	0.64	0.90	-1.13	2.84	0.48	0.13	0.11	0.67
WC	138	0.30	1.03	-1.60	5.01	0.49	0.14	-0.38	0.70
WS	139	0.98	0.95	-0.51	4.90	0.44	0.17	-0.50	0.69

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW= Word Analysis, WC=Writing Conventions, WS=Writing Strategies.

**Table 4.C.7: Summary of Univariate Field-test Item Statistics - Mathematics**

Content area	Number of items	IRT b-value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	1139	0.62	1.03	-2.58	3.85	0.47	0.16	-0.57	0.76
A1	190	1.00	0.85	-1.44	2.96	0.43	0.17	-0.07	0.71
AF	267	0.52	0.93	-2.29	3.59	0.49	0.13	-0.35	0.71
MG	283	0.80	1.20	-1.92	3.85	0.44	0.20	-0.57	0.73
MR	94	0.37	0.98	-2.08	2.68	0.44	0.13	0.10	0.68
NS	261	0.47	0.97	-2.27	3.31	0.50	0.15	-0.31	0.76
PS	138	0.26	0.97	-2.58	3.34	0.48	0.13	-0.01	0.69

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.C.8: IRT Model Data Fit Distribution of Operational Items - ELA**

IRT review Category	Status	Item count	Percent
	A		
B	Use	7	10
C	Use	19	26
D	Review	5	7
F	Do not use	0	0
	Use	1	1
<b>Total</b>		73	100

**Table 4.C.9: IRT Model Data Fit Distribution of Operational Items - Mathematics**

IRT review Category	Status	Item count	Percent
	A		
B	Use	15	19
C	Use	17	21
D	Review	4	5
F	Do not use	3	4
<b>Total</b>		80	100

**Table 4.C.10: IRT Model Data Fit Distribution of Field-test Items - ELA**

IRT review Category	Status	Item count	Percent
	A		
B	Use	158	24
C	Use	258	40
D	Review	34	5
F	Do not use	47	7
<b>Total</b>		655	100

**Table 4.C.11: IRT Model Data Fit Distribution of Field-test Items - Mathematics**

IRT review Category	Status	Item count	Percent
	A		
B	Use	318	28
C	Use	464	41
D	Use	4	0
D	Review	46	4
F	Do not use	119	10
<b>Total</b>		1139	100

**Table 4.C.12: 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		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	2
B-	2	3	0	0	5	7	1	1	3	4	4	5	2	3	1	1		
A	68	93	73	100	62	85	68	93	64	88	64	88	70	96	72	99		
B+	2	3	0	0	2	3	3	4	3	4	2	3	0	0	0	0		
C+	1	1	0	0	3	4	0	0	2	3	2	3	0	0	0	0		
Small N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
TOTAL	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100		

**Table 4.C.13: 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		Native English Learner		Total C-items across all analyses		
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	
	C-	1	1	0	0	3	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4
B-	0	0	0	0	3	4	0	0	3	4	6	8	1	1	0	0	0	0			
A	78	98	80	100	70	88	80	100	76	95	72	90	79	99	80	100	80	100			
B+	0	0	0	0	4	5	0	0	1	1	2	3	0	0	0	0	0	0			
C+	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
Small N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
TOTAL	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100			

**Table 4.C.14: Distribution of Field-Test Item DIF Classifications - ELA**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	6	1	0	0	22	3	0	0	0	0	20	3	12	2	6	1	37
B-	12	2	0	0	50	8	0	0	0	0	46	7	37	6	39	6		
A	610	93	0	0	542	83	0	0	10	2	552	84	602	92	603	92		
B+	26	4	0	0	35	5	0	0	0	0	33	5	3	0	7	1		
C+	1	0	0	0	6	1	0	0	0	0	4	1	1	0	0	0		
Small N	0	0	655	100	0	0	655	100	645	98	0	0	0	0	0	0		
TOTAL	655	100	655	100	655	100	655	100	655	100	655	100	655	100	655	100		

**Table 4.C.15: Distribution of Field-Test Item DIF Classifications - Mathematics**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Native English-English Learner		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	10	1	0	0	25	2	0	0	0	0	16	1	8	1	5	0	0	0	44
B-	31	3	0	0	85	7	0	0	1	0	71	6	49	4	62	5	12	1		
A	1071	94	0	0	898	79	0	0	12	1	955	84	1077	95	1049	92	1112	98		
B+	24	2	0	0	117	10	0	0	0	0	89	8	5	0	21	2	15	1		
C+	3	0	0	0	14	1	0	0	0	0	8	1	0	0	2	0	0	0		
Small N	0	0	1139	100	0	0	1139	100	1126	99	0	0	0	0	0	0	0	0		
TOTAL	1139	100	1139	100	1139	100	1139	100	1139	100	1139	100	1139	100	1139	100	1139	100		

## Appendix 4.D: Item and DIF Statistics for May 2004

**Table 4.D.1: Listing of Operational Items Exhibiting Significant Negative DIF May 2004**

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	FM004751	L10SA528	0	78	A	A	C-	+	B-	C-	A	B-	<i>N/A for</i>
Math	FM006074	M10553	0	4	A	A	C-	+	C-	C-	A	A	A
Math	FM006918	M21563	0	14	A	A	C-	+	A	B-	A	A	A
Math	FM006355	M11861	0	31	A	A	C-	+	A	C-	A	A	A
Math	FM007452	M13872	0	32	A	A	C-	+	A	C-	A	A	A
Math	FM007542	M22077	0	33	A	A	C-	+	A	C-	A	A	A

**Table 4.D.2: Listing of Field-Test Items Exhibiting Significant Negative 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
Math	VC024859	M21821	21	35	C-	+	+	+	+	+	+	+	+
Math	VC023693	M20363	25	18	C-	+	+	+	+	+	+	+	+

**Table 4.D.3: Listing of CR Item Statistics - ELA**

Accession number	FM004506
CAHSEE ID	L1000022
Polyserial correlation	0.82
IRT b Value	.415
Step category 1*	1.087
Step category 2	1.504
Step category 3	2.261
Step category 4	-0.376
Step category 5	-0.058
Step category 6	-2.126
Step category 7	-2.292
DIF category, Male-Female	A
DIF category, White-American Indian	A
DIF category, White-Asian	A
DIF category, White-Pacific Islander	‡
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**	

Notes. \* Step categories refer to the parameters describing each item category in the CR item calibrations.

\*\* This refers to the most disadvantageous DIF category found among all focal groups for which a comparison was made. Positive DIF categories disadvantage the reference group and negative DIF categories disadvantage the focal group.

‡ Focal group sample size too small for analysis.

**Table 4.D.4: Summary Univariate Operational Item Statistics - ELA**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	73	-0.04	0.64	-1.66	1.37	0.58	0.10	0.36	0.82
RC	18	-0.04	0.53	-1.32	0.87	0.57	0.09	0.44	0.72
RL	20	-0.19	0.66	-1.66	0.74	0.57	0.10	0.36	0.72
RW	7	-0.63	0.65	-1.51	0.45	0.63	0.08	0.52	0.76
WA	1	0.42	.	0.42	0.42	0.82	.	0.82	0.82
WC	15	0.02	0.61	-1.41	0.99	0.58	0.09	0.46	0.76
WS	12	0.46	0.49	-0.26	1.37	0.58	0.08	0.46	0.75

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WA=Writing Applications, WC=Writing Conventions, WS=Writing Strategies.

**Table 4.D.5: Summary Univariate Operational Item Statistics - Mathematics**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	80	-0.22	0.72	-1.74	1.84	0.52	0.12	0.20	0.72
A1	12	0.45	0.51	-0.54	1.30	0.46	0.11	0.27	0.66
AF	20	-0.27	0.62	-1.59	0.74	0.54	0.13	0.35	0.70
MG	18	-0.15	0.82	-1.39	1.84	0.54	0.11	0.30	0.69
MR	8	-0.17	0.74	-0.96	0.86	0.45	0.15	0.20	0.67
NS	17	-0.45	0.53	-1.19	0.86	0.49	0.11	0.20	0.72
PS	13	-0.53	0.75	-1.74	1.02	0.58	0.11	0.43	0.71

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.D.6: Summary Univariate Field-Test Item Statistics - ELA**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	172	0.58	0.75	-1.06	2.67	0.49	0.17	-0.17	0.80
RC	10	0.63	0.56	0.15	1.97	0.47	0.18	0.01	0.66
RL	59	0.77	0.67	-0.61	2.40	0.46	0.17	-0.08	0.71
RW	24	0.47	0.75	-0.64	2.06	0.55	0.17	0.12	0.76
WC	31	0.27	0.96	-1.06	2.66	0.52	0.20	-0.09	0.80
WS	48	0.59	0.65	-0.76	2.67	0.47	0.16	-0.17	0.67

Note. RC=Reading Comprehension, RL=Literary Response & Analysis, RW=Word Analysis, WC=Writing Conventions, WS=Writing Strategies.

**Table 4.D.7: Summary Univariate Field-Test Item Statistics - Mathematics**

Content area	Number of items	IRT b value				R-biserial			
		Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
Overall	300	0.31	0.76	-1.29	2.73	0.41	0.18	-0.30	0.77
A1	50	0.71	0.57	-0.73	1.78	0.29	0.16	-0.14	0.60
AF	62	0.28	0.74	-1.15	1.97	0.47	0.17	0.13	0.75
MG	56	0.44	0.80	-1.10	2.73	0.39	0.21	-0.30	0.76
MR	44	0.13	0.75	-1.29	1.85	0.42	0.18	0.01	0.71
NS	74	0.24	0.75	-1.16	2.73	0.43	0.16	-0.02	0.70
PS	58	-0.04	0.71	-1.29	1.78	0.44	0.16	-0.07	0.77

Note. A1=Algebra I, AF=Algebra & Functions, MG=Measurement & Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

**Table 4.D.8: IRT Model Data Fit Distribution of Operational Items - ELA**

IRT review Category	Status	Item count	Percent
A	Use	12	16
B	Use	21	29
C	Use	38	52
D	Review	2	3
<b>Total</b>		73	100

**Table 4.D.9: IRT Model Data Fit Distribution of Operational Items - Mathematics**

IRT review Category	Status	Item count	Percent
A	Use	20	25
B	Use	13	16
C	Use	43	54
D	Review	2	3
F	Do not use	2	3
<b>Total</b>		80	100

**Table 4.D.10: IRT Model Data Fit Distribution of Field-Test Items - ELA**

IRT review Category	Status	Item count	Percent
A	Use	6	3
B	Use	17	10
C	Use	105	61
D	Review	20	12
F	Do not use	24	14
<b>Total</b>		172	100

**Table 4.D.11: IRT Model Data Fit Distribution of Field-Test Items - Mathematics**

IRT review Category	Status	Item count	Percent
A	Use	6	2
B	Use	22	7
C	Use	166	55
D	Review	21	7
F	Do not use	85	28
<b>Total</b>		300	100

**Table 4.D.12: 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		Total C-items across all analyses	
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
	C-	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	1
B-	2	3	2	3	3	4	0	0	9	12	3	4	0	0	3	4		
A	71	97	68	93	62	85	0	0	60	82	65	89	72	99	69	95		
B+	0	0	3	4	7	10	0	0	4	5	4	5	1	1	1	1		
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Small N	0	0	0	0	0	0	73	100	0	0	0	0	0	0	0	0		
TOTAL	73	100	73	100	73	100	73	100	73	100	73	100	73	100	73	100		

**Table 4.D.13: 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		Native English-English Learner		Total C-items across all analyses	
	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	5	6	0	0	1	1	4	5	0	0	0	0	0	0	5
B-	0	0	0	0	10	13	0	0	1	1	5	6	0	0	0	0	0	0		
A	80	100	79	99	52	65	0	0	73	91	64	80	80	100	80	100	80	100		
B+	0	0	1	1	9	11	0	0	5	6	6	8	0	0	0	0	0	0		
C+	0	0	0	0	4	5	0	0	0	0	1	1	0	0	0	0	0	0		
Small N	0	0	0	0	0	0	80	100	0	0	0	0	0	0	0	0	0	0		
TOTAL	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100	80	100		

**Table 4.D.14: Distribution of Field-Test Item DIF Classifications - ELA**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Total C-items across all analyses	
	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
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	3	2	0	0	0	0	0	0	0	0	0	0	3	2	3	2		
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Small N	169	98	172	100	172	100	172	100	172	100	172	100	169	98	169	98		
TOTAL	172	100	172	100	172	100	172	100	172	100	172	100	172	100	172	100		

**Table 4.D.15: Distribution of Field-Test Item DIF Classifications - Mathematics**

DIF category	Male-Female		White-American Indian		White-Asian		White-Pacific Islander		White-Filipino		White-Combined Asian		White-Hispanic		White-African American		Native English-Learner		Total C-items across all analyses			
	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.		
	C-	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1
B-	13	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
A	232	77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
B+	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
Small N	48	16	300	100	300	100	300	100	300	100	300	100	300	100	300	100	300	100	300	100		
TOTAL	300	100	300	100	300	100	300	100	300	100	300	100	300	100	300	100	300	100	300	100		

## **Chapter 5: Summary of Test-Level Analyses**

This chapter summarizes the test-level statistics obtained for the three CAHSEE administrations given in February, March and May 2004. Test-level statistics are discussed for the ELA and mathematics examinations. To simplify the presentation of data, all tables for this section are located in Appendix 5.A to 5.C. The same table numbering system applies in Appendices 5.A (February), 5.B (March) and 5.C (May), as described in the following paragraphs

### **Demographic Distributions**

The target population for the CAHSEE is California high school students who are in at least the 10<sup>th</sup> grade. Summary statistics for the subgroups based on grade, gender, ethnicity, language fluency, economic disadvantage and special education programs are presented in Tables 1 and 2 for February (Appendix 5.A), March (5.B) and May (5.C). These summary statistics are based on all students taking each of the administrations.

Similar summary statistics for demographic groups are presented in Tables 3 and 4 in Appendices 5.A to 5.C. These summary statistics are based on students taking each of the administrations and display the percents at the NCLB cuts of below proficient, proficient and advanced levels.

Selected percentiles, scale-score means and standard deviations for the subgroups mentioned above based on gender, ethnicity, grade, language fluency, economic disadvantage and special education program are presented for all students in Tables 5 and 6 in Appendices 5.A to 5.C.

### **Score Distributions and Summary Statistics**

Frequency distributions of scale scores for ELA and mathematics are presented in Tables 7 and 8 in Appendices 5.A to 5.C. The passing line indicates the pass/Not Pass cuts. Similar distributions of scale scores for ELA and mathematics, with the NCLB cuts indicated are presented in Tables 9 and 10 in Appendices 5.A to 5.C. The first line indicates the Advanced level cut; while the second line indicates the Proficient level cut.

### **Strand Intercorrelations, Reliability and Standard Error of Measurement**

Intercorrelations, reliability and standard error of measurement estimates are reported for each major assessment – mathematics, ELA and the CR item - as well as for each major content area within each assessment: for ELA, word analysis, reading comprehension, literary responses and analysis, writing strategies, writing conventions, and for mathematics, probability and statistics, number sense, algebra and functions, measurement and geometry, and Algebra I. Reliabilities and intercorrelations among the MC English-language arts and mathematics sections, CR items, and reported subscores appear in Table 11 in Appendices 5.A to 5.C. Summary analyses also include information about rater consistency for each CR item in Tables 12 and 13 in Appendices 5.A to 5.C.

Observed-score correlations are the correlations between the raw scores obtained on the different tests or sections by individual test takers. During the February, March and May 2004 administrations, the correlations between the ELA and CR item sections ranged from 0.79 to 0.82, while the correlations between ELA and mathematics and the ELA CR item and mathematics ranged from 0.74 to 0.78 and 0.57 to 0.62, respectively. Correlations between total ELA and the

ELA subscores ranged from 0.74 to 0.92, while correlations between the CR items and the ELA subscores ranged from 0.51 to 0.67. Total mathematics correlated from 0.79 to 0.92 with the mathematics subscores. Correlations between total ELA and the mathematics subscores ranged from 0.52 to 0.74 and correlations between total mathematics and the ELA subscores ranged from 0.56 to 0.70. CR item and mathematics subscores correlations ranged from 0.39 to 0.57.

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 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, which are derived from analysis of the consistency of the performance of individuals on items within a test (internal-consistency reliability). 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 formulas for the internal consistency reliability and for the weighted composite reliability are reported below.

(1) Coefficient Alpha (Cronbach, 1951):

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

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

(2) Reliability of composite scores (Feldt & Brennan, 1989):

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

where  $k$  is the number of part scores in the composite,  $w_j$  is the weight associated with the  $j$ -th part score,  $\sigma_{e_j}$  is the SEM of the  $j$ -th part score, and  $\sigma_c^2$  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 1 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

$$(3) (\text{Corr}_{\text{CR-MC}}/\text{Rel}_{\text{MC}}).$$

Reliability estimates for both the ELA and Mathematics sections across the three administrations ranged from 0.94 to 0.95. Reliabilities for the subsections for ELA ranged from 0.58 to 0.84, while the subscore reliabilities for Mathematics ranged from 0.70 to 0.83. The reliabilities for the CR item score ranged from 0.72 to 0.77.

The formula for computing the standard error of measurement is:

$$(4) \sigma_e = \sigma_x \sqrt{1 - \alpha}$$

where reliability is the reliability estimated using (1), (2), or (3) above, and  $\sigma_x$  is the standard deviation of the score being examined.

The standard errors of measurement (SEM) for the ELA weighted raw scores and the Mathematics raw scores across the February, March and May 2004 administrations ranged from 3.73 to 4.21 and 3.64 to 3.93, respectively. The standard errors of measurement for the subsections for ELA ranged from 0.91 to 1.89, while the standard error of measurement of subsections for Mathematics ranged from 1.42 to 1.96. The standard errors for the CR item ranged from 0.39 to 0.49.

### **CR item and Rater Agreement Summary**

The test form for each administration includes one CR item question. It is either a passage-based prompt or a stand-alone prompt. Two raters using the Online Scoring Network (OSN) score responses to each prompt. An important part of the analysis of any multiple-rated CR item is the degree to which the individual rater scores agree. Table 12 in Appendices 5.A to 5.C presents the possible score combinations which each CR item received after 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 to a CR item, while most ratings assigning different scores differed by only 1 point. Less than 1% of CR item scores assigned by the two raters differed by three score points for the February and March administrations. No CR item scores differed by 3 score points for the May administration. In the three administrations, no raters differed by more than 3 points. Whenever raters differ by more than one point a third rater is used to determine the score.

Table 13 in Appendices 5.A to 5.C presents the mean 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 three administrations ranged from 0.33 to 0.40. The correlation between first and second ratings across the three administrations ranged from 0.72 to 0.82 for the CR item.

Table 13 also summarizes the reasons given for CR items that received a score of zero. Zero scores are given if a CR item is left blank, is illegible, does not address the topic, is a cartoon/drawing, uses inappropriate content, or if a CR item was not written in English. During the three administrations, the percentage of zero scores received on the CR item ranged from 2.1% to 6.5% with the majority of zero scores due to CR items which were either left blank or written off topic.

The lowest percentage of zero scores for the CR item was in the March administration and the largest percentage was in the May 2004 administration.

### **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 4.12). RELCLASS-COMP estimates decision *consistency* using an estimated multivariate distribution of reported classifications on the current form of the exam and classifications on the alternate (parallel) form. RELCLASS-COMP estimates decision *accuracy* using an estimated multivariate distribution of reported classifications on the current form of the exam 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 a  $n$  by  $n$  table (where  $n$  is the number of proficiency levels) and summing the entries in the diagonal.

The reliability classification results for the NCLB cuts at the advanced and proficient levels, for both accuracy and consistency, are reported in Table 14 in Appendices 5.A to 5.C. During the three administrations the decision accuracy for ELA at the advanced level ranged from 0.93 to 0.96, while the decision accuracy for mathematics at the advanced level ranged from 0.95 to 0.98. The decision consistency for ELA at the advanced level ranged from 0.90 to 0.94, while the decision consistency for mathematics at the advanced level ranged from 0.93 to 0.98. The decision accuracy for ELA at the proficient level ranged from 0.92 to 0.94, while the decision accuracy for mathematics at the proficient level ranged from 0.93 to 0.95. The decision consistency for ELA at the proficient level ranged from 0.90 to 0.92, while the decision consistency for mathematics ranged from 0.90 to 0.94 at the proficient level.

The reliability classification results for the cuts at the pass/not pass levels, for both accuracy and consistency are reported in Table 15 in Appendices 5.A to 5.C. During the three administrations the decision accuracy for ELA at the pass/not pass level ranged from 0.94 to 0.95, while the decision accuracy for mathematics at the pass/not pass level ranged from 0.93 to 0.95. The decision consistency for ELA at the pass/not pass level ranged from 0.91 to 0.93, while the decision consistency for mathematics ranged from 0.90 to 0.92. The magnitudes of the numbers (0.90 or above) reflect a high level of accuracy and consistency in the student classifications.

**Appendix 5.A: Demographic Summaries, Frequency Distributions, Intercorrelations,  
and Decision Consistencies—February 2004**

**Table 5.A.1: Demographic Summary For All Examinees<sup>1</sup> ELA February 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Appl. Avg. Score
							WA	RC	LR/A	WS	WC	
							CR item					
<b>Total Examinees</b>	146463	107748	74%	38715	26%	374	78%	72%	73%	63%	72%	2.5
<b>Grade</b>												
Tenth	146328	107686	74%	38642	26%	374	78%	72%	73%	63%	72%	2.5
Eleventh	2	.	.	.	.	.	.	.	.	.	.	.
Twelfth	.	.	.	.	.	.	.	.	.	.	.	.
Adult Education	126	58	46%	68	54%	350	69%	66%	66%	51%	56%	1.9
Unknown	7	.	.	.	.	.	.	.	.	.	.	.
<b>Gender</b>												
Male	74633	51607	69%	23026	31%	369	77%	70%	71%	61%	68%	2.3
Female	71730	56091	78%	15639	22%	379	78%	73%	75%	66%	75%	2.6
Unknown	100	50	50%	50	50%	350	70%	60%	64%	53%	57%	1.9
<b>Race/Ethnicity</b>												
American Indian or Alaska Native	1442	1060	74%	382	26%	371	78%	72%	74%	62%	70%	2.4
Asian	13095	10570	81%	2525	19%	384	79%	76%	76%	69%	78%	2.7
Pacific Islander	1052	760	72%	292	28%	369	75%	69%	71%	60%	71%	2.5
Filipino	4540	3929	87%	611	13%	387	82%	77%	78%	71%	79%	2.8
Hispanic or Latino	57859	35129	61%	22730	39%	359	72%	65%	67%	55%	64%	2.2
African American	13702	8481	62%	5221	38%	360	72%	65%	68%	55%	63%	2.3
White (not of Hispanic origin)	53190	46663	88%	6527	12%	390	84%	80%	80%	73%	79%	2.7
Unknown	1583	1156	73%	427	27%	373	78%	72%	74%	63%	70%	2.4
<b>Language Fluency</b>												
English Only Students	93369	75228	81%	18141	19%	381	81%	76%	77%	68%	75%	2.6
Initially Fluent English Proficient (IFEP)	10891	9085	83%	1806	17%	382	81%	76%	77%	69%	77%	2.6
Redesignated Fluent English Proficient (RFEP)	15239	13337	88%	1902	12%	380	81%	76%	76%	68%	77%	2.6
English Learner Students	26484	9846	37%	16638	63%	340	62%	54%	58%	44%	55%	1.9
Unknown	480	252	53%	228	48%	353	69%	63%	65%	52%	59%	2.0
<b>Economically Disadvantaged</b>												
No	68617	59714	87%	8903	13%	389	84%	79%	80%	72%	79%	2.7
Yes	59516	34962	59%	24554	41%	357	71%	64%	66%	54%	64%	2.2
Unknown	18330	13072	71%	5258	29%	371	77%	71%	73%	62%	70%	2.4
<b>Special Education Program Participation</b>												
Students Receiving Services	12105	3426	28%	8679	72%	333	59%	51%	55%	40%	47%	1.7
Students Not Receiving Services	134358	104322	78%	30036	22%	377	79%	74%	75%	66%	74%	2.5

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> WA — Word Analysis, RC — Reading Comprehension, LR/A — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

**Table 5.A.2: Demographic Summary For All Examinees<sup>1</sup> Mathematics February 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>				
							Average Percent Correct				
							PS	NS	AF	MG	Alg1
<b>Total Examinees</b>	145699	104322	72%	41377	28%	374	68%	67%	67%	67%	57%
<b>Grade</b>											
Tenth	145574	104276	72%	41298	28%	374	68%	67%	67%	67%	57%
Eleventh	2	.	.	.	.	.	.	.	.	.	.
Twelfth	.	.	.	.	.	.	.	.	.	.	.
Adult Education	116	42	36%	74	64%	343	53%	49%	51%	54%	34%
Unknown	7	.	.	.	.	.	.	.	.	.	.
<b>Gender</b>											
Male	74071	52370	71%	21701	29%	374	69%	67%	67%	67%	57%
Female	71527	51905	73%	19622	27%	373	68%	66%	68%	67%	58%
Unknown	101	47	47%	54	53%	351	56%	54%	57%	54%	41%
<b>Race/Ethnicity</b>											
American Indian or Alaska Native	1428	984	69%	444	31%	369	67%	65%	65%	64%	54%
Asian	13079	11449	88%	1630	12%	396	76%	77%	78%	78%	73%
Pacific Islander	1044	737	71%	307	29%	370	66%	65%	67%	65%	56%
Filipino	4534	3907	86%	627	14%	387	74%	73%	75%	74%	67%
Hispanic or Latino	57735	33793	59%	23942	41%	359	61%	59%	60%	60%	49%
African American	13513	7168	53%	6345	47%	355	59%	57%	58%	56%	46%
White (not of Hispanic origin)	52771	45235	86%	7536	14%	388	77%	74%	75%	74%	65%
Unknown	1595	1049	66%	546	34%	369	67%	65%	64%	64%	53%
<b>Language Fluency</b>											
English Only Students	92639	70906	77%	21733	23%	378	72%	69%	70%	69%	59%
Initially Fluent English Proficient (IFEP)	10867	8626	79%	2241	21%	382	72%	71%	72%	71%	62%
Redesignated Fluent English Proficient (RFEP)	15356	12673	83%	2683	17%	381	72%	71%	72%	71%	62%
English Learner Students	26372	11899	45%	14473	55%	350	53%	54%	54%	54%	46%
Unknown	465	218	47%	247	53%	352	57%	55%	56%	56%	43%
<b>Economically Disadvantaged</b>											
No	68363	57523	84%	10840	16%	387	76%	73%	74%	74%	65%
Yes	59188	34500	58%	24688	42%	360	60%	60%	60%	60%	49%
Unknown	18148	12299	68%	5849	32%	370	67%	65%	65%	65%	54%
<b>Special Education Program Participation</b>											
Students Receiving Services	10639	2834	27%	7805	73%	336	47%	47%	45%	45%	36%
Students Not Receiving Services	135060	101488	75%	33572	25%	376	70%	68%	69%	69%	59%

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 5.A.3: NCLB Demographic Summary for All CAHSEE Examinees ELA February 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		146463	78910	54	32163	22	35390	24	67553	46
<b>Grade</b>	Tenth	146328	78796	54	32147	22	35385	24	67532	46
	Eleventh	2	.	.	.	.	.	.	.	.
	Twelfth	0	.	.	.	.	.	.	.	.
	Adult Education	126	108	86	14	11	4	3	18	14
	Unknown	7	.	.	.	.	.	.	.	.
<b>Gender</b>	Male	74633	43995	59	15442	21	15196	20	30638	41
	Female	71730	34839	49	16705	23	20186	28	36891	51
	Unknown	100	76	76	16	16	8	8	24	24
<b>Race/ Ethnicity</b>	American Indian or Alaskan Native	1442	836	58	310	21	296	21	606	42
	Asian	13095	5649	43	2899	22	4547	35	7446	57
	Pacific Islander	1052	621	59	250	24	181	17	431	41
	Filipino	4540	1782	39	1236	27	1522	34	2758	61
	Hispanic or Latino	57859	41101	71	10358	18	6400	11	16758	29
	African American	13702	9517	69	2513	18	1672	12	4185	31
	White (not of Hispanic origin)	53190	18549	35	14263	27	20378	38	34641	65
	Unknown	1583	855	54	334	21	394	25	728	46
<b>Language Fluency</b>	English Only Students	93369	42116	45	22912	25	28341	30	51253	55
	Initially Fluent English Proficient (IFEP)	10891	4915	45	2826	26	3150	29	5976	55
	Redesignated Fluent English Proficient (RFEP)	15239	7395	49	4540	30	3304	22	7844	51
	English Learner Students	26484	24120	91	1813	7	551	2	2364	9
	Unknown	480	364	76	72	15	44	9	116	24
<b>Economic Disadvantage</b>	No	68617	25283	37	17928	26	25406	37	43334	63
	Yes	59516	43444	73	10063	17	6009	10	16072	27
	Unknown	18330	10183	56	4172	23	3975	22	8147	44
<b>Special Education Program Participation</b>	Receiving Services	12105	11073	91	743	6	289	2	1032	9
	Not Receiving Services	134358	67837	50	31420	23	35101	26	66521	50

**Table 5.A.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics February 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		145699	83735	57	43227	30	18737	13	61964	43
<b>Grade</b>	Tenth	145574	83624	57	43214	30	18736	13	61950	43
	Eleventh	2	.	.	.	.	.	.	.	.
	Twelfth	0	.	.	.	.	.	.	.	.
	Adult Education	116	105	91	10	9	1	1	11	9
	Unknown	7	.	.	.	.	.	.	.	.
<b>Gender</b>	Male	74071	41809	56	21724	29	10538	14	32262	44
	Female	71527	41846	59	21485	30	8196	11	29681	41
	Unknown	101	80	79	18	18	3	3	21	21
<b>Race/Ethnicity</b>	American Indian or Alaskan Native	1428	888	62	422	30	118	8	540	38
	Asian	13079	4320	33	4712	36	4047	31	8759	67
	Pacific Islander	1044	628	60	316	30	100	10	416	40
	Filipino	4534	1891	42	1852	41	791	17	2643	58
	Hispanic or Latino	57735	43252	75	11925	21	2558	4	14483	25
	African American	13513	10511	78	2489	18	513	4	3002	22
	White (not of Hispanic origin)	52771	21264	40	21080	40	10427	20	31507	60
	Unknown	1595	981	62	431	27	183	11	614	38
	<b>Language Fluency</b>	English Only Students	92639	47597	51	30979	33	14063	15	45042
Initially Fluent English Proficient (IFEP)		10867	5377	49	3553	33	1937	18	5490	51
Redesignated Fluent English Proficient (RFEP)		15356	7872	51	5371	35	2113	14	7484	49
English Learner Students		26372	22514	85	3251	12	607	2	3858	15
Unknown		465	375	81	73	16	17	4	90	19
<b>Economical Disadvantage</b>		No	68363	28808	42	25885	38	13670	20	39555
	Yes	59188	43764	74	12304	21	3120	5	15424	26
	Unknown	18148	11163	62	5038	28	1947	11	6985	38
<b>Special Education Program Participation</b>	Receiving Services	10639	9687	91	783	7	169	2	952	9
	Not Receiving Services	135060	74048	55	42444	31	18568	14	61012	45

**Table 5.A.5: Demographic Summary For All Examinees<sup>1</sup> ELA – February 2004**

	Percentiles <sup>2</sup>							Mean Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	289	309	346	375	400	434	450	374	38	146463
<b>Grade</b>										
Tenth	289	309	346	375	400	434	450	374	38	146328
Eleventh										0
Twelfth										0
Adult Education	293	305	335	347	366	400	407	350	26	126
Unknown										7
<b>Gender</b>										
Female	284	305	341	371	397	429	450	369	38	74633
Male	293	314	354	380	407	441	450	379	37	71730
Unknown	276	287	319	349	378	409	438	350	39	100
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	291	310	346	373	397	429	450	371	36	1442
Asian	295	316	358	385	411	448	450	384	39	13095
Pacific Islander	293	307	346	371	394	429	450	369	35	1052
Filipino	307	330	364	388	411	441	450	387	33	4540
Hispanic or Latino	282	301	333	360	383	415	441	359	35	57859
African American	279	301	335	360	385	419	441	360	35	13702
White (not of Hispanic origin)	299	328	368	391	415	448	450	390	34	53190
Unknown	289	305	346	375	400	434	450	373	39	1583
<b>Language Fluency</b>										
English Only Students	293	316	358	383	407	441	450	381	37	93369
Initially Fluent English Proficient (IFEP)	299	323	360	383	407	441	450	382	34	10891
Redesignated Fluent English Proficient (RFEP)	312	335	362	380	400	429	450	380	28	15239
English Learner Students	276	295	319	339	360	388	411	340	29	26484
Unknown	279	295	326	350	378	419	441	353	37	480
<b>Economically Disadvantaged</b>										
No	301	328	366	391	415	448	450	389	34	68617
Yes	282	301	332	358	380	415	441	357	35	59516
Unknown	284	305	344	373	397	434	450	371	38	18330
<b>Special Education Program Participation</b>										
Students Receiving Services	275	287	310	330	352	388	415	333	31	12105
Students Not Receiving Services	293	316	352	378	403	434	450	377	36	134358

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.A.6: Demographic Summary For All Examinees<sup>1</sup> Mathematics February 2004**

	Percentiles <sup>2</sup>							Mean Scale Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	302	318	346	371	400	447	450	374	37	145699
<b>Grade</b>										
Tenth	302	318	346	371	400	447	450	374	37	145574
Eleventh										0
Twelfth										0
Adult Education	302	309	325	341	357	389	410	343	25	116
Unknown										7
<b>Gender</b>										
Female	300	314	345	371	403	447	450	374	39	74071
Male	305	320	346	371	397	439	450	373	35	71527
Unknown	297	307	326	345	367	410	433	351	33	101
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	300	314	343	369	392	433	450	369	35	1428
Asian	314	332	369	397	427	450	450	396	38	13079
Pacific Islander	297	316	345	369	394	433	450	370	35	1044
Filipino	312	332	363	386	410	447	450	387	34	4534
Hispanic or Latino	300	314	335	356	380	418	447	359	32	57735
African American	295	309	332	352	375	414	447	355	32	13513
White (not of Hispanic origin)	309	328	363	386	414	450	450	388	35	52771
Unknown	292	312	339	365	397	439	450	369	39	1595
<b>Language Fluency</b>										
English Only Students	302	318	352	377	406	447	450	378	37	92639
Initially Fluent English Proficient (IFEP)	309	324	354	380	410	450	450	382	37	10867
Redesignated Fluent English Proficient (RFEP)	316	332	356	377	403	447	450	381	33	15356
English Learner Students	297	309	330	346	365	403	439	350	28	26372
Unknown	295	307	328	346	371	410	450	352	33	465
<b>Economically Disadvantaged</b>										
No	309	328	361	386	414	450	450	387	36	68363
Yes	300	312	335	356	380	422	450	360	33	59188
Unknown	300	314	341	367	394	439	450	370	37	18148
<b>Special Education Program Participation</b>										
Students Receiving Services	286	300	316	330	352	392	433	336	29	10639
Students Not Receiving Services	307	322	350	373	403	447	450	376	36	135060

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.A.7: Frequency Distributions ELA February 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	2913	2	2913	98
440-449	3683	3	6596	96
430-439	2587	2	9183	94
420-429	6304	4	15487	89
410-419	11629	8	27116	81
400-409	12441	8	39557	73
390-399	12542	9	52099	64
380-389	15454	11	67553	54
370-379	13682	9	81235	45
360-369	14518	10	95753	35
350-359 <sup>1</sup>	11995	8	107748	26
340-349	9576	7	117324	20
330-339	9261	6	126585	14
320-329	6279	4	132864	9
310-319	5920	4	138784	5
300-309	3588	2	142372	3
290-299	2331	2	144703	1
280-289	982	1	145685	1
275-279	778	1	146463	0

<sup>1</sup> Passing Line**Table 5.A.8: Frequency Distributions Mathematics February 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5272	4	5272	96
440-449	2327	2	7599	95
430-439	5383	4	12982	91
420-429	5755	4	18737	87
410-419	9243	6	27980	81
400-409	9362	6	37342	74
390-399	9391	6	46733	68
380-389	15231	10	61964	57
370-379	12028	8	73992	49
360-369	14398	10	88390	39
350-359 <sup>1</sup>	15932	11	104322	28
340-349	12172	8	116494	20
330-339	12331	8	128825	12
320-329	8244	6	137069	6
310-319	4751	3	141820	3
300-309	2977	2	144797	1
290-299	558	0	145355	0
280-289	188	0	145543	0
275-279	156	0	145699	0

<sup>1</sup> Passing Line

**Table 5.A.9: Frequency Distributions ELA February 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	2913	2	2913	98
440-449	3683	3	6596	96
430-439	2587	2	9183	94
420-429	6304	4	15487	89
410-419	11629	8	27116	81
403-409 <sup>1</sup>	8274	6	35390	76
390-402	16709	11	52099	64
380-389 <sup>2</sup>	15454	11	67553	54
370-379	13682	9	81235	45
360-369	14518	10	95753	35
350-359	11995	8	107748	26
340-349	9576	7	117324	20
330-339	9261	6	126585	14
320-329	6279	4	132864	9
310-319	5920	4	138784	5
300-309	3588	2	142372	3
290-299	2331	2	144703	1
280-289	982	1	145685	1
275-279	778	1	146463	0

<sup>1</sup> Advanced Level Cut

<sup>2</sup> Proficient Level Cut

**Table 5.A.10: Frequency Distributions Mathematics February 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	5272	4	5272	96
440-449	2327	2	7599	95
430-439	5383	4	12982	91
422-429 <sup>1</sup>	5755	4	18737	87
410-421	9243	6	27980	81
400-409	9362	6	37342	74
390-399	9391	6	46733	68
380-389 <sup>2</sup>	15231	10	61964	57
370-379	12028	8	73992	49
360-369	14398	10	88390	39
350-359	15932	11	104322	28
340-349	12172	8	116494	20
330-339	12331	8	128825	12
320-329	8244	6	137069	6
310-319	4751	3	141820	3
300-309	2977	2	144797	1
290-299	558	0	145355	0
280-289	188	0	145543	0
275-279	156	0	145699	0

<sup>1</sup> Advanced Level Cut

<sup>2</sup> Proficient Level Cut

**Table 5.A.11: Intercorrelations<sup>1</sup> and Reliability Estimates By Section February 2004**

	ELA <sup>4</sup>	CR item	RW	RC	RL	WS	WC	Math	PS	NS	AF	MG	A1
ELA <sup>2</sup>	1.00	0.79	0.74	0.89	0.88	0.86	0.86	0.78	0.71	0.68	0.74	0.70	0.61
CR item	0.79	1.00	0.51	0.61	0.61	0.58	0.60	0.59	0.53	0.51	0.56	0.53	0.48
Word Analysis	0.74	0.51	1.00	0.64	0.65	0.57	0.58	0.56	0.53	0.50	0.53	0.50	0.41
Reading Comprehension	0.89	0.61	0.64	1.00	0.78	0.73	0.69	0.70	0.65	0.61	0.66	0.63	0.53
Literary Responses & Analysis	0.88	0.61	0.65	0.78	1.00	0.71	0.69	0.67	0.63	0.59	0.64	0.60	0.51
Writing Strategies	0.86	0.58	0.57	0.73	0.71	1.00	0.72	0.68	0.61	0.59	0.64	0.62	0.54
Writing Conventions	0.86	0.60	0.58	0.69	0.69	0.72	1.00	0.69	0.62	0.60	0.66	0.63	0.55
Math <sup>3</sup>	0.78	0.59	0.56	0.70	0.67	0.68	0.69	1.00	0.85	0.88	0.92	0.91	0.84
Probability and Statistics	0.71	0.53	0.53	0.65	0.63	0.61	0.62	0.85	1.00	0.72	0.75	0.72	0.62
Number Sense	0.68	0.51	0.50	0.61	0.59	0.59	0.60	0.88	0.72	1.00	0.76	0.74	0.68
Algebra & Functions	0.74	0.56	0.53	0.66	0.64	0.64	0.66	0.92	0.75	0.76	1.00	0.79	0.73
Measurement & Geometry	0.70	0.53	0.50	0.63	0.60	0.62	0.63	0.91	0.72	0.74	0.79	1.00	0.72
Algebra 1	0.61	0.48	0.41	0.53	0.51	0.54	0.55	0.84	0.62	0.68	0.73	0.72	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	62.41	2.45	5.43	12.94	14.66	7.61	10.74	52.64	8.89	11.34	13.48	12.07	6.87
SD	16.07	0.79	1.48	3.78	3.61	3.08	3.44	15.73	2.88	3.61	4.24	3.93	3.11
Reliability	0.94	0.73	0.58	0.78	0.78	0.77	0.81	0.94	0.74	0.77	0.81	0.81	0.77
SEM	3.86	0.41	0.95	1.76	1.70	1.47	1.50	3.71	1.46	1.72	1.82	1.72	1.49

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WC=Writing Conventions, WS=Writing Strategies, A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

<sup>1</sup> All correlations are for raw scores.

<sup>2</sup> Correlations for ELA section are reported for examinees completing ELA section. Number of examinees: 146463

<sup>3</sup> Correlations for Math section are reported for examinees completing Math section. Number of examinees: 145620

<sup>4</sup> Correlations between ELA and Math are reported for examinees completing both sections. Number of examinees: 40689

**Table 5.A.12: Agreement of First and Second Ratings on CR item**

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	4139	0	0	0	0	4139	3
1	0	6549	3753	172	3	10477	7
2	0	3555	40235	14696	414	58900	40
3	0	109	14518	40689	5791	61107	42
4	0	6	422	6134	5278	11840	8
Total	4139	10219	58928	61691	11486	146463	100
Percent	3	7	40	42	8	100	

**Differences Between First and Second Ratings on CR item**

Difference	Frequency	Percent	Cumulative Percent
0	96890	66	66
1	48447	33	99
2	1117	1	100
3	9	<1	100
4	0	0	100

**Table 5.A.13: Summary Statistics for CR item**

	Mean	Standard Deviation
First Rating	2.45	0.85
Second Rating	2.45	0.84
Mean Absolute Difference Between First and Second Ratings		0.35
Correlation of First and Second Ratings		0.75

**Summary of CR items Receiving Score of Zero**

CR item	
Blank	2168
Illegible	0
Off Topic	1588
Cartoon / Inappropriate	296
Not in English	88
<b>Total</b>	<b>4140</b>

**Table 5.A.14: NCLB Reliability Classifications: February 2004**

**English Language Arts  
Decision Accuracy**

Placement (Raw Score)	Classification on all-forms average*			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (76-90)	0.21	0.03	0.00	0.24
Proficient (68-75)	0.04	0.14	0.04	0.22
Below Proficient (0-67)	0.00	0.04	0.50	0.54
Estimated Proportion Correctly Classified Advanced				0.93
Estimated Proportion Correctly Classified Proficient				0.92

\*True Score

**Decision Consistency**

Placement (Raw Score)	Classification on alternate form			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (76-90)	0.20	0.05	0.00	0.24
Proficient (68-75)	0.05	0.11	0.05	0.22
Below Proficient (0-67)	0.00	0.05	0.49	0.54
Estimated Proportion Correctly Classified at Advanced				0.90
Estimated Proportion Correctly Classified at Proficient				0.90

**Mathematics  
Decision Accuracy**

Placement (Raw Score)	Classification on all-forms average*			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (72-80)	0.10	0.03	0.00	0.13
Proficient (58-71)	0.02	0.24	0.03	0.30
Below Proficient (0-57)	0.00	0.03	0.54	0.58
Estimated Proportion Correctly Classified at Advanced				0.96
Estimated Proportion Correctly Classified at Proficient				0.93

\*True Score

**Decision Consistency**

Placement (Raw Score)	Classification on alternate form			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (72-80)	0.10	0.03	0.00	0.13
Proficient (58-71)	0.03	0.22	0.05	0.30
Below Proficient (0-57)	0.00	0.05	0.53	0.58
Estimated Proportion Correctly Classified at Advanced				0.94
Estimated Proportion Correctly Classified at Proficient				0.90

**Table 5.A.15: Pass/Not Pass Classifications February 2004**

**English Language and Arts  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
54-90	0.70	0.03	0.73
0-53	0.03	0.24	0.27
Estimated Proportion Correctly Classified			0.95

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
54-90	0.69	0.04	0.73
0-53	0.04	0.23	0.27
Estimated Proportion Correctly Classified			0.92

**Mathematics  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
43-80	0.69	0.03	0.71
0-42	0.03	0.25	0.29
Estimated Proportion Correctly Classified			0.94

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
43-80	0.67	0.04	0.71
0-42	0.04	0.24	0.29
Estimated Proportion Correctly Classified			0.92

**Appendix 5.B: Demographic Summaries, Frequency Distributions, Intercorrelations,  
and Decision Consistencies—March 2004**

**Table 5.B.1: Demographic Summary For All Examinees <sup>1</sup> ELA March 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Applic.
							WA	RC	LR/A	WS	WC	CR item
<b>Total Examinees</b>	291382	221494	76%	69888	24%	379	79%	75%	76%	62%	71%	2.4
<b>Grade</b>												
Tenth	290474	221049	76%	69425	24%	379	79%	75%	76%	62%	71%	2.4
Eleventh	1	.	.	.	.	.	.	.	.	.	.	.
Twelfth	116	40	34%	76	66%	343	62%	58%	57%	43%	52%	1.9
Adult Education	536	279	52%	257	48%	352	71%	68%	65%	47%	55%	1.8
Unknown	255	126	49%	129	51%	347	64%	61%	60%	44%	54%	1.9
<b>Gender</b>												
Male	147676	106449	72%	41227	28%	374	79%	74%	74%	60%	68%	2.3
Female	143468	114932	80%	28536	20%	384	80%	76%	78%	64%	74%	2.5
Unknown	238	113	47%	125	53%	345	65%	61%	60%	45%	51%	1.7
<b>Race/Ethnicity</b>												
American Indian or Alaska Native	2460	1839	75%	621	25%	374	79%	74%	75%	60%	67%	2.3
Asian	28752	25032	87%	3720	13%	396	84%	81%	82%	72%	80%	2.7
Pacific Islander	1877	1356	72%	521	28%	373	76%	72%	74%	58%	69%	2.4
Filipino	8488	7475	88%	1013	12%	390	83%	79%	81%	68%	78%	2.6
Hispanic or Latino	120336	75888	63%	44448	37%	362	72%	68%	68%	53%	63%	2.1
African American	20839	13428	64%	7411	36%	363	74%	68%	70%	54%	63%	2.1
White (not of Hispanic origin)	106112	94661	89%	11451	11%	395	86%	83%	84%	71%	77%	2.6
Unknown	2518	1815	72%	703	28%	375	78%	74%	74%	61%	68%	2.3
<b>Language Fluency</b>												
English Only Students	177161	146863	83%	30298	17%	386	83%	79%	80%	66%	74%	2.5
Initially Fluent English Proficient (IFEP)	28203	24239	86%	3964	14%	390	83%	79%	81%	68%	77%	2.6
Redesignated Fluent English Proficient (RFEP)	32347	28076	87%	4271	13%	384	81%	78%	79%	66%	75%	2.5
English Learner Students	52327	21499	41%	30828	59%	343	62%	59%	58%	43%	54%	1.8
Unknown	1344	817	61%	527	39%	361	73%	69%	68%	53%	61%	2.0
<b>Economically Disadvantaged</b>												
No	128145	112608	88%	15537	12%	393	85%	81%	83%	70%	77%	2.6
Yes	115468	71246	62%	44222	38%	361	72%	67%	68%	53%	63%	2.1
Unknown	47769	37640	79%	10129	21%	382	81%	77%	78%	64%	72%	2.4
<b>Special Education Program Participation</b>												
Students Receiving Services	25779	8175	32%	17604	68%	335	59%	55%	54%	39%	47%	1.6
Students Not Receiving Services	265603	213319	80%	52284	20%	383	81%	77%	78%	64%	73%	2.5

<sup>1</sup> Results for groups with less than 11 members are not reported

**Table 5.B.2: Demographic Summary For All Examinees <sup>1</sup> Mathematics March 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>				
							Average Percent Correct				
							PS	NS	AF	MG	Alg1
<b>Total Examinees</b>	289231	219389	76%	69842	24%	379	71%	71%	69%	65%	59%
<b>Grade</b>											
Tenth	288364	218999	76%	69365	24%	379	71%	71%	69%	65%	59%
Eleventh	1	.	.	.	.	.	.	.	.	.	.
Twelfth	123	53	43%	70	57%	348	56%	54%	52%	48%	38%
Adult Education	519	235	45%	284	55%	350	58%	57%	54%	48%	37%
Unknown	224	102	46%	122	54%	349	54%	54%	53%	49%	42%
<b>Gender</b>											
Male	146141	110159	75%	35982	25%	380	71%	71%	69%	65%	58%
Female	142845	109105	76%	33740	24%	378	70%	70%	69%	64%	60%
Unknown	245	125	51%	120	49%	353	59%	56%	56%	49%	44%
<b>Race/Ethnicity</b>											
American Indian or Alaska Native	2414	1717	71%	697	29%	372	69%	68%	65%	61%	54%
Asian	28675	26757	93%	1918	7%	409	82%	84%	83%	80%	77%
Pacific Islander	1874	1373	73%	501	27%	374	68%	68%	67%	62%	57%
Filipino	8467	7451	88%	1016	12%	390	75%	76%	75%	71%	66%
Hispanic or Latino	119557	75954	64%	43603	36%	364	63%	63%	62%	56%	51%
African American	20574	11741	57%	8833	43%	358	61%	59%	59%	52%	48%
White (not of Hispanic origin)	105165	92616	88%	12549	12%	392	79%	78%	75%	73%	65%
Unknown	2505	1780	71%	725	29%	375	69%	68%	66%	63%	55%
<b>Language Fluency</b>											
English Only Students	175438	140613	80%	34825	20%	384	74%	73%	71%	67%	60%
Initially Fluent English Proficient (IFEP)	28130	23593	84%	4537	16%	390	75%	76%	74%	70%	65%
Redesignated Fluent English Proficient (RFEP)	32379	27098	84%	5281	16%	384	73%	73%	72%	68%	63%
English Learner Students	51955	27270	52%	24685	48%	356	56%	57%	59%	51%	48%
Unknown	1329	815	61%	514	39%	363	64%	63%	61%	56%	47%
<b>Economically Disadvantaged</b>											
No	127466	110041	86%	17425	14%	391	77%	77%	75%	72%	65%
Yes	114428	72454	63%	41974	37%	364	63%	63%	63%	56%	52%
Unknown	47337	36894	78%	10443	22%	382	73%	72%	70%	66%	60%
<b>Special Education Program Participation</b>											
Students Receiving Services	23016	7393	32%	15623	68%	341	50%	49%	48%	43%	38%
Students Not Receiving Services	266215	211996	80%	54219	20%	382	73%	72%	71%	67%	61%

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 5.B.3: NCLB Demographic Summary for All CAHSEE Examinees ELA March 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		291382	143677	49	61598	21	86107	30	147705	51
<b>Grade</b>	Tenth	290474	142942	49	61482	21	86050	30	147532	51
	Eleventh	1								
	Twelfth	116	98	84	13	11	5	4	18	16
	Adult Education	536	433	81	71	13	32	6	103	19
	Unknown	255	203	80	32	13	20	8	52	20
<b>Gender</b>	Male	147676	79623	54	30646	21	37407	25	68053	46
	Female	143468	63859	45	30923	22	48686	34	79609	55
	Unknown	238	195	82	29	12	14	6	43	18
<b>Race/Ethnicity</b>	American Indian or Alaskan Native	2460	1324	54	576	23	560	23	1136	46
	Asian	28752	9031	31	5773	20	13948	49	19721	69
	Pacific Islander	1877	1088	58	403	21	386	21	789	42
	Filipino	8488	3154	37	2155	25	3179	37	5334	63
	Hispanic or Latino	120336	81461	68	21913	18	16962	14	38875	32
	African American	20839	13721	66	3979	19	3139	15	7118	34
	White (not of Hispanic origin)	106112	32607	31	26313	25	47192	44	73505	69
	Unknown	2518	1291	51	486	19	741	29	1227	49
<b>Language Fluency</b>	English Only Students	177161	71779	41	41326	23	64056	36	105382	59
	Initially Fluent English Proficient (IFEP)	28203	10892	39	6578	23	10733	38	17311	61
	Redesignated Fluent English Proficient (RFEP)	32347	14434	45	8896	28	9017	28	17913	55
	English Learner Students	52327	45665	87	4584	9	2078	4	6662	13
	Unknown	1344	907	67	214	16	223	17	437	33
<b>Economic Disadvantage</b>	No	128145	42539	33	30729	24	54877	43	85606	67
	Yes	115468	79582	69	20230	18	15656	14	35886	31
	Unknown	47769	21556	45	10639	22	15574	33	26213	55
<b>Special Education Program Participation</b>	Receiving Services	25779	22975	89	1842	7	962	4	2804	11
	Not Receiving Services	265603	120702	45	59756	22	85145	32	144901	55

**Table 5.B.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics March 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		289231	151205	52	96311	33	41715	14	138026	48
<b>Grade</b>	Tenth	288364	150451	52	96220	33	41693	14	137913	48
	Eleventh	1								
	Twelfth	123	112	91	7	6	4	3	11	9
	Adult Education	519	452	87	56	11	11	2	67	13
	Unknown	224	189	84	28	13	7	3	35	16
<b>Gender</b>	Male	146141	74580	51	48313	33	23248	16	71561	49
	Female	142845	76425	54	47963	34	18457	13	66420	46
	Unknown	245	200	82	35	14	10	4	45	18
<b>Race/Ethnicity</b>	American Indian or Alaskan Native	2414	1432	59	787	33	195	8	982	41
	Asian	28675	6037	21	11077	39	11561	40	22638	79
	Pacific Islander	1874	1109	59	582	31	183	10	765	41
	Filipino	8467	3351	40	3568	42	1548	18	5116	60
	Hispanic or Latino	119557	84688	71	29392	25	5477	5	34869	29
	African American	20574	15511	75	4362	21	701	3	5063	25
	White (not of Hispanic origin)	105165	37646	36	45800	44	21719	21	67519	64
	Unknown	2505	1431	57	743	30	331	13	1074	43
<b>Language Fluency</b>	English Only Students	175438	81636	47	65397	37	28405	16	93802	53
	Initially Fluent English Proficient (IFEP)	28130	11878	42	9813	35	6439	23	16252	58
	Redesignated Fluent English Proficient (RFEP)	32379	15336	47	12161	38	4882	15	17043	53
	English Learner Students	51955	41426	80	8624	17	1905	4	10529	20
	Unknown	1329	929	70	316	24	84	6	400	30
<b>Economic Disadvantage</b>	No	127466	48771	38	51293	40	27402	21	78695	62
	Yes	114428	79597	70	28081	25	6750	6	34831	30
	Unknown	47337	22837	48	16937	36	7563	16	24500	52
<b>Special Education Program Participation</b>	Receiving Services	23016	20311	88	2299	10	406	2	2705	12
	Not Receiving Services	266215	130894	49	94012	35	41309	16	135321	51

**Table 5.B.5: Demographic Summary For All Examinees<sup>1</sup> ELA March 2004**

	Percentiles <sup>2</sup>							Mean Scale Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	286	308	351	381	407	445	450	379	40	291382
<b>Grade</b>										
Tenth	286	308	351	381	407	445	450	379	40	290474
Eleventh	.	.	.	.	.	.	.	.	.	0
Twelfth	.	.	.	.	.	.	.	.	.	0
Adult Education	276	302	329	351	374	404	423	352	32	536
Unknown	275	289	314	349	374	407	450	347	39	255
<b>Gender</b>										
Female	284	304	345	376	404	438	450	374	40	147676
Male	293	316	357	386	411	450	450	384	39	143468
Unknown	275	281	314	346	372	411	433	345	39	238
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	281	306	349	376	400	433	450	374	38	2460
Asian	296	325	372	400	428	450	450	396	39	28752
Pacific Islander	291	314	349	372	397	433	450	373	35	1877
Filipino	302	330	369	392	415	445	450	390	34	8488
Hispanic or Latino	281	300	336	363	389	423	445	362	37	120336
African American	279	300	338	365	389	423	445	363	37	20839
White (not of Hispanic origin)	300	330	374	397	419	450	450	395	35	106112
Unknown	275	298	345	379	407	445	450	375	43	2518
<b>Language Fluency</b>										
English Learner Students	291	316	363	389	415	450	450	386	38	177161
Initially Fluent English Proficient (IFEP)	298	327	365	392	415	450	450	390	37	28203
Redesignated Fluent English Proficient (RFEP)	304	332	363	384	404	438	450	384	32	32347
English Only Students	279	293	319	343	365	397	423	343	32	52327
Unknown	276	296	332	363	389	428	450	361	39	1344
<b>Economically Disadvantaged</b>										
No	300	329	372	394	419	450	450	393	36	128145
Yes	281	300	334	361	386	423	450	361	37	115468
Unknown	286	308	357	384	411	445	450	382	40	47769
<b>Special Education Program Participation</b>										
Students Receiving Services	275	286	310	332	359	397	428	335	34	25779
Students Not Receiving Services	293	316	357	384	411	445	450	383	38	265603

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.B.6: Demographic Summary For All Examinees<sup>1</sup> Mathematics March 2004**

	Percentiles <sup>2</sup>							Mean Scale Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	304	320	351	376	407	450	450	379	38	289231
<b>Grade</b>										
Tenth	304	320	351	376	407	450	450	379	38	288364
Eleventh	.	.	.	.	.	.	.	.	.	0
Twelfth	.	.	.	.	.	.	.	.	.	0
Adult Education	283	306	332	347	368	407	437	350	29	519
Unknown	292	306	326	347	366	413	437	349	32	224
<b>Gender</b>										
Female	304	318	351	378	407	450	450	380	39	146141
Male	306	322	351	376	403	450	450	378	36	142845
Unknown	275	306	328	351	372	417	443	353	33	245
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	304	315	347	370	395	437	450	372	35	2414
Asian	320	345	385	413	443	450	450	409	35	28675
Pacific Islander	304	320	349	372	398	437	450	374	35	1874
Filipino	318	334	366	390	413	450	450	390	34	8467
Hispanic or Latino	301	315	340	360	385	421	450	364	32	119557
African American	298	311	336	357	378	417	443	358	32	20574
White (not of Hispanic origin)	311	332	368	392	417	450	450	392	35	105165
Unknown	301	315	345	372	401	450	450	375	39	2505
<b>Language Fluency</b>										
English Only Students	306	322	357	383	410	450	450	384	37	175438
Initially Fluent English Proficient (IFEP)	311	330	360	387	421	450	450	390	38	28130
Redesignated Fluent English Proficient (RFEP)	313	332	360	383	407	450	450	384	34	32379
English Learner Students	298	311	334	351	374	417	450	356	31	51955
Unknown	295	311	338	360	385	431	450	363	35	1329
<b>Economically Disadvantaged</b>										
No	311	332	366	390	417	450	450	391	36	127466
Yes	301	315	340	360	385	426	450	364	34	114428
Unknown	304	320	355	381	410	450	450	382	38	47337
<b>Special Education Program Participation</b>										
Students Receiving Services	289	304	320	336	358	401	437	341	31	23016
Students Not Receiving Services	309	326	355	381	410	450	450	382	36	266215

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.B.7: Frequency Distributions ELA March 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	12801	4	12801	96
440-449	5216	2	18017	94
430-439	12417	4	30434	90
420-429	14892	5	45326	84
410-419	24187	8	69513	76
403-409	16594	6	86107	70
390-402	32190	11	118297	59
380-389	29408	10	147705	49
370-379	25725	9	173430	40
360-369	26868	9	200298	31
350-359 <sup>1</sup>	21196	7	221494	24
340-349	20118	7	241612	17
330-339	13061	4	254673	13
320-329	10541	4	265214	9
310-319	10804	4	276018	5
300-309	7317	3	283335	3
290-299	4180	1	287515	1
280-289	2281	1	289796	1
275-279	1586	1	291382	0

<sup>1</sup>Passing Line

**Table 5.B.8: Frequency Distributions Mathematics 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	17747	6	17747	94
440-449	5720	2	23467	92
430-439	12011	4	35478	88
420-429	12619	4	48097	83
410-419	19448	7	67545	77
400-409	19808	7	87353	70
390-399	25796	9	113149	61
380-389	24877	9	138026	52
370-379	29721	10	167747	42
360-369	27033	9	194780	33
350-359 <sup>1</sup>	24609	9	219389	24
340-349	25620	9	245009	15
330-339	17747	6	262756	9
320-329	13713	5	276469	4
310-319	7470	3	283939	2
300-309	3768	1	287707	1
290-299	989	0	288696	0
280-289	335	0	289031	0
275-279	200	0	289231	0

<sup>1</sup>Passing Line

**Table 5.B.9: Frequency Distributions ELA March 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	12801	4	12801	96
440-449	5216	2	18017	94
430-439	12417	4	30434	90
420-429	14892	5	45326	84
410-419	24187	8	69513	76
403-409 <sup>1</sup>	16594	6	86107	70
390-402	32190	11	118297	59
380-389 <sup>2</sup>	29408	10	147705	49
370-379	25725	9	173430	40
360-369	26868	9	200298	31
350-359	21196	7	221494	24
340-349	20118	7	241612	17
330-339	13061	4	254673	13
320-329	10541	4	265214	9
310-319	10804	4	276018	5
300-309	7317	3	283335	3
290-299	4180	1	287515	1
280-289	2281	1	289796	1
275-279	1586	1	291382	0

<sup>1</sup>Advanced Level Cut

<sup>2</sup>Proficient Level Cut

**Table 5.B.10: Frequency Distributions Mathematics March 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	17747	6	17747	94
440-449	5720	2	23467	92
430-439	12011	4	35478	88
422-429 <sup>1</sup>	6237	2	41715	86
410-421	25830	9	67545	77
400-409	19808	7	87353	70
390-399	25796	9	113149	61
380-389 <sup>2</sup>	24877	9	138026	52
370-379	29721	10	167747	42
360-369	27033	9	194780	33
350-359	24609	9	219389	24
340-349	25620	9	245009	15
330-339	17747	6	262756	9
320-329	13713	5	276469	4
310-319	7470	3	283939	2
300-309	3768	1	287707	1
290-299	989	0	288696	0
280-289	335	0	289031	0
275-279	200	0	289231	0

<sup>1</sup>Advanced Level Cut

<sup>2</sup>Proficient Level Cut

**Table 5.B.11: Intercorrelations<sup>1</sup> and Reliability Estimates By Section March 2004**

	ELA <sup>4</sup>	CR item	CR item	RW	RWC	RCRL	RLWS	WSCW	Math	Math	PS	PS	NS	NSAF	AFMG	MGA1	A1
ELA <sup>2</sup>	1.00	0.82	0.77	0.89	0.92	0.85	0.86	0.78	0.71	0.69	0.71	0.70	0.62				
CR item	0.82	1.00	0.56	0.64	0.67	0.61	0.64	0.62	0.55	0.54	0.57	0.56	0.51				
Word Analysis	0.77	0.56	1.00	0.68	0.70	0.60	0.61	0.59	0.56	0.54	0.53	0.53	0.45				
Reading Comprehension	0.89	0.64	0.68	1.00	0.80	0.71	0.69	0.68	0.64	0.61	0.61	0.61	0.53				
Literary Responses & Analysis	0.92	0.67	0.70	0.80	1.00	0.74	0.72	0.69	0.64	0.62	0.63	0.62	0.54				
Writing Strategies	0.85	0.61	0.60	0.71	0.74	1.00	0.71	0.68	0.61	0.59	0.62	0.62	0.56				
Writing Conventions	0.86	0.64	0.61	0.69	0.72	0.71	1.00	0.69	0.62	0.61	0.64	0.62	0.57				
Math <sup>3</sup>	0.78	0.62	0.59	0.68	0.69	0.68	0.69	1.00	0.85	0.89	0.91	0.91	0.84				
Probability and Statistics	0.71	0.55	0.56	0.64	0.64	0.61	0.62	0.85	1.00	0.74	0.72	0.72	0.62				
Number Sense	0.69	0.54	0.54	0.61	0.62	0.59	0.61	0.89	0.74	1.00	0.76	0.75	0.68				
Algebra & Functions	0.71	0.57	0.53	0.61	0.63	0.62	0.64	0.91	0.72	0.76	1.00	0.77	0.73				
Measurement & Geometry	0.70	0.56	0.53	0.61	0.62	0.62	0.62	0.91	0.72	0.75	0.77	1.00	0.72				
Algebra I	0.62	0.51	0.45	0.53	0.54	0.56	0.57	0.84	0.62	0.68	0.73	0.72	1.00				
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12				
Mean	63.07	2.40	5.54	13.52	15.18	7.45	10.58	53.80	9.20	12.00	13.86	11.65	7.08				
SD	16.52	0.81	1.50	3.53	4.13	3.04	3.36	15.60	2.83	3.71	3.98	4.19	2.92				
Reliability	0.95	0.77	0.63	0.84	0.79	0.78	0.78	0.95	0.75	0.80	0.80	0.83	0.74				
SEM	3.73	0.39	0.91	1.41	1.89	1.42	1.56	3.64	1.42	1.66	1.78	1.70	1.49				

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WC=Writing Conventions, WS=Writing Strategies, A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

<sup>1</sup> All correlations are for raw scores.

<sup>2</sup> Correlations for ELA section are reported for examinees completing ELA section.

Number of examinees: 291382

<sup>3</sup> Correlations for Math section are reported for examinees completing Math section.

Number of examinees: 288990

<sup>4</sup> Correlations between ELA and Math are reported for examinees completing both sections.

Number of examinees: 280831

**Table 5.B.12: Agreement of First and Second Ratings on CR item**

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	6102	0	0	0	0	6102	2
1	0	20149	11086	711	23	31969	11
2	0	10799	76148	30309	1394	118650	41
3	0	628	29789	66883	12261	109561	38
4	0	20	1426	12591	11063	25100	9
Total	6102	31596	118449	110494	24741	291382	100
Percent	2	11	41	38	8	100	

**Differences Between First and Second Ratings on CR item**

Difference	Frequency	Percent	Cumulative Percent
0	180345	62	62
1	106835	37	99
2	4159	1	100
3	43	<1	100
4	0	0	100

**Table 5.B.13: Summary Statistics for CR item**

	Mean	Standard Deviation
First Rating	2.40	0.87
Second Rating	2.40	0.87
Mean Absolute Difference Between First and Second Ratings		0.40
Correlation of First and Second Ratings		0.72

**Summary of CR items Receiving Score of Zero**

CR item	
Blank	2893
Illegible	1
Off Topic	2817
Cartoon / Inappropriate	358
Not in English	38
<b>Total</b>	<b>6107</b>

**Table 5.B.14: NCLB Classifications March 2004**

**ELA  
Decision Accuracy**

Placement (Raw Score)	Classification on all-forms average*			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (75-90)	0.26	0.03	0.00	0.30
Proficient (67-74)	0.04	0.13	0.04	0.21
Below Proficient (0-66)	0.00	0.03	0.46	0.49
Estimated Proportion Correctly Classified as Advanced				0.93
Estimated Proportion Correctly Classified as Proficient				0.93

\* True Score

**Decision Consistency**

Placement (Raw Score)	Classification on alternate form			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (75-90)	0.25	0.04	0.00	0.30
Proficient (67-74)	0.05	0.11	0.05	0.21
Below Proficient (0-66)	0.00	0.04	0.45	0.49
Estimated Proportion Correctly Classified as Advanced				0.90
Estimated Proportion Correctly Classified as Proficient				0.90

**Mathematics  
Decision Accuracy**

Placement (Raw Score)	Classification on all-forms average*			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (71-80)	0.14	0.03	0.00	0.16
Proficient (57-70)	0.02	0.25	0.03	0.31
Below Proficient (0-56)	0.00	0.03	0.50	0.53
Estimated Proportion Correctly Classified as Advanced				0.95
Estimated Proportion Correctly Classified as Proficient				0.93

\*True Score

**Decision Consistency**

Placement (Raw Score)	Classification on alternate form			
	Category Advanced	Category Proficient	Category Below Prof.	Category Total
Advanced (71-80)	0.13	0.03	0.00	0.16
Proficient (57-70)	0.03	0.23	0.05	0.31
Below Proficient (0-56)	0.00	0.05	0.48	0.53
Estimated Proportion Correctly Classified as Advanced				0.93
Estimated Proportion Correctly Classified as Proficient				0.91

**Table 5.B.15: Pass/Not Pass Classifications March 2004**

**ELA  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
52-90	0.75	0.02	0.77
0-51	0.02	0.20	0.23
Estimated Proportion Correctly Classified			0.95

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
52-90	0.74	0.04	0.77
0-51	0.03	0.20	0.23
Estimated Proportion Correctly Classified			0.93

**Mathematics  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
42-80	0.73	0.03	0.75
0-41	0.03	0.22	0.25
Estimated Proportion Correctly Classified			0.95

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
42-80	0.71	0.04	0.75
0-41	0.04	0.21	0.25
Estimated Proportion Correctly Classified			0.92

**Appendix 5.C: Demographic Summaries, Frequency Distributions, Intercorrelations,  
and Decision Consistencies—May 2004**

**Table 5.C.1: Demographic Summary For All Examinees<sup>1</sup> English Language Arts May 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Appl Avg. Score
							WA	RC	LR/A	WS	WC	CR item
<b>Total Examinees</b>	12016	5927	49%	6089	51%	351	70%	60%	63%	51%	59%	2.3
<b>Grade</b>												
Tenth	11929	5889	49%	6040	51%	351	70%	60%	63%	51%	59%	2.3
Eleventh	0	.	.	.	.	.	.	.	.	.	.	.
Twelfth	0	.	.	.	.	.	.	.	.	.	.	.
Adult Education	46	23	50%	23	50%	349	74%	67%	70%	50%	50%	2.1
Unknown	41	15	37%	26	63%	334	60%	52%	54%	40%	48%	2.2
<b>Gender</b>												
Male	6344	2750	43%	3594	57%	344	67%	57%	60%	46%	54%	2.2
Female	5608	3153	56%	2455	44%	359	73%	63%	66%	56%	63%	2.5
Unknown	64	24	38%	40	63%	339	65%	52%	57%	48%	50%	2.1
<b>Race/Ethnicity</b>												
American Indian or Alaska Native	188	99	53%	89	47%	352	73%	63%	65%	51%	58%	2.3
Asian	489	217	44%	272	56%	349	64%	58%	60%	51%	58%	2.2
Pacific Islander	108	41	38%	67	62%	344	66%	55%	57%	45%	55%	2.3
Filipino	219	140	64%	79	36%	367	78%	68%	70%	62%	70%	2.6
Hispanic or Latino	5535	2068	37%	3467	63%	339	63%	54%	57%	44%	53%	2.2
African American	1398	515	37%	883	63%	339	65%	54%	57%	44%	52%	2.2
White (not of Hispanic origin)	3784	2689	71%	1095	29%	372	82%	71%	73%	62%	68%	2.7
Unknown	295	158	54%	137	46%	352	71%	61%	64%	53%	59%	2.2
<b>Language Fluency</b>												
English Only Students	7545	4338	57%	3207	43%	358	75%	64%	67%	55%	62%	2.5
Initially Fluent English Proficient (IFEP)	794	424	53%	370	47%	355	72%	62%	65%	53%	62%	2.5
Redesignated Fluent English Proficient (RFEP)	787	505	64%	282	36%	361	77%	65%	69%	57%	66%	2.6
English Learner Students	2515	461	18%	2054	82%	323	52%	45%	48%	36%	45%	1.8
Unknown	375	199	53%	176	47%	356	72%	62%	64%	55%	59%	2.4
<b>Economically Disadvantaged</b>												
No	4061	2728	67%	1333	33%	368	80%	69%	71%	60%	67%	2.7
Yes	5733	2072	36%	3661	64%	338	63%	53%	56%	44%	52%	2.1
Unknown	2222	1127	51%	1095	49%	353	71%	61%	63%	52%	59%	2.3
<b>Special Education Program Participation</b>												
Students Receiving Services	1544	259	17%	1285	83%	319	52%	43%	46%	33%	40%	1.7
Students Not Receiving Services	10472	5668	54%	4804	46%	356	73%	63%	65%	53%	61%	2.4

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> WA — Word Analysis, RC — Reading Comprehension, LR/A — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

**Table 5.C.2: Demographic Summary For All Examinees<sup>1</sup> Mathematics May 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>				
							Average Percent Correct				
							PS	NS	AF	MG	Alg1
<b>Total Examinees</b>	11377	5190	46%	6187	54%	350	58%	57%	53%	51%	39%
<b>Grade</b>											
Tenth	11287	5158	46%	6129	54%	350	58%	57%	53%	51%	39%
Eleventh	0	.	.	.	.	.	.	.	.	.	.
Twelfth	0	.	.	.	.	.	.	.	.	.	.
Adult Education	49	20	41%	29	59%	345	59%	56%	51%	50%	30%
Unknown	41	12	29%	29	71%	337	45%	51%	43%	46%	36%
<b>Gender</b>											
Male	5872	2544	43%	3328	57%	348	57%	56%	52%	50%	38%
Female	5451	2630	48%	2821	52%	352	60%	58%	55%	52%	41%
Unknown	54	16	30%	38	70%	338	53%	50%	46%	43%	31%
<b>Race/Ethnicity</b>											
American Indian or Alaska Native	165	73	44%	92	56%	347	59%	55%	52%	50%	37%
Asian	446	304	68%	142	32%	370	64%	68%	66%	62%	55%
Pacific Islander	105	43	41%	62	59%	346	55%	54%	51%	49%	38%
Filipino	209	128	61%	81	39%	364	65%	64%	63%	59%	49%
Hispanic or Latino	5325	1827	34%	3498	66%	341	52%	52%	48%	46%	35%
African American	1291	374	29%	917	71%	336	51%	50%	45%	42%	32%
White (not of Hispanic origin)	3545	2301	65%	1244	35%	365	69%	66%	62%	60%	46%
Unknown	291	140	48%	151	52%	351	59%	59%	53%	52%	40%
<b>Language Fluency</b>											
English Only Students	7141	3618	51%	3523	49%	354	62%	59%	55%	53%	41%
Initially Fluent English Proficient (IFEP)	756	366	48%	390	52%	352	59%	58%	55%	52%	39%
Redesignated Fluent English Proficient (RFEP)	771	437	57%	334	43%	357	64%	61%	59%	56%	43%
English Learner Students	2376	591	25%	1785	75%	335	45%	48%	44%	43%	33%
Unknown	333	178	53%	155	47%	356	61%	61%	56%	55%	42%
<b>Economically Disadvantaged</b>											
No	3737	2310	62%	1427	38%	363	67%	64%	61%	59%	46%
Yes	5572	1900	34%	3672	66%	341	52%	52%	48%	46%	35%
Unknown	2068	980	47%	1088	53%	351	59%	58%	54%	52%	39%
<b>Special Education Program Participation</b>											
Students Receiving Services	1340	195	15%	1145	85%	325	41%	42%	37%	36%	28%
Students Not Receiving Services	10037	4995	50%	5042	50%	353	61%	59%	55%	53%	41%

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 5.C.3: NCLB Demographic Summary for All CAHSEE Examinees English Language Arts May 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		12016	8830	73	1707	14	1479	12	3186	27
<b>Grade</b>	Tenth	11929	8755	73	1697	14	1477	12	3174	27
	Eleventh	0	.	.	.	.	.	.	.	.
	Twelfth	0	.	.	.	.	.	.	.	.
	Adult Education	46	38	83	6	13	2	4	8	17
	Unknown	41	37	90	4	10	0	0	4	10
<b>Gender</b>	Male	6344	4995	79	782	12	567	9	1349	21
	Female	5608	3785	67	916	16	907	16	1823	33
	Unknown	64	50	78	9	14	5	8	14	22
<b>Race/Ethnicity</b>	American Indian or Alaskan Native	188	137	73	35	19	16	9	51	27
	Asian	489	357	73	55	11	77	16	132	27
	Pacific Islander	108	84	78	13	12	11	10	24	22
	Filipino	219	126	58	52	24	41	19	93	42
	Hispanic or Latino	5535	4703	85	567	10	265	5	832	15
	African American	1398	1194	85	134	10	70	5	204	15
	White (not of Hispanic origin)	3784	2020	53	811	21	953	25	1764	47
	Unknown	295	209	71	40	14	46	16	86	29
<b>Language Fluency</b>	English Only Students	7545	5056	67	1302	17	1187	16	2489	33
	Initially Fluent English Proficient (IFEP)	794	573	72	116	15	105	13	221	28
	Redesignated Fluent English Proficient (RFEP)	787	543	69	154	20	90	11	244	31
	English Learner Students	2515	2407	96	84	3	24	1	108	4
	Unknown	375	251	67	51	14	73	19	124	33
<b>Economic Disadvantaged</b>	No	4061	2330	57	829	20	902	22	1731	43
	Yes	5733	4924	86	555	10	254	4	809	14
	Unknown	2222	1576	71	323	15	323	15	646	29
<b>Special Education Program Participation</b>	Receiving Services	1544	1465	95	57	4	22	1	79	5
	Not Receiving Services	10472	7365	70	1650	16	1457	14	3107	30

**Table 5.C.4: NCLB Demographic Summary for All CAHSEE Examinees Mathematics May 2004**

Subgroup	Value	N Tested	N Below Proficient	% Below Proficient	N Proficient	% Proficient	N Advanced	% Advanced	N Above Proficient	% Above Proficient
<b>Total Examinees</b>		11377	9094	80	1864	16	419	4	2283	20
<b>Grade</b>	Tenth	11287	9011	80	1857	16	419	4	2276	20
	Eleventh	0	.	.	.	.	.	.	.	.
	Twelfth	0	.	.	.	.	.	.	.	.
	Adult Education	49	45	92	4	8	0	0	4	8
	Unknown	41	38	93	3	7	0	0	3	7
<b>Gender</b>	Male	5872	4714	80	917	16	241	4	1158	20
	Female	5451	4332	79	941	17	178	3	1119	21
	Unknown	54	48	89	6	11	0	0	6	11
<b>Race/Ethnicity</b>	American Indian or Alaskan Native	165	139	84	23	14	3	2	26	16
	Asian	446	254	57	139	31	53	12	192	43
	Pacific Islander	105	92	88	11	10	2	2	13	12
	Filipino	209	140	67	58	28	11	5	69	33
	Hispanic or Latino	5325	4757	89	504	9	64	1	568	11
	African American	1291	1191	92	94	7	6	0	100	8
	White (not of Hispanic origin)	3545	2298	65	979	28	268	8	1247	35
	Unknown	291	223	77	56	19	12	4	68	23
<b>Language Fluency</b>	English Only Students	7141	5437	76	1371	19	333	5	1704	24
	Initially Fluent English Proficient (IFEP)	756	605	80	123	16	28	4	151	20
	Redesignated Fluent English Proficient (RFEP)	771	601	78	145	19	25	3	170	22
	English Learner Students	2376	2216	93	148	6	12	1	160	7
	Unknown	333	235	71	77	23	21	6	98	29
<b>Economic Disadvantage</b>	No	3737	2512	67	960	26	265	7	1225	33
	Yes	5572	4992	90	515	9	65	1	580	10
	Unknown	2068	1590	77	389	19	89	4	478	23
<b>Special Education Program Participation</b>	Receiving Services	1340	1289	96	43	3	8	1	51	4
	Not Receiving Services	10037	7805	78	1821	18	411	4	2232	22

**Table 5.C.5: Demographic Summary For All Examinees<sup>1</sup> English Language Arts May 2004**

	Percentiles <sup>2</sup>							Mean Scale Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	275	289	318	348	380	421	450	351	41	12016
<b>Grade</b>										
Tenth	275	289	318	348	380	421	450	351	41	11929
Eleventh	.	.	.	.	.	.	.	.	.	0
Twelfth	.	.	.	.	.	.	.	.	.	0
Adult Education	275	306	332	349	368	400	421	349	32	46
Unknown	287	299	309	330	358	385	391	334	30	41
<b>Gender</b>										
Female	275	285	311	341	372	416	447	344	40	6344
Male	277	297	327	356	388	432	450	359	41	5608
Unknown	275	283	309	332	374	404	416	339	39	64
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	275	287	324	350	382	416	447	352	38	188
Asian	275	283	311	341	385	439	450	349	47	489
Pacific Islander	275	283	309	339	368	426	450	344	43	108
Filipino	277	299	337	370	394	432	450	367	40	219
Hispanic or Latino	275	285	311	335	364	400	426	339	36	5535
African American	275	287	311	337	362	404	426	339	36	1398
White (not of Hispanic origin)	275	300	343	375	404	439	450	372	42	3784
Unknown	275	283	314	352	385	426	450	352	45	295
<b>Language Fluency</b>										
English Only Students	275	293	327	358	388	426	450	358	42	7545
Initially Fluent English Proficient (IFEP)	275	297	327	352	382	421	450	355	39	794
Redesignated Fluent English Proficient (RFEP)	279	304	337	364	385	412	439	361	34	787
English Learner Students	275	279	302	321	341	377	400	323	29	2515
Unknown	275	287	316	352	391	432	450	356	47	375
<b>Economically Disadvantaged</b>										
No	277	299	339	370	397	439	450	368	41	4061
Yes	275	285	309	334	362	400	426	338	36	5733
Unknown	275	287	318	350	385	426	450	353	43	2222
<b>Special Education Program Participation</b>										
Students Receiving Services	275	277	297	313	337	380	408	319	31	1544
Students Not Receiving Services	275	293	325	354	385	426	450	356	41	10472

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.C.6: Demographic Summary For All Examinees<sup>1</sup> Mathematics May 2004**

	Percentiles <sup>2</sup>							Mean Scale Score	SD <sup>3</sup>	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Total Examinees</b>	287	304	325	345	371	414	448	350	34	11377
<b>Grade</b>										
Tenth	287	304	325	345	371	414	448	350	34	11287
Eleventh	.	.	.	.	.	.	.	.	.	0
Twelfth	.	.	.	.	.	.	.	.	.	0
Adult Education	293	301	330	345	361	382	395	345	23	49
Unknown	293	304	317	334	352	380	418	337	26	41
<b>Gender</b>										
Female	284	301	321	343	371	418	448	348	35	5872
Male	290	306	327	349	373	410	440	352	33	5451
Unknown	287	293	315	332	363	387	410	338	30	54
<b>Race/Ethnicity</b>										
American Indian or Alaska Native	287	308	325	343	365	397	433	347	29	165
Asian	281	308	340	367	397	448	450	370	41	446
Pacific Islander	275	299	323	343	365	400	423	346	32	105
Filipino	304	310	341	363	389	423	448	364	34	209
Hispanic or Latino	287	301	319	338	358	392	423	341	29	5325
African American	277	299	317	332	354	384	410	336	27	1291
White (not of Hispanic origin)	290	308	338	363	389	428	450	365	36	3545
Unknown	275	301	323	347	378	414	448	351	36	291
<b>Language Fluency</b>										
English Only Students	287	304	327	350	378	418	448	354	36	7141
Initially Fluent English Proficient (IFEP)	293	306	327	349	371	410	448	352	33	756
Redesignated Fluent English Proficient (RFEP)	299	310	336	356	375	414	448	357	30	771
English Learner Students	281	301	317	330	349	384	410	335	26	2376
Unknown	275	301	327	354	384	423	450	356	38	333
<b>Economically Disadvantaged</b>										
No	293	310	336	361	387	428	450	363	36	3737
Yes	284	301	319	338	358	392	423	341	29	5572
Unknown	287	304	323	347	375	418	448	351	36	2068
<b>Special Education Program Participation</b>										
Students Receiving Services	275	293	310	321	336	373	403	325	25	1340
Students Not Receiving Services	293	306	329	349	375	418	448	353	34	10037

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> Mean scale score are reported at each percentile

<sup>3</sup> SD - Standard Deviation

**Table 5.C.7: Frequency Distributions CAHSEE ELA May 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	155	1	155	99
440-449	74	1	229	98
430-439	207	2	436	96
420-429	285	2	721	94
410-419	367	3	1088	91
400-409	590	5	1678	86
390-399	660	5	2338	81
380-389	848	7	3186	73
370-379	844	7	4030	66
360-369	923	8	4953	59
350-359 <sup>1</sup>	974	8	5927	51
340-349	940	8	6867	43
330-339	1092	9	7959	34
320-329	1023	9	8982	25
310-319	768	6	9750	19
300-309	991	8	10741	11
290-299	600	5	11341	6
280-289	367	3	11708	3
275-279	308	3	12016	0

<sup>1</sup>Passing Line

**Table 5.C.8: Frequency Distributions CAHSEE Mathematics May 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	83	1	83	99
440-449	106	1	189	98
430-439	70	1	259	98
420-429	160	1	419	96
410-419	286	3	705	94
400-409	347	3	1052	91
390-399	409	4	1461	87
380-389	822	7	2283	80
370-379	699	6	2982	74
360-369	1139	10	4121	64
350-359 <sup>1</sup>	1069	9	5190	54
340-349	1368	12	6558	42
330-339	1218	11	7776	32
320-329	1293	11	9069	20
310-319	1274	11	10343	9
300-309	654	6	10997	3
290-299	250	2	11247	1
280-289	57	1	11304	1
275-279	73	1	11377	0

<sup>1</sup> Passing Line

**Table 5.C.9: Frequency Distributions CAHSEE ELA May 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	155	1	155	99
440-449	74	1	229	98
430-439	207	2	436	96
420-429	285	2	721	94
410-419	367	3	1088	91
403-409 <sup>1</sup>	391	3	1479	88
390-402	859	7	2338	81
380-389 <sup>2</sup>	848	7	3186	73
370-379	844	7	4030	66
360-369	923	8	4953	59
350-359	974	8	5927	51
340-349	940	8	6867	43
330-339	1092	9	7959	34
320-329	1023	9	8982	25
310-319	768	6	9750	19
300-309	991	8	10741	11
290-299	600	5	11341	6
280-289	367	3	11708	3
275-279	308	3	12016	0

<sup>1</sup> Advanced Level Cut

<sup>2</sup> Proficient Level Cut

**Table 5.C.10: Frequency Distributions: CAHSEE Mathematics May 2004**

Scale Score	Frequency	Percent	Cumulative Frequency	Percent Below
450	83	1	83	99
440-449	106	1	189	98
430-439	70	1	259	98
422-429 <sup>1</sup>	160	1	419	96
410-421	286	3	705	94
400-409	347	3	1052	91
390-399	409	4	1461	87
380-389 <sup>2</sup>	822	7	2283	80
370-379	699	6	2982	74
360-369	1139	10	4121	64
350-359	1069	9	5190	54
340-349	1368	12	6558	42
330-339	1218	11	7776	32
320-329	1293	11	9069	20
310-319	1274	11	10343	9
300-309	654	6	10997	3
290-299	250	2	11247	1
280-289	57	1	11304	1
275-279	73	1	11377	0

<sup>1</sup> Advanced Level Cut

<sup>2</sup> Proficient Level Cut

**Table 5.C.11: Intercorrelations<sup>1</sup> and Reliability Estimates By Section May 2004**

	ELA <sup>4</sup>	CR item	RW	RC	RL	WS	WC	Math	PS	NS	AF	MG	A1
ELA <sup>2</sup>	1.00	0.79	0.82	0.91	0.91	0.86	0.85	0.74	0.71	0.64	0.69	0.64	0.52
CR item	0.79	1.00	0.59	0.62	0.64	0.59	0.59	0.57	0.55	0.49	0.53	0.48	0.39
Word Analysis	0.82	0.59	1.00	0.75	0.76	0.65	0.65	0.57	0.58	0.51	0.53	0.49	0.37
Reading Comprehension	0.91	0.62	0.75	1.00	0.82	0.75	0.71	0.68	0.66	0.59	0.63	0.59	0.47
Literary Responses & Analysis	0.91	0.64	0.76	0.82	1.00	0.73	0.70	0.67	0.66	0.59	0.62	0.57	0.46
Writing Strategies	0.86	0.59	0.65	0.75	0.73	1.00	0.74	0.66	0.62	0.57	0.61	0.57	0.49
Writing Conventions	0.85	0.59	0.65	0.71	0.70	0.74	1.00	0.65	0.60	0.55	0.60	0.57	0.47
Math <sup>3</sup>	0.74	0.57	0.57	0.68	0.67	0.66	0.65	1.00	0.85	0.86	0.92	0.89	0.79
Probability and Statistics	0.71	0.55	0.58	0.66	0.66	0.62	0.60	0.85	1.00	0.71	0.74	0.69	0.55
Number Sense	0.64	0.49	0.51	0.59	0.59	0.57	0.55	0.86	0.71	1.00	0.73	0.68	0.61
Algebra & Functions	0.69	0.53	0.53	0.63	0.62	0.61	0.60	0.92	0.74	0.73	1.00	0.76	0.68
Measurement & Geometry	0.64	0.48	0.49	0.59	0.57	0.57	0.57	0.89	0.69	0.68	0.76	1.00	0.67
Algebra 1	0.52	0.39	0.37	0.47	0.46	0.49	0.47	0.79	0.55	0.61	0.68	0.67	1.00
Number of Items	73	1	7	18	20	12	15	80	13	17	20	18	12
Mean	53.67	2.34	4.90	10.81	12.53	6.10	8.79	41.81	7.59	9.67	10.64	9.21	4.70
SD	19.41	0.93	1.88	4.43	4.70	3.30	4.01	16.08	3.24	3.69	4.55	4.23	2.76
Reliability	0.95	0.72	0.70	0.83	0.84	0.80	0.84	0.94	0.78	0.75	0.81	0.82	0.70
SEM	4.21	0.49	1.02	1.80	1.87	1.48	1.61	3.93	1.51	1.84	1.96	1.81	1.52

Note. RC=Reading Comprehension, RL=Literary Response& Analysis, RW=Word Analysis; WC=Writing Conventions, WS=Writing Strategies, A1=Algebra I, AF=Algebra & Functions, MG=Measurement& Geometry, MR=Mathematical Reasoning, NS = Number Sense, PS=Probability & Statistics.

<sup>1</sup> All correlations are for raw scores.

<sup>2</sup> Correlations for ELA section are reported for examinees completing ELA section.

<sup>3</sup> Correlations for Math section are reported for examinees completing Math section.

<sup>4</sup> Correlations between ELA and Math are reported for examinees completing both sections.

Number of examinees: 12016

Number of examinees: 11370

Number of examinees: 7015

**Table 5.C.12: Agreement of First and Second Ratings on CR item**

(Cell Entry = Number of Examinee Responses)

First Rating	Second Rating					Overall	
	0	1	2	3	4	Total	Percent
0	780	0	0	0	0	780	6
1	0	503	415	18	0	936	8
2	0	323	3438	1073	30	4864	40
3	0	19	939	2903	500	4361	36
4	0	0	32	521	522	1075	9
Total	780	845	4824	4515	1052	12016	100
Percent	6	7	40	38	9	100	

**Differences Between First and Second Ratings on CR item**

Difference	Frequency	Percent	Cumulative Percent
0	8146	68	68
1	3771	31	99
2	99	1	100
3	0	0	100
4	0	0	100

**Table 5.C.13: Summary Statistics for CR item**

	Mean	Standard Deviation
First Rating	2.35	0.97
Second Rating	2.35	0.97
Mean Absolute Difference Between First and Second Ratings		0.33
Correlation of First and Second Ratings		0.82

**Summary of CR items Receiving Score of Zero**

CR item	
Blank	410
Illegible	0
Off Topic	307
Cartoon / Inappropriate	53
Not in English	12
<b>Total</b>	<b>782</b>

**Table 5.C.14: NCLB Reliability Classifications May 2004**

**ELA  
Decision Accuracy**

Placement (Raw Score)	Classification on all-forms average*			
	Category	Category	Category	Category
	Advanced	Proficient	Below Prof	Total
Advanced (78-90)	0.10	0.02	0.00	0.12
Proficient (70-77)	0.02	0.09	0.03	0.14
Below Proficient (0-69)	0.00	0.02	0.71	0.73
Estimated Proportion Correctly Classified at Advanced				0.96
Estimated Proportion Correctly Classified at Proficient				0.94

\*True Score

**Decision Consistency**

Placement (Raw Score)	Classification on alternate form			
	Category	Category	Category	Category
	Advanced	Proficient	Below Prof	Total
Advanced (78-90)	0.10	0.03	0.00	0.12
Proficient (70-77)	0.03	0.07	0.04	0.14
Below Proficient (0-69)	0.00	0.03	0.70	0.73
Estimated Proportion Correctly Classified at Advanced				0.94
Estimated Proportion Correctly Classified at Proficient				0.92

**Mathematics  
Decision Accuracy**

Placement Score	Classification on alternate form			
	Category	Category	Category	Category
	Advanced	Proficient	Below Prof	Total
Advanced (72-80)	0.03	0.01	0.00	0.04
Proficient (58-71)	0.01	0.13	0.03	0.16
Below Proficient (0-57)	0.00	0.02	0.78	0.8
Estimated Proportion Correctly Classified at Advanced				0.98
Estimated Proportion Correctly Classified at Proficient				0.95

\*True Score

**Decision Consistency**

Placement Score	Classification on alternate form			
	Category	Category	Category	Category
	Advanced	Proficient	Below Prof	Total
Advanced (72-80)	0.03	0.01	0.00	0.04
Proficient (58-71)	0.01	0.12	0.04	0.16
Below Proficient (0-57)	0.00	0.03	0.77	0.8
Estimated Proportion Correctly Classified at Advanced				0.98
Estimated Proportion Correctly Classified at Proficient				0.94

**Table 5.C.15: Pass/Not Pass Reliability Classifications May 2004**

**ELA  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
56-90	0.46	0.03	0.49
0-55	0.03	0.48	0.51
Estimated Proportion Correctly Classified			0.94

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
56-90	0.45	0.04	0.49
0-55	0.04	0.46	0.51
Estimated Proportion Correctly Classified			0.91

**Mathematics  
Decision Accuracy**

Classification on all-forms average*			
Placement Score	Category Pass	Category Not Pass	Category Total
43-80	0.42	0.03	0.46
0-42	0.03	0.51	0.54
Estimated Proportion Correctly Classified			0.93

\*True Score

**Decision Consistency**

Classification on alternate form			
Placement Score	Category Pass	Category Not Pass	Category Total
43-80	0.41	0.05	0.46
0-42	0.05	0.49	0.54
Estimated Proportion Correctly Classified			0.90

## Chapter 6: Special Analyses

A number of special analyses were conducted for the CAHSEE based on data from the February, March and May 2004 administrations and are summarized in this chapter. In particular, the chapter describes results of summary analyses of student results based on disabilities and special accommodations, generalizability analyses of rater scores on the three CR items, and mark discrimination (or erasure) analyses.

Following the convention of previous chapters, the tables summarizing student results based on disabilities and special accommodations and mark discrimination analyses are listed in Appendix 6.A to 6.C for the February, March and May 2004 administrations. The same table numbering system applies in Appendices 6.A to 6.C.

### Disability/Accommodations Summary

Scores for examinees who reported having a disability or who took CAHSEE with an accommodation or modification were analyzed to investigate the relationships between CAHSEE scores and disability, accommodation, language fluency, and special program participation. Modifications or accommodations for the CAHSEE include the use of a calculator for the mathematics section and the use of a reader or audio presentation for the ELA section. In Appendices 6A to 6C summary statistics are provided in Tables 1 and 2, for ELA and mathematics, respectively, for each disability type and for all accommodated students, as well as separately for standard accommodations versus modifications. It is of note that the largest disability group reported having a specific learning disability. The number of students with specific learning disabilities across the three administrations ranged from 1,202 to 20,757 for ELA and 1,120 to 20,728 in Mathematics. The passing rates for all students tested with accommodations, both standard and non-standard ranged from 11% to 24% for ELA and from 16% to 27% for Mathematics.

Tables 3 and 4 in Appendices 6A to 6C present the passing rates and summary statistics for each accommodation or modification used on the exam. While the accommodation group with the largest percentage of students passing varied across the three administrations, students receiving an accommodation from a Section 504 plan and/or using a Braille or Large Print accommodation most often had the highest percentage passing. Tables 5 and 6 in Appendices 6A to 6C present scale scores at specific percentiles for each accommodation group.

Tables 7 and 8 in Appendices 6A to 6C present summary statistics for the breakdown of each accommodation by reported disability. Tables 9 and 10 in Appendices 6A to 6C present the breakdown of means, standard deviations and passing rates for the comparison of scores for students enrolled in Individualized Education program (IEP) vs. those students with Section 504 plans. The comparison is made within each accommodation category. Across administrations and subject areas, the number of students with IEPs was greater than the number of students with Section 504 plans. The passing rates were typically higher for those students with a Section 504 plan.

Tables 11 and 12 summarize the comparison between language fluency categories within each accommodation/modification group. The most commonly reported categories of accommodation/modification by students with Limited English Proficiency (IFEP, RFEP, ELL)

were additional time, additional breaks, directions read aloud or signed, and the use of a calculator. English-Only speakers outperformed other categories in most but not all instances.

### Generalizability Analyses

Generalizability analyses were performed to estimate variance for the rater and CR item facets. 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). The ELA test features a CR item, which is rated by two raters on a scale between 1 and 4, with non-valid responses scored as zero. A generalizability study (g-study) was performed to estimate variance components for each design facet. 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 Tables 6.1-6.3 below. Generalizability coefficients for the CR item during the three administrations ranged from 0.84 to 0.89. The rater facet was consistently estimated across all three administrations to have no effect on score variance.

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 variance and person variance 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.

**Table 6.1: Generalizability Results February 2004**

#### Person x Rater: CR item Design

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	147438	1.27	0.54274	75.03
Rater (R)	1	0.07	0.00000	0.00
PR, <i>e</i>	147438	0.18	0.18059	24.96
Generalizability Coefficient				0.86
Dependability Coefficient				0.86

**Table 6.2: Generalizability Results March 2004****Person x Rater: CR item Design**

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	292773	1.30	0.54619	72.00
Rater (R)	1	0.75	0.00000	0.00
PR,e	292773	0.21	0.21216	28.00
Generalizability Coefficient				0.84
Dependability Coefficient				0.84

**Table 6.3: Generalizability Results May 2004****Person x Rater: CR item Design**

Facets	Degrees of Freedom	Mean Squares	Variance Components	Percent of Total Variance
Person (P)	12071	1.72	0.77118	81.66
Rater (R)	1	1.64	0.00000	0.00
PR,e	12701	0.17	.17313	18.33
Generalizability Coefficient				0.89
Dependability Coefficient				0.89

**Mark Discrimination Analyses****Background**

The California High School Exit Examination (CAHSEE) has two parts: Mathematics consisting of 80 MC operational items and different sets of 12 embedded field-test items, and ELA consisting of 73 operational items, 72 MC and one CR, and different sets of 7 embedded MC field-test items. Each student record, in addition to the usual item response data, contains right, wrong, and omitted responses for the second darkest mark for each item, if such marks exist. Each student record also contains counts of total number of changed responses, and the total number of changes from wrong-to-right, right-to-wrong, and wrong-to-wrong responses. The CAHSEE contract calls for erasure analysis to be conducted after each administration.

### CAHSEE Answer Document Scanning<sup>4</sup>

The OMR scanners used in processing CAHSEE answer documents are designed for mark detection; they are not designed for erasure detection. The OMR scanners use light and dark as the only indicators that a mark exists. There are 15 levels of mark data that are considered (1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, and F) with '1' being the lowest level and 'F' being the highest level. The scanner determines the mark level for each response gridded. Any mark below level 4 is considered 'no mark'. If there are multiple responses gridded for one item, the scanner sorts the responses by mark level intensity, darkest to lightest. The response with the highest level is formatted in the OMR item record as the first darkest mark. The response with the next highest level is identified as the second darkest mark. If there are two responses with the same mark intensity level this qualifies as a double grid and an asterisk, '\*', is formatted in the OMR item record.

Below is a table of possible cases for item responses within items 1 through 10. The output for each of the cases is given as ITEM OUT for the item area of the OMR record and ERASURE OUT for the erasure part of the OMR record.

ITEM CASES with 15 level value	ITEM OUT	ERASURE OUT
1. 0 0 0 0	0	0
2. E 0 0 0	1	0
3. E A 0 0	1	2
4. E A A 0	1	*
5. E A 8 0	1	2
6. E E 0 0	*	0
7. E E A A	*	*
8. E A A 8	1	*
9. E E 8 0	*	3
10. E D 0 0	*	2

In case 10, two marks are detected in one item of intensity difference of 1 level which qualifies it as a double grid so an '\*' is formatted in the OMR item record. Additionally, since the levels vary by one, there exists a lighter mark in the item so a '2' is formatted in the erasure output area.

#### Purpose

The primary purpose of erasure or mark discrimination analysis is to identify individual students with an unusually high number of changed responses. The second goal is to identify schools with students whose scores changed from Not Pass to Pass as the result of erasing their original responses. From the statistical standpoint, for each test, it is necessary to establish thresholds at the individual level that can be used to identify significantly high patterns of changed responses. The thresholds used in these analyses were empirically determined after examining distributions of erasure variables from the each of the CAHSEE administrations during the February, March & May 2004 administrations. In addition, these thresholds were established with a practical purpose of

---

<sup>4</sup> The information in this section is proprietary to National Computer Systems, Inc.

identifying schools and/or students that had a large number of changed answers as opposed to identifying a large number of students who only changed a few answers.

### **Data Source**

This document contains results of analyses conducted on the February, March, and May 2004 student samples that took the ELA and Mathematics tests under standard testing conditions. Students who took Braille, audio CD, and Large Print test forms were not included in these analyses. The breakdown of the ELA and Mathematics sample sizes was as follows: February 2004, total 152,378 students, ELA 147,178, mathematics 148,185; March 2004, total 300,672 students, ELA 292,090, mathematics 293,116 and May 2004 total 16,510 students, ELA 12,049, mathematics 11,509.

### **Data Analysis**

The tables in Appendix 6.D present frequency distributions, means, and standard deviations of the total number of erasures for ELA and Mathematics tests for the administrations listed above. These tables show that, for most students, the total number of erasures was small and the average number of total erasures was slightly higher for mathematics than it was for ELA. We used a very conservative threshold of the top one percent of students with the largest number of total erasures to be included in the final sample that warranted closer scrutiny. The one percent of the total sample for ELA and mathematics ranged from 6 to 8 erasures.

In Appendix 6.D, Tables 6.D.1- 6.D.6 present frequency distributions, means, and standard deviations of the total number of erasures for ELA and mathematics for each of the administrations. Table 6.D.7 presents the number of schools by test administration where there were students who had a number of erasures greater than the mathematics and ELA thresholds and their scores changed from Not pass to Pass as the result of erasing their original responses. For ELA the percentage of schools who met these criteria varied from 0 to 2.4% for ELA and from 0.4% to 6.6% for mathematics.

The analysis performed on the CAHSEE administrations showed that very few students per school made a significant number of changes in their responses and even the individuals with the largest number of changes might have made them because they misgridded their answer document (e.g., they were off one position) and noticed their mistake before they turned it in. Students taking CAHSEE have time and opportunity to examine their responses carefully because the test is untimed. In addition, these tests are given in a classroom and seating arrangements are not available to examine; therefore, it would be difficult to establish if any copying took place. Furthermore, a relatively small number of scores per school changed from Not Pass to Pass as the result of students erasing their original responses.

## Appendix 6.A: Results of Accommodations and Disability Analyses – February 2004

**Table 6.A.1: Summary Statistics by Test Condition and Disability ELA February 2004**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	4128	328	28	22
Non-Standard	976	323	28	17
All	5104	327	28	21

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	283	300	20	2
Hard of Hearing	123	337	32	35
Deaf	64	320	34	19
Speech or Language Impairment	710	350	38	48
Visual Impairment	51	358	42	65
Emotional Disturbance	870	337	37	36
Orthopedic Impairment	86	342	41	43
Other Health Impairment	489	350	35	50
Specific Learning Disability	10058	330	28	24
Deaf-Blindness	*	*	*	*
Multiple Disabilities	24	318	33	13
Autism	82	358	44	61
Traumatic Brain Injury	32	329	32	28
Unknown	655	357	38	58

\* Results for groups with less than 11 members are not reported

**Table 6.A.2: Summary Statistics by Test Condition and Disability Mathematics February 2004**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	3171	337	27	27
Non-Standard	2419	333	25	22
All	5590	335	26	25

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	271	312	17	3
Hard of Hearing	129	344	29	39
Deaf	64	336	31	30
Speech or Language Impairment	706	356	39	48
Visual Impairment	53	357	40	53
Emotional Disturbance	835	333	32	24
Orthopedic Impairment	88	336	34	33
Other Health Impairment	486	348	33	44
Specific Learning Disability	10117	334	26	23
Deaf-Blindness	*	*	*	*
Multiple Disabilities	23	318	24	9
Autism	82	360	44	57
Traumatic Brain Injury	31	332	22	26
Unknown	631	355	34	50

\* Results for groups with less than 11 members are not reported

**Table 6.A.3: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type ELA February 2004**

Accommodation	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Appl Avg. Score
							WA	RC	LR/A	WS	WC	CR item
Accommodation in Section 504 Plan	108	65	60%	43	40%	358	71%	66%	67%	56%	61%	2.1
Accommodation in an IEP	3677	792	22%	2885	78%	328	56%	47%	52%	37%	44%	1.6
Arithmetic Table	*	*	*	*	*	*	*	*	*	*	*	*
Assistive Device	11	2	18%	9	82%	322	56%	47%	60%	27%	31%	1.4
Directions Read Aloud	528	76	14%	452	86%	322	54%	44%	49%	35%	40%	1.5
Calculator	*	*	*	*	*	*	*	*	*	*	*	*
Dictionary	389	50	13%	339	87%	319	51%	41%	47%	32%	39%	1.5
Scribed Marked Answer Document	31	7	23%	24	77%	319	50%	41%	51%	34%	34%	1.3
Sign Language	11	0	0%	11	100%	307	43%	37%	35%	27%	39%	0.9
Spell Checker or Grammar Checker	115	35	30%	80	70%	335	63%	51%	54%	40%	48%	1.8
Unlisted Modification	281	47	17%	234	83%	321	50%	43%	47%	35%	41%	1.4
Audio CD	204	34	17%	170	83%	324	55%	45%	49%	35%	41%	1.5
Large Print	27	16	59%	11	41%	347	71%	56%	64%	53%	57%	2.0
Braille	11	9	82%	2	18%	379	80%	70%	74%	57%	75%	2.7
English Language Accommodation.	1114	179	16%	935	84%	322	51%	44%	48%	35%	44%	1.5

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> WA — Word Analysis, RC — Reading Comprehension, LR/A — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

**Table 6.A.4: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type  
Mathematics February 2004**

	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>					
							Average Percent Correct					
							PS	NS	AF	MG	Alg1	
<b>Accommodation</b>												
Accommodation in Section 504 Plan	110	65	59%	45	41%	359	63%	60%	59%	59%	44%	
Accommodation in an IEP	4056	840	21%	3216	79%	332	44%	45%	41%	43%	33%	
Arithmetic Table	121	13	11%	108	89%	322	34%	40%	34%	38%	28%	
Assistive Device	*	*	*	*	*	*	*	*	*	*	*	
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*	*	
Calculator	2152	482	22%	1670	78%	333	44%	48%	41%	44%	32%	
Dictionary	111	16	14%	95	86%	326	36%	43%	38%	39%	31%	
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*	*	
Sign Language	*	*	*	*	*	*	*	*	*	*	*	
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*	*	
Unlisted Modification	257	45	18%	212	82%	328	40%	41%	39%	42%	31%	
Audio CD	184	18	10%	166	90%	326	40%	40%	36%	41%	32%	
Large Print	26	12	46%	14	54%	349	52%	49%	53%	57%	47%	
Braille	*	*	*	*	*	*	*	*	*	*	*	
English Language Accommodation.	1046	396	38%	650	62%	345	47%	51%	50%	52%	45%	

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 6.A.5: Percentiles of Scale Scores for Students with Accommodations\* ELA February 2004**

Accommodation	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
Accommodation in Section 504 Plan	275	295	322	364	383	424	434	358	39	108
Accommodation in an IEP	275	287	307	325	344	380	403	328	28	3677
Arithmetic Table	*	*	*	*	*	*	*	*	*	*
Assistive Device	287	287	309	321	326	366	366	322	24	11
Directions Read Aloud	275	287	305	319	337	366	388	322	25	528
Calculator	*	*	*	*	*	*	*	*	*	*
Dictionary	275	282	303	314	330	371	394	319	25	389
Scribed Marked Answer Document	275	276	299	305	337	360	434	319	33	31
Sign Language	275	275	297	301	317	346	346	307	19	11
Spell Checker or Grammar Checker	275	293	310	328	358	397	429	335	32	115
Unlisted Modification	275	279	297	314	341	380	407	321	30	281
Audio CD	275	287	306	321	342	366	383	324	25	204
Large Print	279	301	321	354	368	391	394	347	30	27
Braille	275	275	378	388	419	429	429	379	50	11
English Language Accommodation.	275	284	305	319	339	371	388	322	26	1114

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.A.6: Percentiles of Scale Scores for Students with Accommodations\* Mathematics  
February 2004**

Accommodation	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
Accommodation in Section 504 Plan	302	312	330	359	382	406	450	359	34	110
Accommodation in an IEP	286	300	316	328	345	377	414	332	25	4056
Arithmetic Table	286	297	307	318	330	367	389	322	22	121
Assistive Device	*	*	*	*	*	*	*	*	*	*
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*
Calculator	282	300	316	330	346	377	406	333	24	2152
Dictionary	297	300	309	320	337	365	384	326	22	111
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	282	295	309	322	339	386	418	328	27	257
Audio CD	295	300	314	324	335	356	377	326	17	184
Large Print	286	292	322	344	375	422	447	349	39	26
Braille	*	*	*	*	*	*	*	*	*	*
English Language Accommodation.	297	309	326	341	359	394	433	345	27	1046

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.A.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* ELA February 2004**

<b>Accommodation Type</b>	<b>Disability</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	Specific Learning Disability	15	330	39	27
Accommodation in an IEP	Mental Retardation	87	301	16	0
	Hard of Hearing	30	335	34	30
	Deaf	38	312	25	11
	Speech or Language Impairment	122	326	29	20
	Visual Impairment	30	359	43	70
	Emotional Disturbance	281	334	37	31
	Orthopedic Impairment	26	323	36	23
	Other Health Impairment	124	344	31	43
	Specific Learning Disability	2799	327	26	19
	Autism	26	345	41	46
	Traumatic Brain Injury	12	325	25	17
Audio CD	Speech or Language Impairment	22	323	25	14
	Emotional Disturbance	11	323	33	18
	Specific Learning Disability	153	325	23	16
Dictionary	Speech or Language Impairment	12	313	28	8
	Emotional Disturbance	12	327	33	17
	Specific Learning Disability	207	316	23	9
Directions Read Aloud Signed	Mental Retardation	25	301	13	0
	Speech or Language Impairment	38	319	23	11
	Emotional Disturbance	31	327	34	29
	Other Health Impairment	14	320	26	21
	Specific Learning Disability	343	324	24	15
English Language Accommodation	Specific Learning Disability	32	328	32	28
Large Print	Visual Impairment	16	358	29	81
Scribed Marked Answer Document	Emotional Disturbance	12	298	15	0
Spell Checker or Grammar Checker	Specific Learning Disability	81	327	25	22
Unlisted Modification	Mental Retardation	11	292	12	0
	Emotional Disturbance	11	347	39	27
	Specific Learning Disability	164	317	26	12

\* Results for groups with less than 11 members are not reported

**Table 6.A.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* Mathematics February 2004**

<b>Accommodation Type</b>	<b>Disability</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	Specific Learning Disability	19	344	37	37
Accommodation in an IEP	Mental Retardation	92	309	15	1
	Hard of Hearing	38	340	32	32
	Deaf	38	332	28	26
	Speech or Language Impairment	142	336	29	23
	Visual Impairment	28	361	39	64
	Emotional Disturbance	272	332	30	24
	Orthopedic Impairment	27	321	19	11
	Other Health Impairment	141	340	28	31
	Specific Learning Disability	3124	332	24	19
	Multiple Disabilities	12	312	15	0
	Autism	27	360	44	52
	Traumatic Brain Injury	13	322	21	15
Arithmetic Table	Specific Learning Disability	102	322	21	9
Audio CD	Speech or Language Impairment	16	334	19	19
	Specific Learning Disability	143	326	17	9
Calculator	Mental Retardation	69	309	16	3
	Hard of Hearing	16	342	37	31
	Speech or Language Impairment	110	337	24	23
	Emotional Disturbance	98	332	28	26
	Orthopedic Impairment	14	329	25	29
	Other Health Impairment	87	341	29	33
	Specific Learning Disability	1674	333	23	21
Dictionary	Specific Learning Disability	54	317	14	2
English Language Accommodation	Specific Learning Disability	30	336	31	33
Large Print	Visual Impairment	15	371	36	80
Unlisted Modification	Mental Retardation	11	298	13	0
	Other Health Impairment	11	335	33	36
	Specific Learning Disability	141	325	24	13

\* Results for groups with less than 11 members are not reported

**Table 6.A.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* ELA February 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	*	*	*	*
Assistive Device	Sec504	*	*	*	*
	Special Education	*	*	*	*
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	403	324	26	17
Calculator	Sec504	*	*	*	*
	Special Education	*	*	*	*
Dictionary	Sec504	17	327	30	24
	Special Education	216	315	23	8
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	26	315	26	19
Sign Language	Sec504	*	*	*	*
	Special Education	*	*	*	*
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	90	334	33	31
Unlisted Modification	Sec504	14	335	36	21
	Special Education	195	319	29	15
Audio CD	Sec504	*	*	*	*
	Special Education	177	324	25	18
Large Print	Sec504	*	*	*	*
	Special Education	26	347	30	58
Braille	Sec504	*	*	*	*
	Special Education	*	*	*	*
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	*	*	*	*

\* Results for groups with less than 11 members are not reported

**Table 6.A.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* Mathematics February 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	114	320	21	8
Assistive Device	Sec504	*	*	*	*
	Special Education	*	*	*	*
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	*	*	*	*
Calculator	Sec504	28	356	35	54
	Special Education	1743	333	24	22
Dictionary	Sec504	*	*	*	*
	Special Education	54	321	20	11
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	*	*	*	*
Sign Language	Sec504	*	*	*	*
	Special Education	*	*	*	*
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	*	*	*	*
Unlisted Modification	Sec504	14	333	21	21
	Special Education	176	326	26	15
Audio CD	Sec504	*	*	*	*
	Special Education	170	326	17	9
Large Print	Sec504	*	*	*	*
	Special Education	24	353	39	50
Braille	Sec504	*	*	*	*
	Special Education	*	*	*	*
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	27	343	29	44

\* Results for groups with less than 11 members are not reported

**Table 6.A.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* ELA February 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	English Only	82	366	38	72
	English Learner	20	327	21	10
Accommodation in an IEP	English Only	2578	331	29	25
	Initially Fluent English Proficient	136	334	29	30
	Redesignated Fluent English Proficient	80	334	27	26
	English Learner	871	317	21	8
	Unknown	12	348	27	25
Audio CD	English Only	139	325	24	17
	English Learner	58	320	25	16
Dictionary	English Only	157	319	27	15
	Initially Fluent English Proficient	15	341	33	47
	English Learner	214	317	23	9
Directions Read Aloud Signed	English Only	330	324	27	18
	Initially Fluent English Proficient	13	331	25	15
	English Learner	171	317	20	9
English Language Accommodation	English Only	52	325	31	25
	Redesignated Fluent English Proficient	11	371	26	82
	English Learner	1046	322	25	15
Large Print	English Only	19	351	31	63
Scribed Marked Answer Document	English Only	26	319	35	27
Spell Checker or Grammar Checker	English Only	77	338	33	34
	English Learner	30	321	21	13
Unlisted Modification	English Only	166	325	32	22
	English Learner	91	312	24	7

\* Results for groups with less than 11 members are not reported

**Table 6.A.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* Mathematics February 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	English Only	90	363	33	67
	English Learner	15	331	10	7
Accommodation in an IEP	English Only	2894	333	27	23
	Initially Fluent English Proficient	156	338	27	29
	Redesignated Fluent English Proficient	80	337	22	33
	English Learner	912	326	19	11
	Unknown	14	332	16	21
Arithmetic Table	English Only	99	323	23	12
	English Learner	19	313	18	5
Audio CD	English Only	123	327	18	11
	English Learner	56	324	15	5
Calculator	English Only	1498	335	26	25
	Initially Fluent English Proficient	91	340	24	36
	Redesignated Fluent English Proficient	43	339	26	37
	English Learner	515	327	17	10
Dictionary	English Only	59	321	20	10
	English Learner	51	333	24	20
English Language Accommodation	English Only	47	354	39	51
	English Learner	986	344	26	37
Large Print	English Only	18	350	43	50
Unlisted Modification	English Only	156	328	28	18
	English Learner	82	327	26	16

\* Results for groups with less than 11 members are not reported

**Appendix 6.B: Results of Accommodations and Disability Analyses – March 2004**

**Table 6.B.1: Summary Statistics by Test Condition and Disability ELA**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	8224	330	32	25
Non-Standard	1392	326	29	19
All	9616	329	32	24

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	456	302	20	1
Hard of Hearing	262	342	39	37
Deaf	161	321	40	17
Speech or Language Impairment	1535	349	40	46
Visual Impairment	130	350	48	48
Emotional Disturbance	1709	343	42	42
Orthopedic Impairment	217	349	45	49
Other Health Impairment	1365	355	38	54
Specific Learning Disability	20757	332	31	28
Deaf-Blindness	*	*	*	*
Multiple Disabilities	111	323	28	18
Autism	213	349	45	50
Traumatic Brain Injury	62	329	36	27
Unknown	2933	358	38	58

\* Results for groups with less than 11 members are not reported

**Table 6.B.2: Summary Statistics by Test Condition and Disability Mathematics**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	6190	340	29	31
Non-Standard	4107	335	24	22
All	10297	338	28	27

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	444	313	18	3
Hard of Hearing	263	352	36	45
Deaf	163	336	34	29
Speech or Language Impairment	1539	356	38	49
Visual Impairment	129	353	44	47
Emotional Disturbance	1685	341	34	34
Orthopedic Impairment	210	349	38	43
Other Health Impairment	1368	353	35	47
Specific Learning Disability	20728	338	27	28
Deaf-Blindness	*	*	*	*
Multiple Disabilities	109	331	23	22
Autism	210	354	43	47
Traumatic Brain Injury	61	341	33	31
Unknown	2919	358	33	55

\* Results for groups with less than 11 members are not reported

**Table 6.B.3: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type ELA  
March 2004**

Accommodation	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Appl Avg. Score
							WA	RC	LR/A	WS	WC	CR item
Accommodation in Section 504 Plan	174	142	82%	32	18%	381	84%	77%	78%	64%	71%	2.4
Accommodation in an IEP	7489	1929	26%	5560	74%	330	57%	52%	51%	37%	44%	1.5
Arithmetic Table	*	*	*	*	*	*	*	*	*	*	*	*
Assistive Device	22	9	41%	13	59%	334	71%	55%	60%	43%	46%	1.1
Directions Read Aloud	1172	174	15%	998	85%	322	53%	47%	48%	35%	39%	1.3
Calculator	*	*	*	*	*	*	*	*	*	*	*	*
Dictionary	375	64	17%	311	83%	323	55%	48%	47%	34%	41%	1.3
Scribed Marked Answer Document	68	25	37%	43	63%	340	64%	55%	57%	42%	44%	1.9
Sign Language	26	4	15%	22	85%	322	45%	43%	46%	38%	42%	1.4
Spell Checker or Grammar Checker	196	52	27%	144	73%	337	63%	56%	57%	39%	46%	1.7
Unlisted Modification	349	73	21%	276	79%	323	51%	48%	46%	34%	41%	1.3
Audio CD	497	68	14%	429	86%	323	55%	45%	50%	37%	42%	1.4
Large Print	71	35	49%	36	51%	351	65%	61%	61%	48%	54%	1.9
Braille	17	8	47%	9	53%	345	57%	59%	56%	43%	56%	1.9
English Language Accommodation.	1783	210	12%	1573	88%	319	49%	45%	43%	32%	42%	1.3

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> WA — Word Analysis, RC — Reading Comprehension, LR/A — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

**Table 6.B.4: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type  
Mathematics March 2004**

Accommodation	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>				
							Average Percent Correct				
							PS	NS	AF	MG	Alg1
Accommodation in Section 504 Plan	171	135	79%	36	21%	381	73%	72%	70%	67%	60%
Accommodation in an IEP	8075	1983	25%	6092	75%	336	46%	48%	45%	40%	34%
Arithmetic Table	117	15	13%	102	87%	325	42%	44%	36%	34%	25%
Assistive Device	15	7	47%	8	53%	349	61%	67%	45%	43%	40%
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*	*
Calculator	3754	856	23%	2898	77%	335	45%	49%	44%	39%	32%
Dictionary	103	24	23%	79	77%	337	44%	46%	48%	41%	36%
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	334	57	17%	277	83%	330	42%	44%	41%	36%	30%
Audio CD	335	23	7%	312	93%	324	38%	41%	35%	39%	29%
Large Print	66	35	53%	31	47%	361	61%	59%	59%	55%	48%
Braille	16	7	44%	9	56%	335	41%	54%	43%	44%	34%
English Language Accommodation.	1402	545	39%	857	61%	346	46%	50%	54%	47%	44%

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 6.B.5: Percentiles of Scale Scores for Students with Accommodations\* ELA March 2004**

	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
<b>Accommodation</b>										
Accommodation in Section 504 Plan	291	319	361	380	407	445	450	381	36	174
Accommodation in an IEP	275	284	306	327	351	389	415	330	32	7489
Arithmetic Table	*	*	*	*	*	*	*	*	*	*
Assistive Device	279	287	304	340	359	389	400	334	34	22
Directions Read Aloud	275	284	304	319	338	368	392	322	25	1172
Calculator	*	*	*	*	*	*	*	*	*	*
Dictionary	275	281	302	319	341	374	389	323	28	375
Scribed Marked Answer Document	279	286	310	335	363	407	450	340	38	68
Sign Language	293	293	308	318	329	367	379	322	24	26
Spell Checker or Grammar Checker	279	298	316	335	351	389	428	337	29	196
Unlisted Modification	275	275	300	316	343	381	407	323	32	349
Audio CD	275	287	309	321	337	366	380	323	23	497
Large Print	275	293	312	349	381	438	450	351	46	71
Braille	275	275	303	348	385	424	424	345	48	17
English Language Accommodation.	275	284	302	316	334	367	394	319	26	1783

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.B.6: Percentiles of Scale Scores for Students with Accommodations\* Mathematics  
March 2004**

Accommodation	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
Accommodation in Section 504 Plan	313	326	353	383	407	443	450	381	35	171
Accommodation in an IEP	289	301	318	330	349	385	417	336	27	8075
Arithmetic Table	275	292	311	324	342	360	376	325	22	117
Assistive Device	313	313	324	349	357	398	398	349	26	15
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*
Calculator	289	304	318	332	347	381	410	335	24	3754
Dictionary	289	309	322	338	349	370	392	337	20	103
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	286	298	313	326	342	374	407	330	24	334
Audio CD	289	300	314	322	334	357	381	324	18	335
Large Print	289	302	324	352	398	450	450	361	45	66
Braille	292	292	312	337	359	382	382	335	27	16
English Language Accommodation.	295	309	326	342	360	395	426	346	27	1402

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.B.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* ELA March 2004**

<b>Accommodation Type</b>	<b>Disability</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	Other Health Impairment	13	395	40	92
	Specific Learning Disability	22	381	46	73
Accommodation in an IEP	Mental Retardation	154	301	18	1
	Hard of Hearing	78	330	33	21
	Deaf	113	317	37	14
	Speech or Language Impairment	253	324	28	19
	Visual Impairment	48	339	44	40
	Emotional Disturbance	388	338	41	38
	Orthopedic Impairment	68	347	47	44
	Other Health Impairment	288	347	38	45
	Specific Learning Disability	5589	329	29	23
	Multiple Disabilities	30	317	27	13
	Autism	66	334	40	38
	Traumatic Brain Injury	20	324	32	20
	Unknown	15	324	26	13
Assistive Device	Specific Learning Disability	15	331	30	40
Audio CD	Mental Retardation	27	302	21	4
	Speech or Language Impairment	12	313	17	0
	Specific Learning Disability	413	325	22	14
Braille	Visual Impairment	13	338	47	38
Dictionary	Specific Learning Disability	184	326	26	18
Directions Read Aloud Signed	Mental Retardation	60	304	19	2
	Speech or Language Impairment	38	320	24	8
	Emotional Disturbance	27	329	29	19
	Orthopedic Impairment	15	330	35	27
	Other Health Impairment	25	324	29	24
	Specific Learning Disability	929	323	25	15
	Multiple Disabilities	14	314	23	7
English Language Accommodation	Specific Learning Disability	89	322	22	10
	Unknown	36	329	37	25
Large Print	Visual Impairment	47	355	50	53
Scribed Marked Answer Document	Orthopedic Impairment	14	355	52	43
	Specific Learning Disability	30	339	33	43
Sign Language	Hard of Hearing	15	325	20	13
Spell Checker or Grammar Checker	Specific Learning Disability	149	335	29	24
Unlisted Modification	Emotional Disturbance	25	324	49	36
	Other Health Impairment	14	340	47	36
	Specific Learning Disability	258	323	29	19

\* Results for groups with less than 11 members are not reported

**Table 6.B.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* Mathematics March 2004**

Accommodation Type	Disability	N	Mean	SD	% Passed
Accommodation in Section 504 Plan	Other Health Impairment	12	391	40	75
	Specific Learning Disability	25	381	45	72
Accommodation in an IEP	Mental Retardation	153	314	14	3
	Hard of Hearing	83	345	33	34
	Deaf	123	332	31	24
	Speech or Language Impairment	265	335	26	25
	Visual Impairment	42	351	42	52
	Emotional Disturbance	444	337	30	29
	Orthopedic Impairment	60	343	38	38
	Other Health Impairment	339	345	34	37
	Specific Learning Disability	6130	335	25	23
	Multiple Disabilities	30	330	25	13
	Autism	68	338	35	28
	Traumatic Brain Injury	22	333	26	27
	Unknown	14	321	17	14
Arithmetic Table	Specific Learning Disability	74	326	22	15
Audio CD	Mental Retardation	21	314	9	0
	Specific Learning Disability	275	325	17	7
Braille	Visual Impairment	12	335	28	42
Calculator	Mental Retardation	86	317	15	3
	Hard of Hearing	32	357	39	53
	Speech or Language Impairment	112	333	22	17
	Visual Impairment	11	355	40	55
	Emotional Disturbance	170	339	29	29
	Orthopedic Impairment	31	339	36	39
	Other Health Impairment	155	340	27	32
	Specific Learning Disability	2974	335	23	22
	Multiple Disabilities	22	328	18	14
	Autism	36	334	31	25
Dictionary	Specific Learning Disability	28	333	21	25
English Language Accommodation	Specific Learning Disability	76	331	19	16
	Unknown	31	356	34	55
Large Print	Visual Impairment	39	364	46	62
Unlisted Modification	Emotional Disturbance	25	329	22	24
	Other Health Impairment	11	337	44	18
	Specific Learning Disability	254	329	22	17

\* Results for groups with less than 11 members are not reported

**Table 6.B.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* ELA March 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>%Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	*	*	*	*
Assistive Device	Sec504	*	*	*	*
	Special Education	18	331	32	39
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	1066	322	25	15
Calculator	Sec504	*	*	*	*
	Special Education	*	*	*	*
Dictionary	Sec504	*	*	*	*
	Special Education	222	325	27	16
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	58	344	38	40
Sign Language	Sec504	*	*	*	*
	Special Education	24	321	23	13
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	174	339	28	27
Unlisted Modification	Sec504	*	*	*	*
	Special Education	274	324	32	22
Audio CD	Sec504	*	*	*	*
	Special Education	466	323	23	13
Large Print	Sec504	*	*	*	*
	Special Education	42	344	43	43
Braille	Sec504	*	*	*	*
	Special Education	*	*	*	*
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	14	344	46	43

\* Results for groups with less than 11 members are not reported

**Table 6.B.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* Mathematics March 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>%Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	61	326	23	15
Assistive Device	Sec504	*	*	*	*
	Special Education	14	346	23	43
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	*	*	*	*
Calculator	Sec504	16	369	37	63
	Special Education	3169	335	24	23
Dictionary	Sec504	*	*	*	*
	Special Education	30	339	23	37
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	*	*	*	*
Sign Language	Sec504	*	*	*	*
	Special Education	*	*	*	*
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	*	*	*	*
Unlisted Modification	Sec504	*	*	*	*
	Special Education	259	331	24	19
Audio CD	Sec504	*	*	*	*
	Special Education	304	324	17	6
Large Print	Sec504	*	*	*	*
	Special Education	38	362	39	55
Braille	Sec504	*	*	*	*
	Special Education	12	338	30	50
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	78	331	20	15

\* Results for groups with less than 11 members are not reported

**Table 6.B.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* ELA March 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in Section 504 Plan	English Only	163	383	35	84
Accommodation in an IEP	English Only	4983	334	33	31
	Initially Fluent English Proficient	337	332	30	27
	Redesignated Fluent English Proficient	296	343	32	39
	English Learner	1853	317	24	10
	Unknown	20	314	29	20
Assistive Device	English Only	11	337	38	36
Audio CD	English Only	332	326	25	17
	Initially Fluent English Proficient	12	336	15	17
	Redesignated Fluent English Proficient	11	327	19	9
	English Learner	142	316	19	6
Braille	English Only	13	353	43	54
Dictionary	English Only	147	324	31	22
	Initially Fluent English Proficient	13	335	20	15
	Redesignated Fluent English Proficient	11	341	28	36
	English Learner	204	321	25	13
Directions Read Aloud Signed	English Only	746	325	26	18
	Initially Fluent English Proficient	34	324	24	18
	Redesignated Fluent English Proficient	28	336	25	25
	English Learner	358	315	21	8
English Language Accommodation	English Only	102	323	32	16
	Initially Fluent English Proficient	16	341	32	44
	Redesignated Fluent English Proficient	14	330	28	14
	English Learner	1641	319	25	11
Large Print	English Only	48	358	45	56
	English Learner	15	314	31	13
Scribed Marked Answer Document	English Only	50	345	40	40
	English Learner	11	319	32	18
Sign Language	English Only	13	322	26	15
Spell Checker or Grammar Checker	English Only	144	342	30	33
	English Learner	42	320	20	7
Unlisted Modification	English Only	233	325	34	23
	Initially Fluent English Proficient	14	324	33	21
	English Learner	92	315	29	14

\* Results for groups with less than 11 members are not reported

**Table 6.B.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* Mathematics March 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>%Passed</b>
Accommodation in Section 504 Plan	English Only	160	382	35	79
Accommodation in an IEP	English Only	5435	338	28	28
	Initially Fluent English Proficient	379	338	26	24
	Redesignated Fluent English Proficient	300	346	31	39
	English Learner	1938	327	21	13
	Unknown	23	334	19	13
Arithmetic Table	English Only	60	324	22	12
	English Learner	48	325	20	13
Audio CD	English Only	198	326	20	10
	Redesignated Fluent English Proficient	13	326	15	8
	English Learner	116	321	14	2
Braille	English Only	11	334	27	45
Calculator	English Only	2577	337	25	26
	Initially Fluent English Proficient	178	337	23	24
	Redesignated Fluent English Proficient	152	343	26	34
	English Learner	834	327	19	12
	Unknown	13	326	14	0
Dictionary	English Only	28	334	24	25
	English Learner	69	339	19	23
English Language Accommodation	English Only	97	336	26	28
	Initially Fluent English Proficient	15	351	39	33
	Redesignated Fluent English Proficient	15	335	22	20
	English Learner	1265	347	27	40
Large Print	English Only	44	367	45	61
	English Learner	13	323	28	8
Unlisted Modification	English Only	221	328	23	15
	Initially Fluent English Proficient	13	345	32	38
	English Learner	88	330	26	19

\* Results for groups with less than 11 members are not reported

## Appendix 6.C: Results of Accommodations and Disability Analyses - May 2004

**Table 6.C.1: Summary Statistics by Test Condition and Disability ELA**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	324	311	29	11
Non-Standard	56	315	32	14
All	380	311	30	11

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	47	293	15	0
Hard of Hearing	12	319	27	17
Deaf	*	*	*	*
Speech or Language Impairment	43	327	35	21
Visual Impairment	*	*	*	*
Emotional Disturbance	169	333	39	35
Orthopedic Impairment	19	331	42	37
Other Health Impairment	68	328	41	31
Specific Learning Disability	1202	317	28	13
Multiple Disabilities	*	*	*	*
Autism	12	331	46	33
Traumatic Brain Injury	*	*	*	*
Unknown	473	355	44	53

\* Results for groups with less than 11 members are not reported

**Table 6.C.2: Summary Statistics by Test Condition and Disability Mathematics**

<b>Accommodation</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Standard	234	326	27	15
Non-Standard	146	324	25	18
All	380	325	26	16

<b>Disability*</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Percent Passed</b>
Mental Retardation	49	308	15	2
Hard of Hearing	*	*	*	*
Deaf	*	*	*	*
Speech or Language Impairment	42	332	29	19
Visual Impairment	*	*	*	*
Emotional Disturbance	155	329	30	23
Orthopedic Impairment	21	338	35	38
Other Health Impairment	56	333	33	27
Specific Learning Disability	1120	324	23	13
Multiple Disabilities	*	*	*	*
Autism	12	337	46	17
Traumatic Brain Injury	*	*	*	*
Unknown	423	353	36	50

\* Results for groups with less than 11 members are not reported

**Table 6.C.3: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type ELA May 2004**

Accommodation	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Reading <sup>2</sup> Avg. % Correct			Writing <sup>2</sup> Avg. % Correct		Writing Appl Avg. Score
							WA	RC	LR/A	WS	WC	CR item
Accommodation in Section 504 Plan	*	*	*	*	*	*	*	*	*	*	*	*
Accommodation in an IEP	295	39	13%	256	87%	314	47%	40%	43%	31%	36%	1.5
Arithmetic Table	*	*	*	*	*	*	*	*	*	*	*	*
Assistive Device	*	*	*	*	*	*	*	*	*	*	*	*
Directions Read Aloud	65	4	6%	61	94%	307	42%	36%	38%	27%	30%	1.4
Calculator	*	*	*	*	*	*	*	*	*	*	*	*
Dictionary	*	*	*	*	*	*	*	*	*	*	*	*
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	33	5	15%	28	85%	314	46%	40%	44%	35%	38%	1.3
Audio CD	12	1	8%	11	92%	315	48%	36%	43%	35%	33%	1.6
Large Print	*	*	*	*	*	*	*	*	*	*	*	*
Braille	*	*	*	*	*	*	*	*	*	*	*	*
English Language Accommodation.	74	2	3%	72	97%	301	38%	35%	37%	27%	35%	0.8

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> WA — Word Analysis, RC — Reading Comprehension, LR/A — Literary Responses/Analysis, WS — Writing Strategies, WC — Writing Conventions

**Table 6.C.4: Demographic Summary For All Examinees<sup>1</sup> by Accommodation Type  
Mathematics May 2004**

Accommodation	N Tested	N Pass	% Pass	N Not Pass	% Not Pass	Mean Scale Score	Strands for Mathematics <sup>2</sup>					
							Average Percent Correct					
							PS	NS	AF	MG	Alg1	
Accommodation in Section 504 Plan	*	*	*	*	*	*	*	*	*	*	*	*
Accommodation in an IEP	289	28	10%	261	90%	320	37%	41%	34%	32%	24%	24%
Arithmetic Table	*	*	*	*	*	*	*	*	*	*	*	*
Assistive Device	*	*	*	*	*	*	*	*	*	*	*	*
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*	*	*
Calculator	126	19	15%	107	85%	321	39%	43%	33%	33%	23%	23%
Dictionary	*	*	*	*	*	*	*	*	*	*	*	*
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	24	5	21%	19	79%	328	44%	45%	40%	35%	30%	30%
Audio CD	*	*	*	*	*	*	*	*	*	*	*	*
Large Print	*	*	*	*	*	*	*	*	*	*	*	*
Braille	*	*	*	*	*	*	*	*	*	*	*	*
English Language Accommodation.	66	23	35%	43	65%	342	42%	56%	47%	50%	41%	41%

<sup>1</sup> Results for groups with less than 11 members are not reported

<sup>2</sup> PS — Probability/ Statistics, NS — Number Sense, AF — Algebra & Functions, MG — Measurement/Geometry, Alg1 — Algebra 1

**Table 6.C.5: Percentiles of Scale Scores for Students with Accommodations\* ELA May 2004**

Accommodation	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
Accommodation in Section 504 Plan	*	*	*	*	*	*	*	*	*	*
Accommodation in an IEP	275	275	293	306	330	375	408	314	30	295
Arithmetic Table	*	*	*	*	*	*	*	*	*	*
Assistive Device	*	*	*	*	*	*	*	*	*	*
Directions Read Aloud	275	275	289	306	323	356	385	307	26	65
Calculator	*	*	*	*	*	*	*	*	*	*
Dictionary	*	*	*	*	*	*	*	*	*	*
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	275	275	281	314	337	382	408	314	35	33
Audio CD	275	275	305	314	324	385	385	315	27	12
Large Print	*	*	*	*	*	*	*	*	*	*
Braille	*	*	*	*	*	*	*	*	*	*
English Language Accommodation.	275	275	285	295	314	343	360	301	20	74

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.C.6: Percentiles of Scale Scores for Students with Accommodations\* Mathematics  
May 2004**

Accommodation	Percentiles							Mean Scale Score	SD**	N Tested
	1%	5%	25%	50%	75%	95%	99%			
Accommodation in Section 504 Plan	*	*	*	*	*	*	*	*	*	*
Accommodation in an IEP	277	293	306	317	329	361	387	320	21	289
Arithmetic Table	*	*	*	*	*	*	*	*	*	*
Assistive Device	*	*	*	*	*	*	*	*	*	*
Directions Read Aloud	*	*	*	*	*	*	*	*	*	*
Calculator	284	293	306	317	330	360	375	321	21	126
Dictionary	*	*	*	*	*	*	*	*	*	*
Scribed Marked Answer Document	*	*	*	*	*	*	*	*	*	*
Sign Language	*	*	*	*	*	*	*	*	*	*
Spell Checker or Grammar Checker	*	*	*	*	*	*	*	*	*	*
Unlisted Modification	284	293	305	319	346	387	392	328	29	24
Audio CD	*	*	*	*	*	*	*	*	*	*
Large Print	*	*	*	*	*	*	*	*	*	*
Braille	*	*	*	*	*	*	*	*	*	*
English Language Accommodation.	275	296	321	337	361	392	448	342	33	66

\* Results for groups with less than 11 members are not reported

\*\* SD - Standard Deviation

**Table 6.C.7: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* ELA May 2004**

<b>Accommodation Type</b>	<b>Disability</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in an IEP	Mental Retardation	20	291	14	0
	Emotional Disturbance	42	318	33	17
	Other Health Impairment	13	328	46	38
	Specific Learning Disability	189	312	26	10
Directions Read Aloud Signed	Specific Learning Disability	43	312	25	7
Unlisted Modification	Specific Learning Disability	15	312	32	7

\* Results for groups with less than 11 members are not reported

**Table 6.C.8: Summary of Scale Scores and Passing Rates by Disability and Accommodation Type\* Mathematics May 2004**

<b>Accommodation Type</b>	<b>Disability</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in an IEP	Mental Retardation	17	312	15	6
	Emotional Disturbance	40	320	23	15
	Specific Learning Disability	198	321	20	9
Calculator	Emotional Disturbance	15	323	23	20
	Specific Learning Disability	91	321	21	14

\* Results for groups with less than 11 members are not reported

**Table 6.C.9: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* ELA May 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	*	*	*	*
Assistive Device	Sec504	*	*	*	*
	Special Education	*	*	*	*
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	57	308	27	7
Calculator	Sec504	*	*	*	*
	Special Education	*	*	*	*
Dictionary	Sec504	*	*	*	*
	Special Education	*	*	*	*
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	*	*	*	*
Sign Language	Sec504	*	*	*	*
	Special Education	*	*	*	*
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	*	*	*	*
Unlisted Modification	Sec504	*	*	*	*
	Special Education	31	311	34	13
Audio CD	Sec504	*	*	*	*
	Special Education	*	*	*	*
Large Print	Sec504	*	*	*	*
	Special Education	*	*	*	*
Braille	Sec504	*	*	*	*
	Special Education	*	*	*	*
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	*	*	*	*

\* Results for groups with less than 11 members are not reported

**Table 6.C.10: Summary of Scale Scores and Passing Rates by Special Education and Accommodation Type\* Mathematics May 2004**

<b>Accommodation</b>		<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Arithmetic Table	Sec504	*	*	*	*
	Special Education	*	*	*	*
Assistive Device	Sec504	*	*	*	*
	Special Education	*	*	*	*
Directions Read Aloud	Sec504	*	*	*	*
	Special Education	*	*	*	*
Calculator	Sec504	*	*	*	*
	Special Education	107	320	21	14
Dictionary	Sec504	*	*	*	*
	Special Education	*	*	*	*
Scribed Marked Answer Document	Sec504	*	*	*	*
	Special Education	*	*	*	*
Sign Language	Sec504	*	*	*	*
	Special Education	*	*	*	*
Spell Checker or Grammar Checker	Sec504	*	*	*	*
	Special Education	*	*	*	*
Unlisted Modification	Sec504	*	*	*	*
	Special Education	22	323	25	14
Audio CD	Sec504	*	*	*	*
	Special Education	*	*	*	*
Large Print	Sec504	*	*	*	*
	Special Education	*	*	*	*
Braille	Sec504	*	*	*	*
	Special Education	*	*	*	*
English Language Accommodation.	Sec504	*	*	*	*
	Special Education	*	*	*	*

\* Results for groups with less than 11 members are not reported

**Table 6.C.11: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* ELA May 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in an IEP	English Only	217	316	32	17
	Redesignated Fluent English Proficient	13	312	35	8
	English Learner	48	305	18	0
Directions Read Aloud Signed	English Only	42	311	28	10
	English Learner	16	307	20	0
English Language Accommodation	English Learner	69	298	16	0
Unlisted Modification	English Only	21	309	32	14

\* Results for groups with less than 11 members are not reported

**Table 6.C.12: Summary of Scale Scores and Passing Rates by Language Fluency and Accommodation Type\* Mathematics May 2004**

<b>Accommodation Type</b>	<b>Language Fluency</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>% Passed</b>
Accommodation in an IEP	English Only	217	321	22	12
	English Learner	49	319	14	0
Calculator	English Only	100	321	22	17
	English Learner	18	321	18	6
English Language Accommodation	English Learner	61	346	30	38
Unlisted Modification	English Only	16	325	25	19

\* Results for groups with less than 11 members are not reported

### Appendix 6.D: Mark Discrimination Analysis

**Table 6.D.1: ELA February 2004 Frequency Distribution of Total Number of Erasures**

Total Erasures	Frequency	Percent	Cumulative Frequency	Cumulative Percent
39	1	0.00	1	0.00
34	1	0.00	2	0.00
25	1	0.00	3	0.00
22	2	0.00	5	0.00
21	1	0.00	6	0.00
20	1	0.00	7	0.00
19	1	0.00	8	0.01
18	8	0.01	16	0.01
17	3	0.00	19	0.01
16	4	0.00	23	0.02
15	10	0.01	33	0.02
14	11	0.01	44	0.03
13	17	0.01	61	0.04
12	27	0.02	88	0.06
11	39	0.03	127	0.09
10	67	0.05	194	0.13
9	84	0.06	278	0.19
8	156	0.11	434	0.29
7	264	0.18	698	0.47
6	478	0.32	1176	0.80
5	877	0.60	2053	1.39
4	1743	1.18	3796	2.58
3	3799	2.58	7595	5.16
2	9049	6.15	16644	11.31
1	24335	16.53	40979	27.84
0	106199	72.16	147178	100.00

Mean    0.51    Standard Deviation    1.13

**Table 6.D.2: Mathematics February 2004 Frequency Distribution of Total Number of Erasures**

<b>Total Erasures</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
51	1	0.00	1	0.00
39	1	0.00	2	0.00
37	1	0.00	3	0.00
33	1	0.00	4	0.00
28	1	0.00	5	0.00
26	1	0.00	6	0.00
25	2	0.00	8	0.01
24	3	0.00	11	0.01
23	5	0.00	16	0.01
22	2	0.00	18	0.01
21	3	0.00	21	0.01
20	4	0.00	25	0.02
19	2	0.00	27	0.02
18	7	0.00	34	0.02
17	10	0.01	44	0.03
16	7	0.00	51	0.03
15	25	0.02	76	0.05
14	37	0.02	113	0.08
13	43	0.03	156	0.11
12	63	0.04	219	0.15
11	92	0.06	311	0.21
10	125	0.08	436	0.29
9	200	0.13	636	0.43
8	300	0.20	936	0.63
7	508	0.34	1444	0.97
6	830	0.56	2274	1.53
5	1387	0.94	3661	2.47
4	2603	1.76	6264	4.23
3	4970	3.35	11234	7.58
2	10380	7.00	21614	14.59
1	24991	16.86	46605	31.45
0	101580	68.55	148185	100.00

Mean 0.65 Standard Deviation 1.41

**Table 6.D.3: ELA March 2004 Frequency Distribution of Total Number Erasures**

<b>Total Mark Changes</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
54	1	0.00	1	0.00
50	2	0.00	3	0.00
35	1	0.00	4	0.00
29	1	0.00	5	0.00
28	2	0.00	7	0.00
26	2	0.00	9	0.00
25	2	0.00	11	0.00
24	5	0.00	16	0.01
23	2	0.00	18	0.01
22	1	0.00	19	0.01
21	2	0.00	21	0.01
20	4	0.00	25	0.01
19	4	0.00	29	0.01
18	4	0.00	33	0.01
17	15	0.01	48	0.02
16	17	0.01	65	0.02
15	22	0.01	87	0.03
14	24	0.01	111	0.04
13	42	0.01	153	0.05
12	56	0.02	209	0.07
11	85	0.03	294	0.10
10	113	0.04	407	0.14
9	210	0.07	617	0.21
8	306	0.10	923	0.32
7	548	0.19	1471	0.50
6	905	0.31	2376	0.81
5	1735	0.59	4111	1.41
4	3554	1.22	7665	2.62
3	7360	2.52	15025	5.14
2	17354	5.94	32379	11.09
1	47552	16.28	79931	27.37
0	212159	72.63	292090	100.00
Mean	0.50	Std Deviation	1.15	

**Table 6.D.4: Mathematics March 2004 Frequency Distribution of Total Number Erasures**

<b>Total Mark Changes</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
68	1	0.00	1	0.00
62	1	0.00	2	0.00
37	1	0.00	3	0.00
32	1	0.00	4	0.00
30	3	0.00	7	0.00
28	2	0.00	9	0.00
26	5	0.00	14	0.00
25	2	0.00	16	0.01
24	6	0.00	22	0.01
23	9	0.00	31	0.01
22	5	0.00	36	0.01
21	14	0.00	50	0.02
20	6	0.00	56	0.02
19	19	0.01	75	0.03
18	20	0.01	95	0.03
17	31	0.01	126	0.04
16	50	0.02	176	0.06
15	41	0.01	217	0.07
14	66	0.02	283	0.10
13	90	0.03	373	0.13
12	131	0.04	504	0.17
11	203	0.07	707	0.24
10	299	0.10	1006	0.34
9	410	0.14	1416	0.48
8	659	0.22	2075	0.71
7	1101	0.38	3176	1.08
6	1862	0.64	5038	1.72
5	3064	1.05	8102	2.76
4	5472	1.87	13574	4.63
3	10617	3.62	24191	8.25
2	21762	7.42	45953	15.68
1	51047	17.42	97000	33.09
0	196116	66.91	293116	100.00

Mean 0.70 Standard Deviation 1.47

**Table 6.D.5: ELA May 2004 Frequency Distribution of Total Number Erasures**

<b>Total Mark Changes</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
31	1	0.01	1	0.01
19	1	0.01	2	0.02
14	2	0.02	4	0.03
13	5	0.04	9	0.07
12	2	0.02	11	0.09
11	4	0.03	15	0.12
10	6	0.05	21	0.17
9	7	0.06	28	0.23
8	8	0.07	36	0.30
7	18	0.15	54	0.45
6	38	0.32	92	0.76
5	86	0.71	178	1.48
4	138	1.15	316	2.62
3	337	2.80	653	5.42
2	800	6.64	1453	12.06
1	2217	18.40	3670	30.46
0	8379	69.54	12049	100.00

Mean 0.55                      Std Deviation 1.17

**Table 6.D.6: Mathematics May 2004 Frequency Distribution of Total Number Erasures**

<b>Total Mark Changes</b>	<b>Frequency</b>	<b>Percent</b>	<b>Cumulative Frequency</b>	<b>Cumulative Percent</b>
39	1	0.01	1	0.01
22	1	0.01	2	0.02
20	2	0.02	4	0.03
15	1	0.01	5	0.04
14	2	0.02	7	0.06
12	4	0.03	11	0.10
11	4	0.03	15	0.13
10	7	0.06	22	0.19
9	15	0.13	37	0.32
8	18	0.16	55	0.48
7	31	0.27	86	0.75
6	62	0.54	148	1.29
5	103	0.89	251	2.18
4	213	1.85	464	4.03
3	439	3.81	903	7.85
2	945	8.21	1848	16.06
1	2220	19.29	4068	35.35
0	7441	64.65	11509	100.00

Mean 0.69      Standard Deviation 1.37

**Table 6.D.7: Number and Percentage\* of Schools Meeting the Threshold Where Scores Changed from Not Pass to Pass as a Result of Erasures**

<b>Administration</b>	<b>Number of Schools</b>	<b>ELA</b>	<b>Mathematics</b>
<b>February 2004</b>	757	18 (2.38%)	36 (4.76%)
<b>March 2004</b>	1647	34 (2.06%)	109 (6.62%)
<b>May 2004</b>	1271	0 (0%)	5 (0.39%)

\* Percents are based on the number of schools testing in each administration.

## Chapter 7: Standard Setting<sup>5</sup>

The California High School Exit Examination (CAHSEE) is a criterion-based assessment designed to measure a student's competency in English language arts (ELA) and mathematics prior to graduating from California's public high schools. The mathematics and reading portions consist entirely of MC items that have undergone thorough content review. Blueprints were modified in July 2003 resulting in an ELA test that will be used beginning in 2004 that contains a single writing prompt. The mathematics test was modified slightly to align the measurement of the mathematics standards to be a more appropriate assessment of what the test is intending to measure. Content committees and the State Board deemed some of the concepts and items, such as using box-and-whisker plots, overly difficult given the small amount of time spent on the concepts in the classroom and the lower relevance of the material in post-high school settings. After decisions were made on modifying the blueprints, the California Department of Education (CDE) tasked the Educational Testing Service (ETS) to convene standard setting panels for the purpose of reviewing the passing scores marking passing performance in ELA and mathematics. This document describes the standard setting workshop and the results presented to SBE.

### Overview of the Process

ETS was directed to replicate the standard setting conducted on CAHSEE in 2001 to the greatest extent possible. Thus, ETS used the Bookmark Method, a procedure for setting passing scores that has widespread support in the measurement profession, has been used in more than 20 states, and has withstood legal challenges (see, for example, Lewis, et al., 1999 & Mitzel, et al. 2001). The Bookmark is typically a three-round standard setting process in which panelists work through a test booklet that has been re-ordered from easiest to hardest based on item difficulty. That is, the items are ordered based on how well students performed on them, so the items that students answered correctly most often will be followed by those they were more likely to answer incorrectly. Panelists are 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 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.

The workshops took place beginning Thursday, September 18, 2003 through Saturday, September 20, 2003. Thursday evening began with a welcome dinner, an overview of the CAHSEE, and an introduction to the Bookmark process. Friday commenced with panelists taking an abbreviated version of the test on which they would be setting standards. They were also given a chance to practice placing a bookmark, using CAHSEE items from a previously administered version of the test. Once panelists were comfortable with the procedure, ETS began the operational standard setting. One round of standard setting was conducted on Friday and the remaining two rounds and

---

<sup>5</sup> This chapter is based on the CAHSEE Technical Report from the Standard Setting Workshop, Final Report (November 6, 2003).

the follow up activities were completed on Saturday. Please see Appendix 7.A for a detailed agenda of the workshop.

### Characteristics of Standard Setting Panels

Two standard setting panels for each subject were convened for the workshop. One panel in each subject was composed primarily of California educators e.g., English or mathematics teachers, while the other was composed primarily of community members, e.g., community members, business representatives, district or school administrators, and college professors.

CDE worked with ETS to recruit four panels to participate in the workshop. The panels were comprised of 18-23 panelists that were divided among three tables of 6–9 panelists per table. CDE nominated one person from each table to be the table leader. The table leaders met two hours before the scheduled dinner on Thursday to receive extra training on how to facilitate table discussions. 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. Also, in the interest of equity, representatives from diverse geographic regions and from gender and major racial/ethnic subgroups were asked to participate. The resulting panels are described in Table 7.1.

**Table 7.1: Characteristics of the Standard Setting Panelists**

Educators			Community members		
	ELA	Math		ELA	Math
TOTAL NUMBER	23	22	TOTAL NUMBER	18	23
Panelist characteristics			Panelist characteristics		
<i>Gender</i>			<i>Gender</i>		
Male	4	9	Male	4	9
Female	19	13	Female	14	14
<i>Race/ethnicity</i>			<i>Race/ethnicity</i>		
White	18	18	White	12	17
African American	1	2	African American	1	1
Hispanic	2	2	Hispanic	5	3
Asian	1	0	Asian	0	2
Characteristics of their school or district <sup>1,2</sup>			Type of Community member		
<i>High Minority Population</i>			School or District Administrator	11	11
High percentage of Hispanics	11	10	College Instructor	2	3
High percentage of African American	7	4	Business Representative	2	5
High Poverty	7	11	Community Leader	3	1
High percentage of EL	9	10	Parent/Other*	0	3

<sup>1</sup>"High" is defined as above the state average.

<sup>2</sup>Categories are not mutually exclusive.

\*"Other" includes consultants and tutors.

### Materials

Prior to the standard setting workshop, panel members were provided with a letter describing the purpose and procedures of the passing score study, an agenda for the meeting (see Appendix 7.A), and the test blueprints for the ELA and mathematics sections of the CAHSEE. At the standard setting workshop, panelists received training materials, a set of practice materials, and a set of operational materials, including operational test items. Items were kept secure by assigning

panelists an individual identification number and giving them material marked with the same number. Each panelist was required to sign an affidavit of nondisclosure, check the material out and in each day, and was responsible for controlling all documents labeled with his or her ID number. ETS staff monitored each room to ensure no materials left the rooms. In addition to the live test items, panelists also received an item map, a panelist recording form, two practice tests, and an ordered test booklet containing items that were consistent with the test blueprints but ordered in terms of difficulty. The item map and ordered item booklet are described more fully below.

#### **Item Map**

The item map is a summary document displaying relevant information regarding each item. It shows the item number, the type of item (MC or CR), the correct answer for the MC items and the point value for the writing tasks, the scale score of the item, and the content strand measured by the item. The item map is ordered by difficulty in the same manner as the ordered test booklet. See Appendix 7.B for the Item Maps used for ELA and mathematics.

#### **Ordered Item Booklet**

The ordered *item booklet* contains all the information about the item that panelists need to complete the bookmark task. For the MC items, each page of the ordered item booklet shows the item, along with any passage or graphic, and the possible responses. The panelists could find the key to each item on the item map. For the writing task, the initial prompt, the scoring guide, and a sample student response for each possible score were provided.

#### **Evaluation Forms**

Panelists also received evaluation forms throughout the process both to gauge their understanding of the procedure and to collect various types of evidence for future reports (see Appendix 7.C for copies of the evaluation forms). One purpose of an evaluation form is to demonstrate the procedural validity of the Bookmark process. As described in a recent textbook on standard setting “Information about the panelists’ level of satisfaction with the performance descriptors, training, standard-setting process, and final standards is an important piece of the evidence for establishing the validity of the performance standards.” (Hambleton, 2001, p. 104) ETS was interested in knowing as soon as possible if panelists were not satisfied with the level of training they received. Therefore the first evaluation form was given to the panelists at the end of the first night to gauge their current understanding of the process and their comfort level with the tasks they would be performing. The evaluation forms were analyzed that night, and the next morning, facilitators reviewed with the panelists certain materials that appeared to be unclear. Although the overall ratings of their understanding were high, some panelists had issues they wanted addressed the following day. As a result of this review and discussion, panelists indicated they were comfortable with the process and ready to proceed. Specific comments and questions included:

- The item map is less than clear.
- How does the CR item score fit into the ELA item map and final score?
- How do we go from the final bookmark placement to the 350 scale score?
- Need more information to understand the sample item map and the scale score.
- It might have been good to include the full content standards alongside the CAHSEE blueprint standards for the people to review.
- Would like a definition sheet with words like “criterion-referenced test,” “SEM,” “median score” etc.

- What is the relationship between this bookmark process (for minimum pass rate) and the NCLB requirement for 3 performance levels?
- When do you predict the State Board of Education deciding on a new (if any) score for the CAHSEE?
- In placing the Bookmark, are we looking for the grade level of the standard? The 7<sup>th</sup>-grade knowledge of a 7<sup>th</sup>-grader? The 7<sup>th</sup>-grade standard that should be known by a 10<sup>th</sup>-grader? By a high school graduate?
- Need clarification on the scale score on the item map.
- Review the purpose of the discussion—do we have to reach consensus?

A second purpose of the evaluation forms is to provide information on how panelists made their judgments. This issue is particularly important in this standard setting workshop since two separate panels made judgments on the same set of items. In the event that the two panels provided different recommendations, ETS wanted to further understand how they came to their passing score and provide that information to the policy decision-makers. Therefore, at the end of each round of standard setting, the panelists were asked to complete a short survey rating the factors they considered when placing their bookmark. After round 3, the survey included questions regarding the influence of the impact data on their judgments. Finally, at the end of the workshop, the panelists completed a summative evaluation addressing the procedural validity of the standard setting process. All forms were submitted anonymously with only the subject area and panel (educator vs. community member) identified. A complete analysis of these evaluation forms is provided in the Results section.

## **Procedure**

The standard setting workshop was conducted using the Bookmark procedure in which panelists were instructed to set one passing score at the just sufficient level on the test they reviewed (ELA or mathematics). That is, the panelists were asked to find the point in the ordered item booklet, which defined the knowledge and skills needed to just pass the CAHSEE. The following sections describe the training, including practice for the panelists, the judgments they made, the analysis of these judgments, and the feedback data they received.

## **Training**

On the first evening, a general orientation session was held for the entire group (i.e., all four panels) where the need for a criterion-referenced passing score was explained. Geno Flores, the Deputy Superintendent for Accountability and Assessment of CDE gave an introductory speech orienting the panelists to the CAHSEE and updating them on recent changes both to the test and related policy. Deborah Franklin, the liaison to SBE also gave a short welcoming speech and explained how SBE would use the information provided by the panelists. Marianne Perie, a Senior Research Scientist at ETS and the Director of standard setting for CAHSEE, then introduced the Bookmark approach for setting a criterion-referenced passing score along with the expectations for panel members' participation. Day 2 provided an opportunity for more in-depth and hands-on training on the Bookmark method. At the beginning of Day 2 panelists were asked to take a test about half the length of the regular CAHSEE, including a writing prompt in ELA. The purpose of taking the test was to allow the panelists to experience the test under similar conditions to the student. No key was provided at this point. The facilitators then reviewed the answers with the panelists and went over

the rubric and sample responses for the writing prompt. In addition, now that the panelists were in their subject-specific groups, ETS content leads trained the panelists on the content standards measured by each test and facilitated a discussion on what it meant to have the knowledge and skills necessary to just pass the CAHSEE in terms of the content standards.

#### **Table Leader Training**

Prior to the welcome dinner ETS trained the table leaders in a two-hour session. The training began with a description of a table leader's role and responsibilities. They received instruction on the following tasks:

- Lead the review of the Ordered Item Booklet
  - Ensure group understands the activity
  - Keep group focused on the same item at the same time
  - Keep the group working together
  - Ensure all members are participating
  - Notify group leader of any problems
- Ensure panelists understand the task of placing bookmarks
  - Discuss understanding of content standards with panelists
  - Lead discussion on understanding of what it means to have sufficient knowledge and skills
  - Make sure panelists understand where to place bookmark
- Facilitate discussion in each round
  - Using the table of results, direct the panelists' attention to the items between the highest and lowest bookmark
  - Lead a discussion on what those items are measuring and whether a student with sufficient knowledge and skills should be able to answer them
- Collect and check all rating forms for completeness and accuracy
  - Collect group's rating forms
  - Check to see the forms have been filled out correctly by comparing the number on the reporting form to the bookmark in the ordered item booklet
  - Give forms to large-group leader
- Review and present feedback data to your table
  - Return forms to panelists
  - Show table of results, focusing on high and low scores
- Control secure materials
  - Before any break, instruct panelists to leave all green (secure) material on the table
  - Ask panelists to stack their material in the order provided by the form in their folder, with the ordered item booklet on the top

#### **Practice**

Prior to working with the full ordered-item booklet, panelists first had an opportunity to practice placing a bookmark on a short ordered item booklet made up of released items. A test of 16 MC items for mathematics and 12 MC items along with one writing prompt for ELA was rated to practice the modified Bookmark procedure. Panelists were repeatedly reminded that they were looking for the point in the ordered test book at which they felt students have demonstrated sufficient knowledge and skills to just pass the CAHSEE in that subject area.

Following the bookmark placement in the sample ordered-item booklet, ETS facilitators lead a discussion on why panelists placed the bookmark where they did. When the panelists were comfortable with the process, they returned their practice material and signed out the operational material.

## **Ratings**

Once the panelists were ready to begin the Bookmark task, they returned all practice materials and signed out the ordered-item booklet and item map.

They were then told to review the ordered item booklets in their table groups. During this review, they were instructed to answer and discuss two questions:

1. What does this item measure?
2. What makes it more difficult than the previous item?

The table leaders facilitated this discussion. The purpose of this exercise was for the panelists to gain a common understanding about what knowledge and skills each item requires. This stage is considered essential to setting a reasonable standard based on the amount of knowledge and skills students should know to pass the CAHSEE. In this phase of the workshop, however, panelists were cautioned not to discuss the placement of the bookmark but only to focus on individual items.

After they reviewed the ordered item booklet, the panelists were asked to place their first bookmark independently. They were told to “place a bookmark after the hardest item for which two-thirds of those students with sufficient knowledge and skills to just pass the CAHSEE would be able to answer correctly.” They were also told to consider the following pieces of information:

- California content standards, which specify the content that should be learned by students given implementation of the standards.
- CAHSEE blueprint, which maps the California content standards to the CAHSEE.
- Group discussions, which focused on the amount of knowledge and skills required for students to just pass the CAHSEE.

The panelists completed this bookmark task in three rounds.

## **Analysis**

After completing each bookmark placement, the panelists recorded the page number and item number on which they placed their bookmark. ETS then entered the item number into a spreadsheet. The spreadsheet calculated the median scale score related to these item placements and highlighted the highest and lowest recommended scores. This analysis was completed for each individual table after Rounds 1 and 2 and for the room as a whole after Rounds 2 and 3. The results of the ratings are presented in the Results section.

## **Feedback and Discussion**

Feedback was given to the panelists after each round, and they were given an opportunity to discuss the feedback in a group setting. After Round 1, ETS facilitators provided feedback to each table on the lowest, highest, and median bookmark rating within their table. Panelists then were given an opportunity to share with each other why they placed their bookmarks where they did.

After Round 2, panelists were given the same information at the table level, but they also received the same information for the room as a whole. That is, they were told what the median of all the tables was and the highest and lowest bookmark value in the room. Then, each table leader gave a

4–5 minute presentation on the types of considerations and concerns that were being discussed at their table.

Finally, consequence, or impact, data based on the census 10<sup>th</sup>-grade population who took the test in March and May 2003 were provided to the panelists. That is, ETS facilitators told the panelists the predicted percentage of students who would pass the assessment given the current median cut point. The panelists were then told to discuss within their tables all the information they had heard and place their third and final bookmark when they were ready. Consensus was not required for this workshop.

## Results

For both ELA and mathematics, ETS created tables showing every possible raw score point, the percentage correct on the test represented by that score, and the percentage of students who would score at or above it. The percentage of students scoring at or above the score point is based on pooled grade 10 data from the March and May 2003 administrations. The percentage shown is the predicted percentage of *first-time test-takers*, not the percentage of students predicted to pass that portion of the CAHSEE prior to graduation. The procedures used to develop the predicted consequence data are described in Appendix 7.D.

In addition, disaggregated consequence data are provided. That is, for an abbreviated range of scores the percentage of students predicted to score at or above that score point are displayed by race/ethnicity, gender, school performance, and special student populations. For race/ethnicity, the data were disaggregated by the four largest categories California uses for reporting Adequate Yearly Progress (AYP): African American, Hispanic, white (non-Hispanic), and Asian. School performance was calculated using CAHSEE data. Schools were divided into quintiles<sup>6</sup> based on their CAHSEE scores and then displayed the predicted percentage of students who attend schools falling in the highest and lowest quintiles who would score at or above each score point. The percentage of students meeting or exceeding each score point was calculated for special student populations in each of the following categories:

- Special education, which includes students with disabilities.
- English learners, as identified by the school-level test administrator.
- Socio-economically disadvantaged students, which is based on participation in the National School Lunch Program (NSLP).

The passing scores recommended by the panelists are identified by a labeled arrow in Tables 7.4, 7.5, and 7.6 for ELA and 7.7, 7.8, and 7.9 for mathematics. The recommended passing score shown for each group is the median passing score for that group. That is, the median passing score for both the educator group and the community group are shown as two separate recommended passing scores. Also identified are the scores representing the recommendation +/- one and two standard errors of judgment (SEJ). The SEJ is a measure of the degree of variability among the sets of panelists. Finally, the recommended passing scores +/- two standard errors of measurement (SEM) are identified. The SEM describes the expected variation between the examinees' true score and their observed score. By bracketing the median passing score by 2 SEMs, one can identify the 95% confidence interval, which can be used to estimate the effects of false positives (passing students who may not actually have sufficient knowledge and skills) or false negatives (failing students who

---

<sup>6</sup> These were California quintiles, not national quintiles, so each quintile represents 20% of the California population.

do have sufficient knowledge and skills). These standard errors can aid in identifying the appropriate passing score taking into consideration variance in the human judgments and imprecision in the test itself.

This report first presents the results of each round of the Bookmark procedure for each table. It then provides summary information for each subject, which was presented in a preliminary report used to make a final recommendation to SBE. Finally it presents the results of the evaluation forms. Because the ELA and mathematics groups were kept separate throughout the process, the results are presented in two separate sections below.

### **English Language Arts (ELA)**

Forty-one panelists participated in the Bookmark exercise for the ELA portion of the CAHSEE, 23 in the Educator’s group and 18 in the Community member’s group. Tables 7.2a and 7.2b display the high, low, and median scores for each table after each round. The median of the table medians was selected as the recommended passing score for each group and will be discussed further in the next section. The information displayed in Tables 7.2a and 7.2b, shows the type of information given to the panelists during the feedback stage of rounds 2 and 3. However, the panelists received the feedback in terms of the number of items referred to by the placement of the bookmark, while these data below show those bookmarks converted into raw score points with a maximum possible 90 points. See Appendix 7.E for a detailed explanation of how the bookmark placements were converted into raw scores.

**Table 7.2a ELA Educator Table Results by Round**

Minimum Score out of 90	Round 1	Round 2	Round 3
Table 1			
High	74	62	62
Low	56	56	56
Median	62	58	58
Table 2			
High	69	68	68
Low	63	63	62
Median	65	65	65
Table 3			
High	60	56	56
Low	55	55	55
Median	58	55	56

**Table 7.2b ELA Community Table Results by Round**

Minimum Score out of 90	Round 1	Round 2	Round 3
Table 4			
High	66	61	61
Low	54	58	58
Median	60	60	60
Table 5			
High	72	67	66
Low	60	62	62
Median	66	65	63
Table 6			
High	66	61	61
Low	58	60	59
Median	62	61	60

**Summary of ELA Panel Results**

Table 7.3a displays the median passing scores after the panelists’ bookmarks were converted to a raw score through an IRT model using an RP67. That is, ETS converted the item indicated by a panelist’s bookmark to the score point at which students would have a 67% probability of answering correctly the item indicated by the bookmark.<sup>7</sup> Table 7.3a displays the median raw score of the two ELA panels after each round. For example, row 1 shows that the ELA educators recommended a raw score of 62 points out of a maximum possible of 90 as a passing score after Round 1. They then dropped their median score to 58 in Round 2 and stayed at that same median score for Round 3.

**Table 7.3a: Summary Median Passing Scores for ELA Shown as the Minimum Number of Points Out of 90**

Panel	Round		
	1	2	3
ELA educators	62	58	58
ELA community	62	61	60

NOTE: Round 3 ratings are the final median passing scores.

These numbers represent the panelist bookmark placements converted to a raw score using the response probability (RP) 67.

The numbers in Table 7.3b represent the standard errors of judgment (SEJs) for each round. These numbers are the measures of variability in individual panelist ratings for each group of panelists from one round to the next. The panelists placed their bookmark on the item they thought separated the items into the skills and knowledge required to pass the CAHSEE from those not required, so the standard errors were also calculated in units of numbers of items. (See Appendix 7.E for a full definition of the standard error of judgment.) In general, panelists achieved a high degree of convergence over time.

<sup>7</sup> For more information on this conversion, see Attachment E entitled “Technical Note on Converting Panelist Bookmarks to Passing scores on the CAHSEE.”

**Table 7.3b: Standard Errors of Judgment of Bookmarks Placements by Round**

Panel	Round		
	1	2	3
ELA educators	2.33	2.39	2.23
ELA community	2.57	1.44	1.45

Table 7.4 shows the raw score, the percentage correct (raw score/maximum score points converted to a percentage), and the predicted percentage of students at or above each score. The raw score is the number of points a student must receive. In ELA this score is the number correct plus the weighted points on the writing task. The percentage correct is the percentage of points answered correctly. To calculate this percentage, the Raw Score was used as the numerator, and the maximum possible score was in the denominator, (90 in ELA). Finally, the predicted percentage of students at or above each passing score was calculated using the pooled scores of the 10<sup>th</sup>-graders who took the CAHSEE for the first time in March and May 2003. The numerator is the number of students receiving that score or higher, and the denominator is the number of test takers. It can be seen in Table 7.4 that the educators' median passing score (58) corresponds to receiving 64% of the maximum possible points, and 71% of students are predicted to achieve that level on their first try. The left hand column shows the community and educator median passing scores; these numbers are bolded. The error bands around the median passing scores are also indicated, identifying the score corresponding to +/- 1 SEJ, 2 SEJs, and 2 SEMs. The two standard errors of measurement (SEMs) always bracket the other scores and indicate the 95% confidence interval around the median passing score.

**Table 7.4: ELA Consequence Data for Every Score Point**

	Raw Score	Percentage Correct	Predicted Percentage of Students At or Above Each Score*
	90	100%	0%
	89	99%	0%
	88	98%	1%
	87	97%	2%
	86	96%	3%
	85	94%	5%
	84	93%	7%
	83	92%	10%
	82	91%	13%
	81	90%	15%
	80	89%	19%
	79	88%	22%
	78	87%	25%
	77	86%	28%
	76	84%	31%
	75	83%	34%
	74	82%	37%
	73	81%	40%
	72	80%	43%
	71	79%	45%
	70	78%	48%
	69	77%	50%
Community passing score + 2 SEMs ▶	68	76%	53%
	67	74%	55%
Educator passing score + 2 SEMs ▶	66	73%	57%
	65	72%	59%
	64	71%	61%
Community passing score + 2 SEJs ▶	63	70%	63%
Educator passing score + 2 SEJs ▶	62	69%	65%
Community passing score + 1 SEJ ▶	61	68%	66%
<b>Community median passing score ▶</b>	<b>60</b>	<b>67%</b>	<b>68%</b>
Community passing score - 1 SEJ ▶	59	66%	69%

**Table 7.4: ELA Consequence Data for Every Score Point (continued)**

	Raw Score	Percentage Correct	Predicted Percentage of Students At or Above Each Score*
<b>Educators' median passing score ►</b>	<b>58</b>	<b>64%</b>	<b>71%</b>
Community passing score - 2 SEJs ►	57	63%	72%
Educator passing score - 1 SEJ ►	56	62%	74%
	55	61%	75%
Educator passing score - 2 SEJs ►	54	60%	76%
	53	59%	78%
Community passing score - 2 SEMs ►	52	58%	79%
	51	57%	80%
Educator passing score - 2 SEMs ►	50	56%	81%
	49	54%	82%
	48	53%	83%
	47	52%	84%
	46	51%	85%
	45	50%	86%
	44	49%	87%
	43	48%	87%
	42	47%	88%
	41	46%	89%
	40	44%	90%
	39	43%	91%
	38	42%	91%
	37	41%	92%
	36	40%	93%
	35	39%	93%
	34	38%	94%
	33	37%	94%
	32	36%	95%
	31	34%	95%
	30	33%	96%
	29	32%	96%
	28	31%	97%
	27	30%	97%
	26	29%	97%
	25	28%	98%
	24	27%	98%
	23	26%	98%
	22	24%	99%
	21	23%	99%
	20	22%	99%
	19	21%	99%
	18	20%	99%
	17	19%	100%

**Table 7.4: ELA Consequence Data for Every Score Point (continued)**

Raw Score	Percentage Correct	Predicted Percentage of Students At or Above Each Score*
16	18%	100%
15	17%	100%
14	16%	100%
13	14%	100%
12	13%	100%
11	12%	100%
10	11%	100%
9	10%	100%
8	9%	100%
7	8%	100%
6	7%	100%
5	6%	100%
4	4%	100%
3	3%	100%
2	2%	100%
1	1%	100%
0	0%	100%

\*Percentage of grade 10 first-time test-takers based on the combined March and May 2003 results.

Table 7.5 is similar to Table 7.4, but it disaggregates the consequence data by race/ethnicity and school performance. The first two columns as well as the markers are identical to Table 7.4. The four largest race/ethnicity categories are displayed both in Table 7.5 and in Figure 7.1 for ELA. Figure 7.1 shows the predicted percentage of students who would pass at each potential passing score, where the passing score is shown as the percent correct on the test. There is one curve for each racial-ethnic group, allowing the reader to choose a percent correct point and draw a line up to each curve to see the percentage of students of each major ethnic group who would pass. This figure also allows the reader to see the gaps between the racial-ethnic groups at different passing scores. The school-level category in Table 7.5 was calculated from school-level analyses of March and May CAHSEE data for 10<sup>th</sup> graders. Thus, the bottom quintile shows the score distributions for students in the 20% of California schools with the lowest average CAHSEE ELA performance, and the top quintile shows the score distributions for students in the 20% of California schools with the highest average CAHSEE ELA performance. Then, the predicted percentage of students in each of these schools who would pass the CAHSEE as grade 10 first-time test takers was calculated for elected score points.

**Table 7.5: ELA Consequence Data by Race/ethnicity and School Performance**

	Raw Score	Percent Correct	Predicted Percentage of Students At or Above Each Score <sup>1</sup>				Lowest <sup>2</sup> 20% of Schools	Highest 20% of Schools
			African American	Asian	Hispanic	White		
	72	80%	26%	56%	24%	61%	5%	65%
	71	79%	29%	58%	26%	64%	5%	68%
	70	78%	31%	60%	28%	67%	6%	70%
	69	77%	34%	62%	31%	69%	7%	72%
Community passing score + 2 SEMs ▶	68	76%	36%	64%	33%	72%	8%	75%
	67	74%	39%	66%	35%	74%	9%	76%
Educator passing score + 2 SEMs ▶	66	73%	41%	68%	38%	76%	10%	78%
	65	72%	44%	70%	40%	77%	11%	80%
	64	71%	46%	72%	42%	79%	12%	81%
Community passing score + 2 SEJs ▶	63	70%	48%	73%	44%	81%	14%	83%
Educator passing score + 2 SEJs ▶	62	69%	50%	74%	46%	82%	15%	84%
Community passing score + 1 SEJ ▶	61	68%	53%	76%	48%	83%	17%	85%
<b>Community median passing score ▶</b>	<b>60</b>	<b>67%</b>	<b>55%</b>	<b>77%</b>	<b>50%</b>	<b>84%</b>	<b>18%</b>	<b>86%</b>
Community passing score - 1 SEJ ▶	59	66%	57%	78%	52%	85%	19%	87%
<b>Educators' median passing score ▶</b>	<b>58</b>	<b>64%</b>	<b>58%</b>	<b>79%</b>	<b>54%</b>	<b>86%</b>	<b>21%</b>	<b>88%</b>
Community passing score - 2 SEJs ▶	57	63%	60%	80%	56%	87%	22%	89%
Educator passing score - 1 SEJ ▶	56	62%	62%	81%	58%	88%	24%	89%
	55	61%	64%	82%	60%	89%	26%	90%
Educator passing score - 2 SEJs ▶	54	60%	66%	83%	62%	90%	27%	91%
	53	59%	67%	84%	63%	91%	29%	91%
Community passing score - 2 SEMs ▶	52	58%	69%	85%	65%	91%	31%	92%
	51	57%	70%	86%	67%	92%	33%	93%
Educator passing score - 2 SEMs ▶	50	56%	72%	87%	68%	92%	35%	93%
	49	54%	73%	87%	70%	93%	36%	93%
	48	53%	75%	88%	72%	93%	38%	94%
	47	52%	76%	89%	73%	94%	40%	94%
	46	51%	78%	90%	74%	94%	42%	95%
	45	50%	79%	90%	76%	95%	45%	95%
	44	49%	80%	91%	77%	95%	47%	96%
	43	48%	81%	92%	79%	95%	49%	96%
	42	47%	83%	92%	80%	96%	51%	96%
	41	46%	84%	93%	81%	96%	53%	96%
	40	44%	84%	93%	82%	96%	55%	97%
	39	43%	85%	94%	83%	97%	57%	97%
	38	42%	86%	94%	84%	97%	59%	97%
	37	41%	87%	95%	85%	97%	61%	97%
	36	40%	88%	95%	86%	97%	64%	98%

<sup>1</sup>Percentage of Grade 10 first-time test-takers based on the combined March and May 2003 results.

<sup>2</sup>Lowest and highest 20% of schools is determined by school-level performance on the March and May 2003 CAHSEE ELA tests and includes all students in the lowest and highest performing schools, regardless of their individual scores.

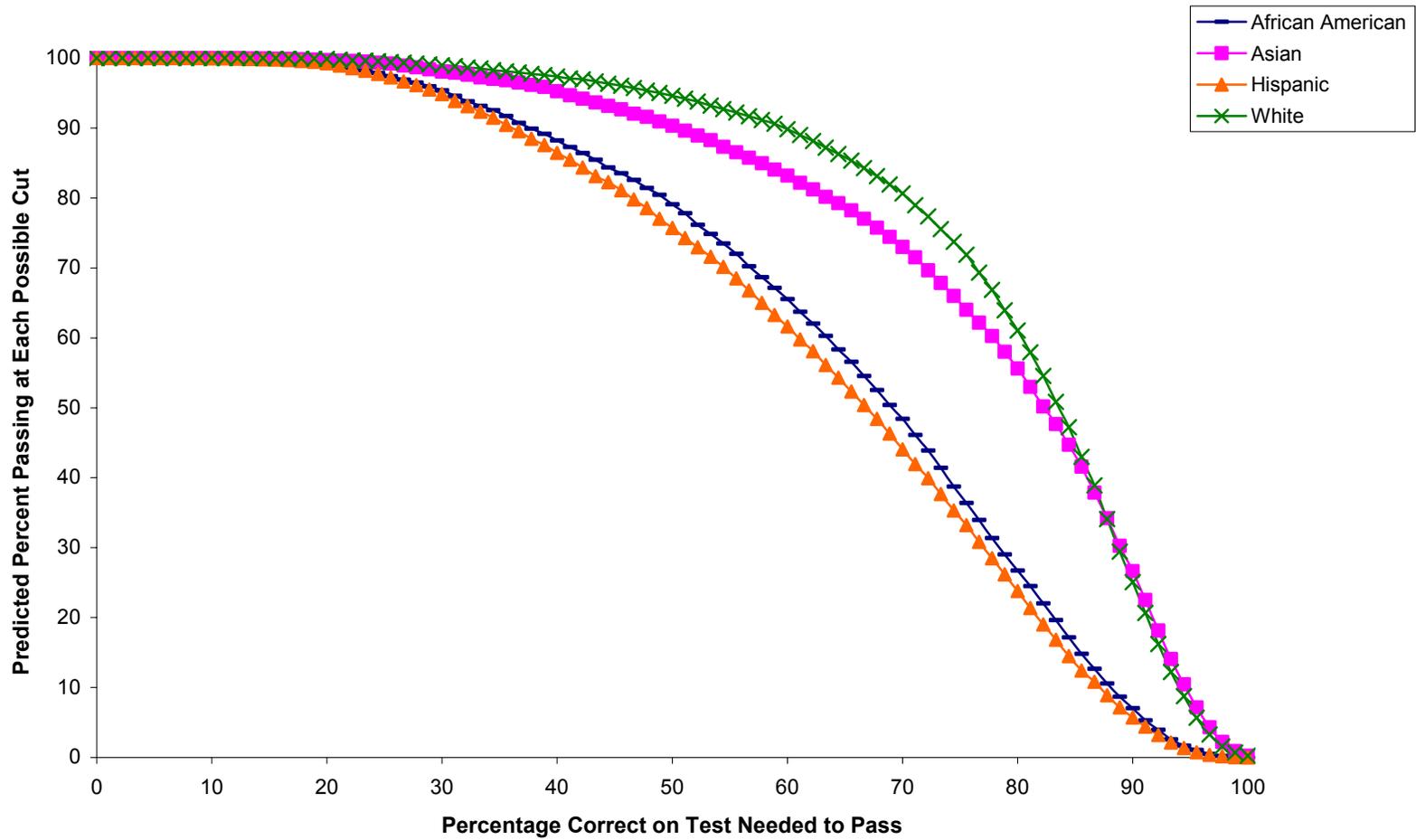
Table 7.6 is similar to Tables 7.4 and 7.5, but the consequence data are disaggregated by gender and special populations. The first two columns as well as the markers are identical to Tables 7.4 and 7.5. The next two columns report the percentage of males and females expected to pass at selected passing scores. The last three columns focus on the passing rate of special populations. The NSLP column shows the percentage of students who indicated that they receive free and reduced lunch through the National School Lunch Program. The next column refers to those students identified as English Learners, and the last column represents students with disabilities.

**Table 7.6: ELA Consequence Data by Gender and Special Populations**

	Raw Score	Percent Correct	Predicted Percentage of Students At or Above Each Score*				
			Male	Female	NSLP Students	English Learners	Special Education
	72	80%	37%	48%	22%	6%	12%
	71	79%	40%	50%	25%	7%	14%
	70	78%	42%	53%	27%	8%	15%
	69	77%	45%	55%	29%	10%	16%
Community passing score + 2 SEMs ▶	68	76%	47%	57%	32%	11%	17%
	67	74%	49%	60%	34%	12%	19%
Educator passing score + 2 SEMs ▶	66	73%	51%	62%	36%	14%	20%
	65	72%	53%	63%	38%	15%	21%
	64	71%	55%	65%	40%	17%	22%
Community passing score + 2 SEJs ▶	63	70%	57%	67%	42%	19%	24%
Educator passing score + 2 SEJs ▶	62	69%	59%	69%	44%	20%	25%
Community passing score + 1 SEJ ▶	61	68%	61%	70%	46%	22%	27%
<b>Community median passing score ▶</b>	<b>60</b>	<b>67%</b>	<b>63%</b>	<b>72%</b>	<b>48%</b>	<b>24%</b>	<b>28%</b>
Community passing score - 1 SEJ ▶	59	66%	64%	73%	50%	26%	29%
<b>Educators' median passing score ▶</b>	<b>58</b>	<b>64%</b>	<b>66%</b>	<b>75%</b>	<b>52%</b>	<b>28%</b>	<b>31%</b>
Community passing score - 2 SEJs ▶	57	63%	67%	76%	55%	30%	32%
Educator passing score - 1 SEJ ▶	56	62%	69%	78%	57%	32%	34%
	55	61%	70%	79%	58%	34%	36%
Educator passing score - 2 SEJs ▶	54	60%	71%	80%	60%	37%	37%
	53	59%	73%	81%	61%	39%	39%
Community passing score - 2 SEMs ▶	52	58%	74%	82%	63%	41%	41%
	51	57%	75%	83%	65%	44%	42%
Educator passing score - 2 SEMs ▶	50	56%	76%	84%	66%	46%	44%
	49	54%	78%	85%	68%	48%	46%
	48	53%	79%	86%	69%	51%	48%
	47	52%	80%	87%	71%	53%	49%
	46	51%	81%	88%	73%	55%	51%
	45	50%	82%	89%	74%	57%	53%

\*Percentage of grade 10 first-time test-takers based on the combined March and May 2003 results.

Figure 7.1: ELA Passing Scores by Race/ethnicity



## Mathematics

Forty-five panelists participated in the Bookmark exercise for the mathematics portion of the CAHSEE, 22 in the educator group and 23 in the community group. Tables 7.7a and 7.7b display the high, low, and median scores for each table after each round. The median of the table medians was selected as the recommended passing score for each group and will be discussed further in the next section. The information displayed in Tables 7.8a and 7.8b shows the type of information given to the panelists during the feedback stage of rounds 2 and 3. However, the panelists received the feedback in terms of the number of items referred to by the placement of the bookmark, while these data below show those bookmarks converted into raw score points with a maximum of 80 points. See Appendix 7.E for a detailed explanation of how the bookmark placements were converted into raw scores.

**Table 7.7a Mathematics Educator Table Results by Round**

Minimum Score out of 80	Round 1	Round 2	Round 3
Table 1			
High	60	55	55
Low	50	53	53
Median	55	55	53
Table 2			
High	60	58	57
Low	51	52	50
Median	56.5	55	52
Table 3			
High	58	56	55
Low	50	49	49
Median	56.5	55	53

**Table 7.7b Mathematics Community Table Results by Round**

Minimum Score out of 80	Round 1	Round 2	Round 3
Table 4			
High	62	59	56
Low	56	56	52
Median	60	58	56
Table 5			
High	58	56	56
Low	47	51	50
Median	52.5	51.5	52
Table 6			
High	60	58	58
Low	53	56	53
Median	57.5	57	56

### Summary of Mathematics Panel Results

The entries in Table 7.8a show the median passing scores after the panelists' bookmarks were converted to a raw score through an IRT model using an RP67.<sup>8</sup> Table 7.8a displays the median raw score of the two panels after each round. Table 7.8b represents the standard errors of judgment (SEJs) for each round. These numbers are the measures of variability in individual panelist ratings for each group of panelists from one round to the next. The standard errors were calculated in units of numbers of items, because the panelists placed their bookmark on the item they thought separated the items in those skills and knowledge required to pass the CAHSEE from those not required. In general, panelists achieved some degree of convergence over time. For example, in the last row one can see that the mathematics community panel started out with a higher variability (2.26 items) after the first round but grew more similar in their ratings over rounds as their variability decreased in Round 2 to 1.60 and in Round 3 to 1.29.

**Table 7.8a: Summary Median Mathematics Passing scores by Round**

Panel	Round		
	1	2	3
Mathematics educators	57	55	53
Mathematics community	58	58	56

NOTE: Round 3 ratings are the final recommended passing scores.

These numbers represent the panelist bookmark placements converted to a raw score using the response probability (RP) 0.67.

**Table 7.8b: Standard Errors of Judgment of Bookmark Placements by Round**

Panel	Round		
	1	2	3
Mathematics educators	1.78	1.13	1.27
Mathematics community	2.26	1.60	1.29

Table 7.9 shows the raw score, the percentage correct, and the predicted percentage of students at or above each score. The raw score is the number of points a student must receive. In mathematics this translates to the number of items answered correctly. The percentage correct is the percentage of points answered correctly using the Raw Score as the numerator and the maximum possible score in the denominator (80 in mathematics). Finally, the predicted percentage of students at or above each passing score was calculated using the pooled scores of the 10<sup>th</sup>-graders who took the CAHSEE for the first time in March and May 2003. The numerator is the number of students receiving that score or higher, and the denominator is the number of test takers. The table shows that the educators' median passing score (53) corresponds to receiving 66% of the maximum possible points, and 52% of students are predicted to achieve that level on their first try. The left hand column shows the community and educator median passing scores; these numbers are bolded. The table also indicates the error bands around the median passing scores, identifying the score

<sup>8</sup> For more information on this conversion, see Appendix 7E entitled "Technical Note on Converting Panelist Bookmarks to Passing scores on the CAHSEE."

corresponding to +/- 1 SEJ, 2 SEJs, and 2 SEMs. The two standard errors of measurement (SEMs) always bracket the other scores and indicated the 95% confidence interval around the median passing score.

**Table 7.9: Mathematics Consequence Data for Every Score Point**

	Raw Score	Percentage Correct	Predicted Percentage of Students At or Above Each Score*
	80	100%	0%
	79	99%	1%
	78	98%	3%
	77	96%	5%
	76	95%	7%
	75	94%	9%
	74	93%	11%
	73	91%	13%
	72	90%	15%
	71	89%	17%
	70	88%	19%
	69	86%	21%
	68	85%	23%
	67	84%	25%
	66	83%	27%
	65	81%	29%
	64	80%	31%
Community passing score + 2 SEMs ▶	63	79%	33%
	62	78%	35%
	61	76%	37%
Educator passing score + 2 SEMs ▶	60	75%	38%
Community passing score + 2 SEJs ▶	59	74%	40%
	58	73%	42%
Community passing score + 1 SEJ ▶	57	71%	44%
<b>Community median passing score ▶</b>	<b>56</b>	<b>70%</b>	<b>47%</b>
Community passing score - 1 SEJ ▶	55	69%	48%
Educator passing score + 1 SEJ ▶	54	68%	50%
<b>Educators' median passing score ▶</b>	<b>53</b>	<b>66%</b>	<b>52%</b>
Educator passing score - 1 SEJ ▶	52	65%	54%
	51	64%	55%
Educator passing score - 2 SEJs ▶	50	63%	57%
Community passing score - 2 SEMs ▶	49	61%	59%
	48	60%	61%
	47	59%	63%
Educator passing score - 2 SEMs ▶	46	58%	64%
	45	56%	66%
	44	55%	67%
	43	54%	69%

**Table 7.9: Mathematics Consequence Data for Every Score Point (continued)**

Raw Score	Percentage Correct	Predicted Percentage of Students At or Above Each Score*
42	53%	71%
41	51%	73%
40	50%	74%
39	49%	76%
38	48%	78%
37	46%	80%
36	45%	81%
35	44%	83%
34	43%	84%
33	41%	85%
32	40%	87%
31	39%	88%
30	38%	90%
29	36%	91%
28	35%	92%
27	34%	94%
26	33%	95%
25	31%	95%
24	30%	96%
23	29%	97%
22	28%	98%
21	26%	98%
20	25%	99%
19	24%	99%
18	23%	99%
17	21%	99%
16	20%	100%
15	19%	100%
14	18%	100%
13	16%	100%
12	15%	100%
11	14%	100%
10	13%	100%
9	11%	100%
8	10%	100%
7	9%	100%
6	8%	100%
5	6%	100%
4	5%	100%
3	4%	100%
2	3%	100%
1	1%	100%
0	0%	100%

\*Percentage of Grade 10 first-time test-takers based on the combined March and May 2003 results.

Table 7.10 is similar to Table 7.9, but it displays the consequence data disaggregated by race/ethnicity and school performance. The first two columns as well as the markers are identical to Table 7.9, followed by the four largest race/ethnicity categories. Data for these groups are also shown in Figure 7.2, which is identical to Figure 7.1, except it displays mathematics data. The reader can choose a percent correct point on the x-axis and draw a line up to each curve to identify the percentage of students of different race/ethnicities who would pass and to see the gaps between the racial-ethnic groups at different scores. The school-level category in Table 7.10 was calculated from the March and May CAHSEE data for 10<sup>th</sup> graders. Then, the predicted percentage of students in each of these schools who would pass the CAHSEE as first-time test-takers was calculated for a range of passing scores.

**Table 7.10: Mathematics Consequence Data by Race/ethnicity and School Performance**

	Raw Score	Percent Correct	Predicted Percentage of Students At or Above Each Score <sup>1</sup>				Lowest <sup>2</sup> 20% of Schools	Highest 20% of Schools
			African American	Asian	Hispanic	White		
Community passing score + 2 SEMs	64	80%	13%	60%	15%	44%	1%	42%
▶	63	79%	14%	62%	16%	47%	1%	46%
	62	78%	15%	64%	17%	49%	2%	48%
	61	76%	17%	66%	19%	52%	2%	51%
Educator passing score + 2 SEMs	60	75%	18%	68%	21%	55%	2%	54%
Community passing score + 2 SEJs	59	74%	20%	70%	22%	57%	3%	57%
▶	58	73%	21%	71%	24%	59%	3%	59%
Community passing score + 1 SEJ	57	71%	22%	73%	26%	61%	3%	61%
<b>Community median passing score</b>	<b>56</b>	<b>70%</b>	<b>24%</b>	<b>74%</b>	<b>28%</b>	<b>63%</b>	<b>4%</b>	<b>64%</b>
Community passing score - 1 SEJ	55	69%	26%	76%	29%	65%	4%	66%
Educator passing score + 1 SEJ	54	68%	27%	77%	32%	67%	5%	68%
<b>Educators' median passing score</b>	<b>53</b>	<b>66%</b>	<b>29%</b>	<b>79%</b>	<b>34%</b>	<b>69%</b>	<b>5%</b>	<b>70%</b>
Educator passing score - 1 SEJ	52	65%	31%	80%	36%	71%	6%	72%
	51	64%	33%	81%	38%	73%	7%	74%
Educator passing score - 2 SEJs	50	63%	35%	82%	40%	75%	8%	76%
Community passing score - 2 SEMs	49	61%	37%	83%	42%	76%	9%	78%
▶	48	60%	38%	84%	44%	77%	10%	79%
	47	59%	40%	85%	46%	79%	11%	80%
Educator passing score - 2 SEMs	46	58%	43%	86%	48%	80%	13%	82%
▶	45	56%	45%	87%	50%	81%	14%	83%
	44	55%	47%	87%	52%	83%	16%	85%
	43	54%	49%	88%	55%	84%	17%	86%
	42	53%	51%	89%	57%	85%	19%	87%
	41	51%	53%	90%	59%	86%	21%	88%
	40	50%	55%	91%	62%	87%	23%	89%
	39	49%	57%	92%	64%	88%	25%	89%
	38	48%	59%	92%	67%	89%	27%	90%
	37	46%	61%	93%	69%	90%	29%	91%
	36	45%	64%	94%	71%	91%	32%	92%

<sup>1</sup>Percentage of Grade 10 first-time test-takers based on the combined March and May 2003 results.

<sup>2</sup>Lowest and highest 20% of schools is determined by school-level performance on the March and May 2003 CAHSEE ELA tests and includes all students in the lowest and highest performing schools, regardless of their individual scores.

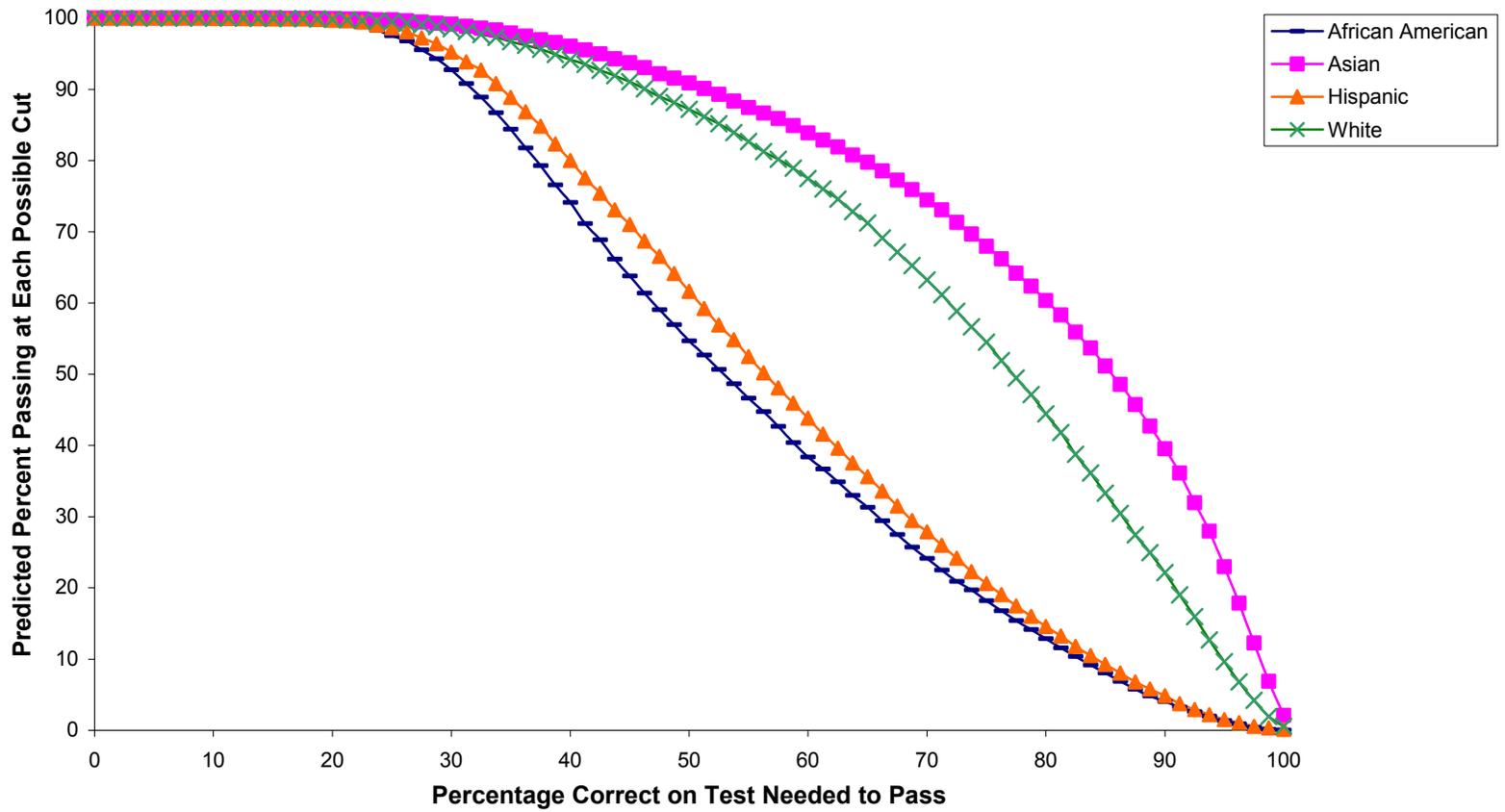
Table 7.11 is similar to Tables 7.9 and 7.10, but it displays the consequence data disaggregated by gender and special populations. The first two columns as well as the markers are identical to Tables 7.9 and 7.10. The next two columns report the percentage of males and females expected to pass at selected passing scores. The last three columns focus on the passing rate of special populations. The NSLP column shows the percentage of students who indicated that they receive free and reduced lunch through the National School Lunch Program. The next column refers to those students identified as English Learners, and the last column represents those students with disabilities.

**Table 7.11: Mathematics Consequence Data by Gender and Special Populations**

	Raw Score	Percent Correct	Predicted Percentage of Students At or Above Each Score				
			Male	Female	NSLP Students	English Learners	Special Education
	64	80%	32%	30%	16%	9%	9%
Community passing score + 2 SEMs ▶	63	79%	34%	32%	17%	10%	9%
	62	78%	36%	34%	19%	11%	10%
	61	76%	38%	36%	20%	12%	11%
Educator passing score + 2 SEMs ▶	60	75%	40%	38%	22%	13%	12%
Community passing score + 2 SEJs ▶	59	74%	41%	40%	24%	14%	13%
	58	73%	43%	42%	26%	16%	14%
Community passing score + 1 SEJ ▶	57	71%	45%	44%	27%	17%	14%
<b>Community median passing score ▶</b>	<b>56</b>	<b>70%</b>	<b>47%</b>	<b>46%</b>	<b>29%</b>	<b>19%</b>	<b>15%</b>
Community passing score - 1 SEJ ▶	55	69%	49%	48%	31%	20%	17%
Educator passing score + 1 SEJ ▶	54	68%	51%	49%	33%	22%	18%
<b>Educators' median passing score ▶</b>	<b>53</b>	<b>66%</b>	<b>52%</b>	<b>52%</b>	<b>35%</b>	<b>23%</b>	<b>19%</b>
Educator passing score - 1 SEJ ▶	52	65%	54%	53%	37%	25%	20%
	51	64%	56%	55%	39%	27%	21%
Educator passing score - 2 SEJs ▶	50	63%	57%	57%	41%	29%	22%
Community passing score - 2 SEMs ▶	49	61%	59%	59%	42%	30%	24%
	48	60%	61%	61%	45%	32%	25%
	47	59%	63%	63%	47%	34%	26%
Educator passing score - 2 SEMs ▶	46	58%	64%	64%	49%	36%	28%
	45	56%	66%	66%	51%	38%	29%
	44	55%	67%	68%	53%	40%	31%
	43	54%	69%	70%	55%	43%	32%
	42	53%	70%	72%	57%	45%	34%
	41	51%	72%	73%	59%	47%	36%
	40	50%	73%	75%	61%	50%	38%
	39	49%	75%	76%	64%	52%	39%
	38	48%	77%	78%	66%	55%	41%
	37	46%	78%	80%	68%	58%	43%
	36	45%	80%	82%	71%	61%	45%

\*Percentage of Grade 10 first-time test takers based on the combined March and May 2003 results.

**Figure 7.2: Mathematics Passing Scores by Race/Ethnicity**



## Evaluation of the Bookmark Process

Panelists were asked to complete an evaluation form after each round rating the usefulness of different training exercises and the influence of various factors on their bookmark placement. One question asked about the degree of influence of various factors on their ratings, and this question included new factors after each round. For example, in Round 1, the panelists considered the influence of the standards, their knowledge of the test and the items, and their familiarity with the student population. During Round 2, they also had information about where their peers placed their bookmarks, and they had participated in a table discussion. Finally, for Round 3 they had information about the large-group bookmarks and the percentage of first-time test-takers likely to pass, in addition to all the other information from previous rounds. For each of the factors, they indicated whether that factor was “very influential,” “somewhat influential,” or “not influential.”

### Evaluation Results from the ELA panel

Table 7.12 displays the percentage of ELA panelists who felt each factor was very influential.

**Table 7.12: Percentage of ELA Panelists Reporting that Various Factors Were Very Influential in Placing the Bookmark**

Factor	Educators			Community members		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
State Content Standards	55%	58%	70%	63%	71%	61%
Difficulty of the items	90%	83%	96%	74%	94%	61%
Personal experiences with students	90%	88%	91%	74%	82%	78%
Consequences of the test	55%	46%	61%	26%	47%	44%
Table Discussion <sup>1</sup>	—	92%	100%	—	76%	83%
Bookmark placements of other panelists <sup>1</sup>	—	29%	26%	—	24%	33%
Large Group Discussion <sup>2</sup>	—	—	48%	—	—	50%
Percentage of students likely to pass <sup>2</sup>	—	—	52%	—	—	39%

<sup>1</sup>Only relevant after Round 1 placement

<sup>2</sup>Only relevant after Round 2 placement

As seen in Table 7.12, throughout the process, ELA panelists were influenced by the difficulty of the items and their personal experience with students, although these factors were generally more influential to educators than to community members. Likewise, in the first two rounds, community members were more influenced by the state content standards than were the educators, although over half of all panelists said that the state content standards were very influential at all times. In both Rounds 2 and 3, all panelists were highly influenced by the table discussions. The large group discussion in Round 3 was less influential. Finally, the consequences of the test and the percentage of students likely to pass had more influence on the educators’ judgments than on the community members.

In the comments section of the evaluations, ELA panelists indicated that they appreciated the group discussion because it allowed them to hear “different points of view.” Others mentioned that they found the table discussions to be both “clarifying” and “reassuring.” In addition, one panelist noted

after Round 3 that “The large group discussion and the student passing rate were both influential. They helped me to narrow in on the bookmark that I feel is most appropriate.”

Panelists also filled out a final evaluation that was analyzed for evidence of validity of the standard setting process. Specific questions used to examine the validity of the process, again based on Hambleton (2001), were

1. The bookmark training was appropriate in giving me the information I needed to complete my assignment. (Strongly agree, agree, disagree, strongly disagree)
2. How comfortable are you with the location of the final group median Bookmark? (very comfortable, somewhat comfortable, somewhat uncomfortable, very uncomfortable)
3. How comfortable would you be defending this process to your peers? (very comfortable, somewhat comfortable, somewhat uncomfortable, very uncomfortable)

Nineteen out of the 23 ELA educators indicated that they strongly agreed that the training was appropriate, and the remaining four agreed. Thirteen of the 18 ELA community members also strongly agreed and the other five also agreed. Of the 21 ELA educators who responded to question 8, 16 of them were comfortable with the location of the final median bookmark placement (9 of those very comfortable). Of those who were uncomfortable, they provided the following comments: “I am concerned about the amount of students not passing.” and “Our table placed the bookmark much higher than two other panels.” So, one panelist who said he was somewhat uncomfortable with the final placement thought it was too high, while another thought it was too low. All but one ELA community member was comfortable with the final median bookmark placement. The one who expressed some discomfort also indicated that the median bookmark was placed too low. “I feel it was set too low. I’m an administrator and responsible for the remediation of students who don’t pass but I feel strongly that we need to set a respectable bar.” Finally, for question 10, only one ELA educator indicated that she would be somewhat uncomfortable defending the process to her peers, because “It [is] terribly difficult to explain a process...to a person not involved in the process.” One ELA community member was concerned that it was a very subjective process. Everyone else who responded to this item (21 educators and 15 community members) indicated that they would be either somewhat or very comfortable defending the process to their peers.

#### **Evaluation Results from the Mathematics Panel**

As with ELA, the mathematics panelists indicated whether various factors were “very influential,” “somewhat influential,” or “not influential.” Table 7.13 displays the percentage of mathematics panelists who felt each factor was very influential.

**Table 7.13: Percentage of Mathematics Panelists Reporting that Various Factors Were Very Influential in Placing the Bookmark**

Factor	Educators			Community members		
	Round 1	Round 2	Round 3	Round 1	Round 2	Round 3
State Content Standards	64%	48%	45%	48%	50%	48%
Difficulty of the items	64%	76%	75%	64%	63%	60%
Personal experiences with students	77%	81%	90%	60%	75%	64%
Consequences of the test	59%	67%	80%	48%	54%	64%
Table Discussion <sup>1</sup>	—	71%	75%	—	71%	72%
Bookmark placements of other panelists <sup>1</sup>	—	33%	35%	—	13%	24%
Large Group Discussion <sup>2</sup>	—	—	75%	—	—	44%
Percentage of students likely to pass <sup>2</sup>	—	—	75%	—	—	64%

<sup>1</sup>Only relevant after Round 1 placement

<sup>2</sup>Only relevant after Round 2 placement

As seen in Table 7.13, mathematics panelists were influenced most by their personal experience with students, although these factors were generally more influential to educators than to community members. The difficulty of the items also influenced both groups throughout the process. The state content standards had more influence in Round 1 for the educators than in subsequent rounds. In both Rounds 2 and 3, all panelists were highly influenced by the table discussions. The large group discussion in Round 3 was less influential for the community members and equally influential as the table discussions for the educators. Finally, the consequences of the test and the percentage of students likely to pass had more influence on the educators’ judgments than on the community members’.

When asked what materials, information, or procedures were most influential, one educator commented after Round 1 “The idea of 2/3 of the students w/sufficient knowledge & skills. The idea of the content standard in comparison to the actual problem.” After Round 2, a community member commented that he was influenced by “more detailed discussion w/table discussions. What would it take for 67% students to pass CAHSEE?” Many of the mathematics educators commented on the helpfulness of the table discussions after both Rounds 2 and 3. After Round 3, one educator commented that the impact data had a large influence on her. “Its so frustrating to see such a low pass rate despite all the years CAHSEE has been given. But change is slow.” Another expressed a similar viewpoint. “To hear how all teachers felt was helpful even though there were differing opinions. Each teacher justified his or her bookmark placements well. An illustrative comment was “I think it helped us expand our own thinking. Hearing the consequences, finding out that the expected pass rate was actually lower was surprising. Somewhat distressful, also.” A community member described what influenced her most this way: “A combination of factors. Since the subjective element is as valid as the technical data, that had a profound effect in the end. However, the most useful factor was when we focused on the kid who was last in line for his diploma- ‘Could he have gotten this question correct?’”

Panelists also filled out a final evaluation form that was analyzed for evidence of validity of the standard setting process. Specific questions used to examine the validity of the process were

1. The bookmark training was appropriate in giving me the information I needed to complete my assignment. (Strongly agree, agree, disagree, strongly disagree)
2. How comfortable are you with the location of the final group median Bookmark? (very comfortable, somewhat comfortable, somewhat uncomfortable, very uncomfortable)
3. How comfortable would you be defending this process to your peers? (very comfortable, somewhat comfortable, somewhat uncomfortable, very uncomfortable)

Seventeen out of the 22 mathematics educators indicated that they strongly agreed that the training was appropriate, and the remaining five agreed. Sixteen of the 23 mathematics community members also strongly agreed and the other seven also agreed. Of the 19 mathematics educators who responded to question 8, 18 of them were comfortable with the location of the final median bookmark (7 of those very comfortable). The one who was uncomfortable wrote: “too high for reality.” Three mathematics community members were somewhat uncomfortable with the final median bookmark placement. Two of them indicated that they felt the median bookmark was too high, and the other wrote, “I believe that there is a wide difference of expectations for what students will be able to do.” Finally, for question 10, all educators and community members indicated that they would be at least somewhat comfortable defending the process to their peers, and more than half indicated they would be very comfortable.

Most of the overall comments on the process were very positive. Several panelists indicated that they learned something from this experience and that they were now more confident in the CAHSEE. Some comments relevant to the standard setting workshop included:

1. Provide impact data earlier.
2. Include high school students or recent graduates in the group of panelists.
3. Bring the groups together at the end so that mathematics and ELA can benefit from hearing each other’s discussions.

## **Conclusions**

Based on these results, the CDE provided specific recommendations about the placement of the CAHSEE passing scores to the SBE. In November 2003, after reviewing the results of this study, the California State Board of Education voted to maintain the current CAHSEE passing score of 60% of the items correct on ELA (54 out of 90) and 55% of the items correct on mathematics (44 out of 80).

## Appendix 7.A: CAHSEE Bookmark Workshop Agenda

### Thursday, September 18, 2003

3:00 p.m. – 4:45 p.m.	Table Leader Training
5:15 p.m. – 6:15 p.m.	Welcome Dinner
6:15 p.m. – 7:00 p.m.	Opening Remarks
7:00 p.m. – 8:30 p.m.	Overview of Goals and Introduction to the Bookmark Procedure

### Friday, September 19, 2003

8:00 a.m. – 8:30 a.m.	Continental Breakfast
8:30 a.m. – 9:00 a.m.	Review Goals, Agenda, & the Bookmark Procedure
9:00 a.m. – 10:00 p.m.	Take Test and Review Responses
10:00 a.m. – 11:00 a.m.	Introduce Subject-specific Content Standards
11:00 a.m. – 11:15 a.m.	Break
11:15 a.m. – 12:15 a.m.	Practice Placing Bookmarks
12:15 a.m. – 12:30 a.m.	Answer Any Last Questions on the Bookmark Procedure
12:30 p.m. – 1:30 p.m.	Lunch
1:30 p.m. – 3:30 p.m.	Review Ordered Test Booklet
3:30 p.m. – 3:45 p.m.	Break
3:45 p.m. – 4:30 p.m.	Place Round 1 Bookmark and Complete Evaluation
4:30 p.m. – 5:00 p.m.	Review Tomorrow's Agenda and Dismiss

## **Agenda (continued)**

### **Saturday, September 20, 2003**

8:00 a.m. – 8:30 a.m.	Continental Breakfast
8:30 a.m. – 8:45 a.m.	Review of Goals & Agenda
8:45 a.m. – 10:00 a.m.	Provide Feedback on Round 1 Placements and Facilitate Small-Group Discussions
10:00 a.m. – 10:30 a.m.	Place Round 2 Bookmark and Complete Evaluation
10:30 a.m. – 10:50 a.m.	Break
10:50 a.m. – 11:30 a.m.	Provide Feedback on Round 2 Placements and Facilitate Large-Group Discussion
11:30 a.m. – 12:15 p.m.	Provide Consequence Data and Facilitate Small-Group Discussion
12:15 p.m. – 12:30 p.m.	Place Round 3 Bookmark and Complete Evaluation
12:30 p.m. – 1:30 p.m.	Lunch
1:30 p.m. – 3:00 p.m.	Final Debriefing, Evaluation, and Dismissal

## Appendix 7.B: Item Maps

### Item Map for CAHSEE English Language Arts

Original Position	New Position Number	Key	Scale Score*	Content Standard	Comments
CR	1	1	13	Writing Applications	
3	2	A	20	Literary Response & Analysis	
72	3	A	24	Writing Conventions	
68	4	B	25	Writing Conventions	
24	5	A	25	Literary Response & Analysis	
4	6	D	28	Literary Response & Analysis	
2	7	B	29	Literary Response & Analysis	
69	8	D	30	Writing Conventions	
1	9	A	30	Word Analysis	
61	10	B	31	Writing Conventions	
22	11	B	33	Literary Response & Analysis	
8	12	C	34	Reading Comprehension	
31	13	B	34	Literary Response & Analysis	
CR	14	2	34	Writing Applications	
44	15	C	35	Reading Comprehension	
13	16	A	37	Word Analysis	
26	17	B	40	Word Analysis	
64	18	D	40	Writing Conventions	
6	19	B	40	Literary Response & Analysis	
11	20	C	41	Reading Comprehension	
41	21	B	41	Reading Comprehension	
12	22	B	43	Word Analysis	
16	23	B	43	Reading Comprehension	
62	24	A	43	Writing Conventions	
42	25	D	44	Reading Comprehension	
25	26	D	44	Literary Response & Analysis	
65	27	C	46	Writing Conventions	
49	28	C	47	Writing Strategies	
28	29	C	47	Word Analysis	
30	30	C	48	Literary Response & Analysis	
60	31	C	48	Writing Conventions	
59	32	C	48	Writing Conventions	
63	33	B	49	Writing Conventions	
38	34	C	49	Word Analysis	
70	35	A	49	Writing Conventions	
27	36	B	51	Literary Response & Analysis	
67	37	C	52	Writing Conventions	
58	38	C	52	Writing Strategies	
7	39	C	52	Reading Comprehension	
45	40	A	52	Reading Comprehension	
18	41	D	55	Literary Response & Analysis	

33	42	D	55	Reading Comprehension	
54	43	C	55	Writing Strategies	
23	44	C	56	Literary Response & Analysis	
47	45	D	57	Writing Strategies	
9	46	D	57	Reading Comprehension	
51	47	D	58	Writing Strategies	
32	48	C	58	Reading Comprehension	
66	49	B	58	Writing Conventions	
5	50	D	59	Literary Response & Analysis	
43	51	B	60	Reading Comprehension	
34	52	B	61	Reading Comprehension	
14	53	D	61	Reading Comprehension	
15	54	A	62	Reading Comprehension	
46	55	B	67	Writing Conventions	
57	56	A	69	Writing Strategies	
50	57	A	69	Writing Strategies	
21	58	B	70	Literary Response & Analysis	
48	59	A	71	Writing Strategies	
53	60	A	71	Writing Strategies	
55	61	C	74	Writing Strategies	
40	62	A	75	Literary Response & Analysis	
35	63	A	78	Reading Comprehension	
52	64	B	78	Writing Strategies	
19	65	A	78	Literary Response & Analysis	
10	66	C	79	Reading Comprehension	
17	67	A	79	Reading Comprehension	
39	68	A	79	Literary Response & Analysis	
37	69	A	83	Word Analysis	
29	70	D	85	Literary Response & Analysis	
36	71	A	88	Literary Response & Analysis	
CR	72	3	89	Writing Applications	
71	73	B	90	Writing Conventions	
56	74	D	91	Writing Strategies	
20	75	B	105	Literary Response & Analysis	
CR	76	4	160	Writing Applications	

Note. \*This is an arbitrary scale used during the standard setting only.

## Item Map for CAHSEE Mathematics

Original Position	New Position Number	Key	Scale Score*	Primary Content Standard	Comments
22	1	A	23	Statistics, Data Analysis & Probability	
17	2	B	25	Number Sense	
72	3	C	26	Measurement & Geometry	
39	4	B	34	Algebra & Functions	
54	5	D	35	Algebra & Functions	
64	6	B	35	Measurement & Geometry	
61	7	D	35	Measurement & Geometry	
2	8	D	36	Number Sense	
55	9	A	38	Algebra & Functions	
37	10	B	41	Algebra & Functions	
1	11	D	42	Number Sense	
34	12	B	44	Statistics, Data Analysis & Probability	
53	13	D	45	Algebra & Functions	
33	14	A	49	Statistics, Data Analysis & Probability	
41	15	A	49	Algebra & Functions	
56	16	D	50	Algebra & Functions	
75	17	C	50	Measurement & Geometry	
50	18	A	50	Algebra & Functions	
42	19	D	51	Algebra & Functions	
23	20	C	52	Statistics, Data Analysis & Probability	
5	21	D	54	Number Sense	
38	22	A	55	Algebra & Functions	
91	23	B	55	Algebra I	
16	24	D	55	Number Sense	
65	25	C	56	Measurement & Geometry	
68	26	A	57	Measurement & Geometry	
45	27	C	57	Algebra & Functions	
30	28	C	60	Statistics, Data Analysis & Probability	
66	29	D	60	Measurement & Geometry	
25	30	B	61	Statistics, Data Analysis & Probability	
78	31	C	61	Measurement & Geometry	
62	32	C	62	Measurement & Geometry	
14	33	C	63	Number Sense	
29	34	B	64	Statistics, Data Analysis & Probability	
47	35	B	65	Algebra & Functions	
81	36	A	65	Algebra I	
24	37	C	66	Statistics, Data Analysis & Probability	

3	38	A	66	Number Sense
67	39	C	66	Measurement & Geometry
31	40	B	67	Statistics, Data Analysis & Probability
63	41	C	69	Measurement & Geometry
4	42	B	69	Number Sense
71	43	C	69	Measurement & Geometry
40	44	D	70	Algebra & Functions
43	45	B	70	Algebra & Functions
9	46	D	71	Number Sense
15	47	C	73	Number Sense
7	48	B	74	Number Sense
49	49	D	75	Algebra & Functions
46	50	C	75	Algebra & Functions
92	51	B	77	Algebra I
44	52	D	78	Algebra & Functions
12	53	C	80	Number Sense
48	54	B	81	Algebra & Functions
74	55	C	81	Measurement & Geometry
89	56	B	81	Algebra I
82	57	C	82	Algebra I
27	58	C	83	Statistics, Data Analysis & Probability
11	59	A	85	Number Sense
6	60	C	86	Number Sense
85	61	C	86	Algebra I
32	62	B	86	Statistics, Data Analysis & Probability
28	63	B	86	Statistics, Data Analysis & Probability
8	64	C	87	Number Sense
76	65	A	87	Measurement & Geometry
88	66	B	92	Algebra I
13	67	A	92	Number Sense
69	68	B	95	Measurement & Geometry
51	69	B	95	Algebra & Functions
73	70	D	97	Measurement & Geometry
84	71	A	97	Algebra I
83	72	A	99	Algebra I
90	73	C	100	Algebra I
87	74	A	101	Algebra I
86	75	B	101	Algebra I
77	76	B	108	Measurement & Geometry
26	77	B	113	Statistics, Data Analysis & Probability
10	78	C	115	Number Sense
70	79	A	119	Measurement & Geometry
52	80	A	121	Algebra & Functions

Note. \*This is an arbitrary scale used during the standard setting only.

## Appendix 7.C: Evaluation Forms

### Initial Evaluation of the Training on the Bookmark Procedure for the California High School Exit Examination 9/17/03

The purpose of this first evaluation form is to secure your feedback about the training you have received so far on the Bookmark process. Your feedback will provide a basis for determining what to review prior to starting the Bookmark process tomorrow.

Please complete the information below. Do not put your name on the form as we want your feedback to be anonymous.

**Content area:** Mathematics  English Language Arts

**Group representing:** Teachers  Educational Administrators  Higher Education   
Business/Community/Other

1. Please read each of the following statements carefully. Place a check mark (✓) under one category (Strongly agree, Agree, Disagree, or Strongly disagree) to indicate the degree to which you agree with each statement.

	Strongly agree	Agree	Disagree	Strongly disagree
a. I understand the purpose of this workshop.				
b. The trainer explained things clearly.				
c. The item map seems easy to understand.				
d. The ordered item booklet seems easy to understand.				
e. The training on the Bookmark process was sufficient in giving me the information I need to place a bookmark.				
f. I know what I will be doing the next two days.				
g. I need more information about the process before I would feel comfortable placing a bookmark.				
h. After being given a chance to practice tomorrow, I am sure I will be ready to place a bookmark on the CAHSEE				

2. Have you participated in a Bookmark or other standard setting workshop before today?  
\_\_\_\_\_ Yes \_\_\_\_\_ No
3. Do you have any suggestions for areas that we should spend more time training on tomorrow?

**Round 1 Evaluation of the Bookmark Workshop for the  
California High School Exit Examination 9/19/03**

The purpose of this second evaluation form is to secure your feedback about the how you made your judgment on where you placed your bookmark. Your feedback will provide a basis for evaluating the training, methods, and materials in the Bookmark process.

Please complete the information below. Do not put your name on the form, as we want your feedback to be anonymous.

**Content area:** Mathematics  English Language Arts

**Group representing:** Teachers  Educational Administrators  Higher Education   
Business/Community/Other

1. Please rate the usefulness of the following materials or procedures in placing your bookmark.

	Very useful	Somewhat useful	Not at all useful
a. Taking the exam prior to placing a bookmark			
b. Reviewing the ordered item booklet at my table			
c. Information in the item map			
d. Practicing the procedure with real items			

2. How influential was each of the following factors in placing your bookmark?

	Very influential	Somewhat influential	Not influential
a. The description of evidence			
b. State content standards			
c. Your perception of the difficulty of the items			
d. Your personal experiences with students or young adults			
e. The consequences of the test			

3. What materials, information, or procedures were most influential in your placement of the bookmark? Why?

**Round 2 Evaluation of the Bookmark Workshop for the  
California High School Exit Examination 9/20/03**

The purpose of this evaluation form is to secure your feedback about the judgments you made in placing your second bookmark. Your feedback will provide a basis for evaluating the training, methods, and materials in the Bookmark process.

Please complete the information below. Do not put your name on the form, as we want your feedback to be anonymous.

**Content area:** Mathematics  English Language Arts

**Group representing:** Teachers  Educational Administrators  Higher Education   
Business/Community/Other

1. Please rate the usefulness of the following materials or procedures in placing the second bookmark.

	Very useful	Somewhat useful	Not at all useful
a. Content standards			
b. The description of evidence			
c. Information in the item map			
d. Table discussions			

2. How influential was each of the following factors in placing your bookmark in this round?

	Very influential	Somewhat influential	Not influential
a. The description of evidence			
b. State content standards			
c. Your perception of the difficulty of the items			
d. Your personal experiences with students or young adults			
e. Table discussion			
f. The bookmark placements of other panelists			
g. The consequences of the test			

3. What materials, information, or procedures were most influential in your second placement of the bookmark? Why?

**Round 3 Evaluation of the Bookmark Workshop for the  
California High School Exit Examination 9/20/03**

The purpose of this evaluation form is to secure your feedback about the judgments you made in placing your final bookmark. Your feedback will provide a basis for evaluating the training, methods, and materials in the Bookmark process.

Please complete the information below. Do not put your name on the form, as we want your feedback to be anonymous.

**Content area:** Mathematics  English Language Arts   
**Group representing:** Teachers  Educational Administrators  Higher Education   
 Business/Community/Other

1. Please rate the usefulness of the following materials or procedures in placing this last bookmark.

	Very useful	Somewhat useful	Not at all useful
a. The description of evidence			
b. State content standards			
c. Whole group discussions			
d. Impact information (% of students passing)			
e. Table discussions			

2. How influential was each of the following factors in placing your final bookmark?

	Very influential	Somewhat influential	Not influential
a. The description of evidence			
b. State content standards			
c. Your perception of the difficulty of the items			
d. Your personal experiences with students or young adults?			
e. Table discussion			
f. Large group discussion			
g. The bookmark placements of other panelists			
h. The percentage of students who will probably pass			
i. The consequences of the test for the students			

3. What materials, information, or procedures were most influential in your final placement of the bookmark? Why?

## Final Evaluation of the Bookmark Workshop for the California High School Exit Examination 9/20/03

The purpose of this final evaluation form is to secure your feedback about the overall Bookmark process. Your feedback will provide a basis for evaluating the training, methods, and materials in the Bookmark process.

Please complete the information below. Do not put your name on the form as we want your feedback to be anonymous.

**Content area:** Mathematics  English Language Arts

**Group representing:** Teachers  Educational Administrators  Higher Education   
Business/Community/Other

1. Please read each of the following statements carefully. Place a check mark (✓) under one category (Strongly agree, Agree, Disagree, or Strongly disagree) to indicate the degree to which you agree with each statement.

	Strongly agree	Agree	Disagree	Strongly disagree
a. I understood the purpose of this workshop.				
b. The training guide contained all the information I needed to complete my assignment.				
c. The item map was easy to understand.				
d. The ordered test booklet was easy to understand.				
e. The bookmark training was appropriate in giving me the information I needed to complete my assignment.				
f. The training on the description of evidence was appropriate in giving me the information I needed to complete my assignment.				
g. The training on how to place the bookmark was appropriate in giving me the information I needed to complete my assignment.				

2. Please rate the clarity of the following materials used in the Bookmark process.

	Very clear	Somewhat clear	Somewhat unclear	Very unclear
a. Instructions provided in the training materials				
b. Instructions provided by the facilitators				
c. Description of evidence				

3. Please rate the usefulness of the following materials or procedures in completing the Bookmark process.

	Very useful	Somewhat useful	Not at all useful
a. Taking the exam prior to placing a Bookmark			
b. Reviewing the ordered item booklet at my table			
c. Information in the item map			
d. The description of evidence			
e. Practicing the procedure with real items			
f. Table discussions			
g. Whole group discussions			
h. Impact information (% of students passing)			

4. How influential was each of the following factors in placing your bookmark?

	Very influential	Somewhat influential	Not influential
a. The description of evidence			
b. State content standards			
c. My perception of the difficulty of the items			
d. My personal experiences with students or young adults			
e. Table discussion			
f. Large group discussion			
g. The bookmark placements of other panelists			
h. The percentage of students who will probably pass			
i. The consequences of the test for the students			

5. Were there any materials or procedures that became more (or less) influential in your placement of the bookmark from one round to another? If so, which ones? Why?

6. How appropriate was the amount of time you were given to complete the different components of the Bookmark process?

	Too much time	About right	Too little time
a. Training on bookmark			
b. Taking the test			
c. Scoring the test			
d. Reviewing the ordered items			
e. Small group discussion after 1 <sup>st</sup> bookmark			
f. Large group discussion after 2 <sup>nd</sup> bookmark			
g. Final small group discussion after 2 <sup>nd</sup> bookmark			

7. Do you have additional comments about this process?

8. How comfortable are you with the location of the final group median Bookmark?

\_\_\_ Very comfortable \_\_\_ Somewhat comfortable \_\_\_ Somewhat uncomfortable \_\_\_ Very uncomfortable

If you checked uncomfortable, please explain why.

9. What proportion of just sufficient students do you think will pass the CAHSEE if the Board adopts the final group median Bookmark as the passing score? \_\_\_\_\_

10. How comfortable would you be defending this process to your peers?

\_\_\_ Very comfortable \_\_\_ Somewhat comfortable \_\_\_ Somewhat uncomfortable \_\_\_ Very uncomfortable

If you checked uncomfortable, please explain why.

11. Do you have any suggestions on how to improve the training and implementation of the Bookmark workshop?

## **Appendix 7.D: Technical Note on Procedures to Generate Predicted Consequences Data for the CAHSEE Standard Setting**

The CAHSEE standard setting utilized two test forms that were developed according to the revised CAHSEE blueprints. These forms had not been administered previously. To inform policy decisions related to the standard setting, ETS carried out a series of analyses to predict score distributions on these new test forms, using data for 10<sup>th</sup> graders from the March and May 2003 CAHSEE administrations. This technical note describes these analyses.

The analyses involved the following steps:

1. Ability distributions for a total group of 10<sup>th</sup> grade students and for 10<sup>th</sup> grade students in various demographic groups (i.e., gender, ethnicity, English language learner, etc.) were estimated using an Item Response Theory (IRT) procedure developed by Mislevy (1984). A computer program called RESOLVE was used to accomplish this estimation. The program was run on data for 10<sup>th</sup> grade students from the March and May 2003 administrations. Program inputs also included Rasch difficulty estimates for the items calibrated in these administrations.
2. The program outputs were listed for each of 41 quadrature points on the ability scale that ranged from  $-4.0$  to  $+4.0$  in increments of  $0.2$ . Associated with each quadrature point was a weight that represented the proportion of students in a group estimated to be at that ability level. The quadrature weights for March and May were combined in proportion to the sample sizes for each administration. The result of this combining process was a single set of weights for each test and demographic group.
3. IRT-based procedures were employed to generate simulated data using difficulty parameter estimates for the two standard setting forms and student abilities based on the quadrature points and weights. The computer program RESGEN 4.0 (Muraki, 2000) was used in these simulations. A total of 10,000 cases were generated with true abilities in proportion to the quadrature weights obtained in Step 2. Each simulated case consisted of scores of 1 or 0 on each MC test item and a score on the CR item of 0, 1, 1.5, 2, 2.5, 3, 3.5, or 4.
4. For each simulated case, number-correct scores were obtained by summing item-level scores. The distribution of number-correct scores was compiled for the simulations for all students and for each demographic subgroup. For ELA, a total score was calculated that applied a weight of 4.5 to the CR item score (for a total maximum score of 90).
5. IRT true-score equating procedures were used to equate scores on the test forms used for standard setting to the CAHSEE scale. Based on this equating, scale scores on the current 250-450 CAHSEE scale were assigned.

Table 7.D.1 presents the quadrature weights for the total groups. This table includes the weights from the March sample, the weights from the May sample, and the combined weights. Tables 7.D.2 and 7.D.3 for mathematics and ELA, respectively, present the combined March and May weights for each of the demographic subgroups included in the study. In the simulations conducted for each group, the number of cases at a given true ability level was equal to the quadrature weight times 10,000 and rounded to the nearest integer. For this reason, the combined weights shown in Tables

7.D.1 to 7.D.3 have also been multiplied by 10,000, although they have not been rounded but rather reflect the full precision displayed in the RESOLVE program output.

As a check on the accuracy of the simulations, the quadrature weights in Tables 7.D.1 to 7.D.3 were used with the item parameters for the March 2003 Mathematics and ELA forms to generate simulated data. These simulated data were then compared to the real March 2003 data. Figure 7.D.1 provides one result of these checks. In this figure, the predicted and actual percentages of students at each raw score point on the MC portion of the English language arts test are plotted. There are two plots, one for Males and one for Females. The predicted and actual percentages in both plots are very similar and are symmetrically distributed around the identity line. Similar plots were obtained for all demographic subgroups across English language arts and Mathematics. In general, the predictions of the March 2003 score distributions for the various subgroups appeared to be reasonably accurate.

A second check on the accuracy of the simulations was done once the true-score equatings were completed and predicted distributions of CAHSEE scale scores were generated. We compared the predicted percentages of students passing the standard-setting forms under the current passing standard to the observed combined March and May 2003 passing rates. Table 7.D.4 summarizes these comparisons and indicates that the actual and predicted passing rates are nearly identical across all of the comparison groups shown. Based on these comparisons, ETS is confident that the predicted consequences data provide an accurate basis for considering the impact of alternate CAHSEE passing scores.

**Table 7.D.1: Quadrature Weights for March and May 2003 10<sup>th</sup> Graders**

Quad Pt	<u>Mathematics</u>			<u>English Language Arts</u>		
	Mar 03 Wt	May 03 Wt	Comb Wt	Mar 03 Wt	May 03 Wt	Comb Wt
-4.0	0.000008	0.000043	0.100	0.000000	0.000000	0.000
-3.8	0.000014	0.000055	0.163	0.000000	0.000000	0.000
-3.6	0.000022	0.000079	0.252	0.000000	0.000000	0.000
-3.4	0.000032	0.000127	0.374	0.000001	0.000000	0.009
-3.2	0.000042	0.000209	0.515	0.000002	0.000001	0.019
-3.0	0.000049	0.000314	0.640	0.000003	0.000006	0.032
-2.8	0.000056	0.000388	0.748	0.000005	0.000017	0.057
-2.6	0.000066	0.000392	0.845	0.000008	0.000033	0.094
-2.4	0.000090	0.000381	1.065	0.000014	0.000049	0.159
-2.2	0.000163	0.000467	1.802	0.000034	0.000060	0.354
-2.0	0.000407	0.000861	4.327	0.000103	0.000075	1.014
-1.8	0.001283	0.002291	13.401	0.000349	0.000146	3.377
-1.6	0.004234	0.007309	44.081	0.001105	0.000580	10.759
-1.4	0.012046	0.021908	126.042	0.002981	0.003210	29.937
-1.2	0.026244	0.047671	274.568	0.006532	0.012593	68.685
-1.0	0.042487	0.067763	439.177	0.011448	0.025737	122.414
-0.8	0.053449	0.072330	545.177	0.016387	0.032053	172.568
-0.6	0.057508	0.072509	583.571	0.020495	0.033311	212.066
-0.4	0.058418	0.075377	593.779	0.024282	0.034816	248.669
-0.2	0.059079	0.079001	602.066	0.028484	0.039794	291.120
0.0	0.060144	0.080572	613.003	0.033311	0.047262	340.856
0.2	0.060936	0.077218	618.576	0.038398	0.053383	392.300
0.4	0.060705	0.068320	611.360	0.043378	0.059779	442.886
0.6	0.059623	0.057017	594.755	0.048642	0.064950	495.475
0.8	0.057572	0.047359	569.939	0.054556	0.067932	552.987
1.0	0.054108	0.041164	533.753	0.061047	0.071553	616.303
1.2	0.049631	0.036261	488.742	0.067772	0.074381	681.390
1.4	0.044787	0.030530	439.800	0.074018	0.072943	739.583
1.6	0.039775	0.024624	389.174	0.078724	0.066358	780.374
1.8	0.034874	0.019687	340.144	0.080314	0.057880	790.684
2.0	0.030451	0.015742	296.184	0.077148	0.049592	756.180
2.2	0.026553	0.012447	257.546	0.068727	0.041309	672.047
2.4	0.022980	0.009691	222.278	0.056165	0.032451	548.483
2.6	0.019552	0.007489	188.692	0.041785	0.023131	407.493
2.8	0.016262	0.005812	156.705	0.028200	0.014795	274.557
3.0	0.013217	0.004553	127.266	0.017285	0.008799	168.138
3.2	0.010524	0.003600	101.321	0.009650	0.005121	93.985
3.4	0.008220	0.002868	79.171	0.004938	0.002981	48.293
3.6	0.006288	0.002290	60.617	0.002304	0.001667	22.686
3.8	0.004695	0.001831	45.329	0.000996	0.000855	9.882
4.0	0.003405	0.001451	32.944	0.000410	0.000394	4.091

Notes. Combined weights are based on March and May sample sizes (N).

For March, Mathematics N is 389,612; ELA N is 376,509.

For May, Mathematics N is 23,376; ELA N is 22,134.

**Table 7.D.2: Quadrature Weights for Combined March and May 2003 10<sup>th</sup> Graders by Subgroup - Mathematics**

Quad	Female	Male	Am Ind	Asian	Pac Isl	Filipino	Hisp	Black	White	Disable	NSLP	Lower 20%	Upper 20%	Eng-Lmrr
-4.0	0.083	0.489	3.839	0.000	0.000	0.000	0.091	1.007	0.010	0.707	0.073	7.647	0.000	0.102
-3.8	0.157	0.243	1.349	0.000	0.000	0.000	0.203	0.420	0.031	1.208	0.170	2.360	0.001	0.148
-3.6	0.285	0.282	0.361	0.000	0.000	0.000	0.378	0.634	0.088	1.832	0.379	1.620	0.002	0.243
-3.4	0.432	0.495	0.133	0.000	0.000	0.000	0.570	1.160	0.230	2.542	0.750	2.447	0.054	0.426
-3.2	0.542	0.960	0.124	0.000	0.000	0.000	0.784	2.046	0.472	3.412	1.207	6.147	0.621	0.679
-3.0	0.590	1.943	0.304	0.000	0.000	0.029	0.934	2.965	0.726	4.573	1.551	16.572	1.382	0.931
-2.8	0.564	3.357	1.254	0.000	0.000	0.183	0.951	3.573	0.881	5.799	1.612	32.523	0.552	1.109
-2.6	0.540	3.909	4.086	0.001	0.000	0.530	0.911	3.755	0.936	6.572	1.468	40.765	0.073	1.169
-2.4	0.573	3.355	6.626	0.041	0.000	0.771	0.938	2.896	0.984	6.410	1.399	36.329	0.023	1.240
-2.2	0.825	3.244	7.324	0.191	0.146	0.897	1.211	2.866	1.254	6.178	1.699	29.943	0.062	1.608
-2.0	1.774	5.085	8.200	0.457	6.739	1.380	2.319	5.475	2.190	8.952	3.151	32.428	0.592	3.152
-1.8	5.816	13.395	8.549	1.269	25.977	3.735	7.189	16.388	5.492	24.863	9.248	63.087	3.481	10.106
-1.6	24.265	48.538	22.560	5.673	34.573	12.926	32.572	64.126	16.750	113.141	38.664	222.003	9.251	46.587
-1.4	92.629	174.337	105.075	28.783	56.467	37.525	148.377	245.676	48.674	528.542	165.008	847.034	26.826	215.329
-1.2	248.704	430.699	323.638	88.516	206.964	101.544	443.339	638.520	111.901	1396.434	471.358	1740.475	106.649	649.499
-1.0	439.125	622.768	589.974	149.417	523.097	205.224	758.147	988.188	191.254	1681.581	776.513	1723.059	248.566	1080.145
-0.8	554.689	620.014	670.838	176.960	611.746	275.119	867.430	993.264	255.880	1179.889	858.762	1331.420	242.718	1157.613
-0.6	588.665	573.355	577.618	199.794	587.535	306.517	847.985	857.840	301.025	779.545	817.041	1053.453	197.841	1047.378
-0.4	601.917	581.317	538.774	251.280	652.872	362.101	821.998	789.339	346.310	646.515	784.218	752.386	246.437	941.441
-0.2	621.709	623.686	620.959	308.734	689.355	464.263	810.144	763.576	402.534	602.039	776.288	546.746	392.902	856.587
0.0	646.845	656.695	737.552	330.658	667.283	571.089	803.869	711.579	468.794	529.113	768.132	486.342	530.782	767.305
0.2	658.204	647.067	695.037	351.293	660.186	623.294	764.616	646.193	537.809	427.106	730.886	409.441	552.830	653.688
0.4	642.431	600.570	619.491	405.360	582.960	634.462	674.482	600.898	599.738	347.007	650.044	255.047	594.661	534.066
0.6	617.243	556.965	655.938	464.422	528.556	650.076	579.664	531.640	653.697	314.369	560.922	116.043	726.897	434.896
0.8	590.388	525.451	667.720	498.713	671.438	675.600	495.285	437.260	690.077	274.586	488.198	51.912	824.294	349.213
1.0	548.911	489.388	580.271	508.823	806.645	699.646	413.300	360.833	693.290	212.189	417.517	40.339	814.534	267.266
1.2	496.239	440.880	452.722	502.083	566.636	707.727	342.861	300.010	666.719	169.085	345.513	63.149	757.783	200.077
1.4	442.089	390.776	381.163	511.572	333.528	660.232	283.799	245.754	621.492	145.035	283.624	55.650	676.100	158.593
1.6	384.620	346.496	354.285	523.928	268.076	552.449	226.170	195.326	564.258	118.191	231.213	12.954	562.113	131.416
1.8	327.204	303.907	331.386	512.428	287.784	448.423	172.705	153.735	503.158	91.237	185.625	3.220	468.777	106.338
2.0	278.327	259.700	273.459	498.655	316.281	387.770	131.569	120.922	445.094	74.735	148.555	2.363	425.708	83.566
2.2	240.010	217.920	201.631	501.678	273.983	354.799	102.537	92.427	392.062	65.951	119.530	2.302	394.099	66.547
2.4	208.123	182.329	145.737	504.217	179.415	320.577	80.071	68.303	342.182	57.439	95.600	1.791	328.505	55.155
2.6	178.169	152.234	104.755	477.533	113.182	271.975	60.313	49.328	292.700	43.985	74.680	2.658	241.775	46.682
2.8	148.522	126.137	74.361	418.764	86.092	215.548	43.171	35.369	242.515	31.874	56.580	5.121	171.764	38.417
3.0	120.225	103.648	51.778	353.098	67.969	161.953	29.660	25.241	193.105	23.546	41.916	2.610	128.244	29.743
3.2	94.740	84.745	37.325	303.527	56.792	116.203	19.822	17.518	147.034	18.177	30.781	0.445	101.255	21.768
3.4	73.025	69.285	33.939	278.252	49.075	79.482	13.012	11.476	107.045	15.081	22.702	0.134	80.957	15.505
3.6	54.354	55.822	35.535	271.510	39.564	50.492	8.295	6.879	73.563	13.481	16.639	0.022	62.240	10.912
3.8	39.104	44.157	37.759	278.353	29.189	29.503	5.167	3.736	47.956	13.076	12.073	0.002	45.770	7.751
4.0	27.372	34.354	36.561	294.019	19.904	15.976	3.181	1.833	30.052	14.004	8.676	0.000	32.891	5.613
N	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000

**Table 7.D.3: Quadrature Weights for Combined March and May 2003 10<sup>th</sup> Graders by Subgroup - ELA**

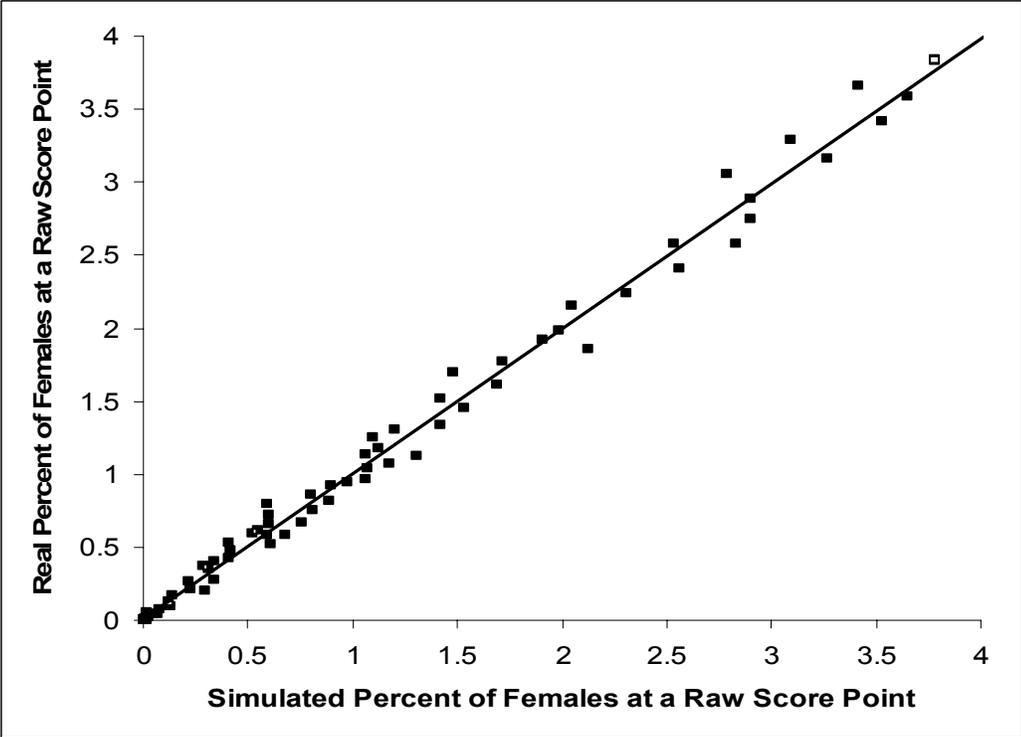
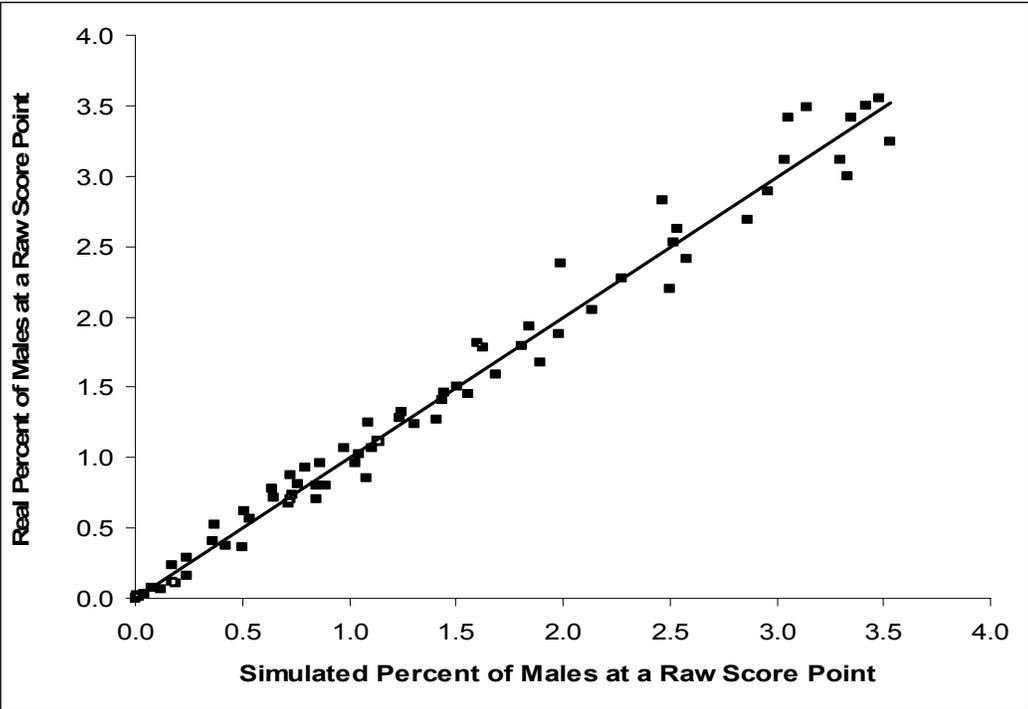
Quad	Female	Male	Am Ind	Asian	Pac Isl	Filipino	Hisp	Black	White	Disable	NSLP	Lower 20%	Upper 20%	Eng-Lmnr
-4.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-3.8	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.047	0.000	0.000
-3.6	0.000	0.009	0.000	0.000	0.000	0.000	0.000	0.000	0.010	0.009	0.018	0.248	0.000	0.000
-3.4	0.009	0.019	0.000	0.000	0.000	0.000	0.000	0.019	0.029	0.095	0.036	0.683	0.010	0.000
-3.2	0.019	0.049	0.000	0.000	0.000	0.000	0.011	0.047	0.048	0.417	0.073	0.877	0.029	0.000
-3.0	0.029	0.083	0.000	0.000	0.000	0.019	0.035	0.123	0.039	0.976	0.100	0.842	0.058	0.028
-2.8	0.040	0.111	0.000	0.000	0.000	0.183	0.076	0.240	0.019	1.327	0.119	0.841	0.087	0.093
-2.6	0.044	0.118	0.000	0.000	0.000	0.395	0.125	0.433	0.019	1.191	0.127	0.659	0.107	0.167
-2.4	0.059	0.155	0.000	0.000	0.000	0.193	0.191	0.405	0.019	1.285	0.186	0.634	0.117	0.216
-2.2	0.092	0.256	0.000	0.020	0.000	0.029	0.320	0.406	0.039	1.458	0.337	0.907	0.136	0.522
-2.0	0.243	0.655	0.000	0.186	0.000	0.010	0.753	0.791	0.126	1.640	0.838	2.032	0.214	0.975
-1.8	0.935	2.416	0.000	0.963	0.000	0.029	2.760	2.698	0.534	4.976	3.076	7.851	0.570	3.032
-1.6	3.878	9.876	0.370	3.565	0.683	0.742	11.396	11.181	2.272	23.508	12.630	35.418	1.924	15.389
-1.4	13.822	35.135	7.932	10.395	14.414	6.692	40.965	40.201	8.066	107.328	44.816	138.601	6.311	67.359
-1.2	38.337	94.539	76.498	29.756	58.733	18.921	113.380	109.389	21.164	359.027	121.119	442.746	16.682	212.034
-1.0	79.065	176.693	180.443	70.207	87.682	40.401	218.670	202.331	40.593	700.434	233.031	815.495	33.410	429.835
-0.8	123.922	236.467	182.362	114.344	151.311	67.230	304.392	262.538	59.988	812.348	329.643	855.897	51.919	589.743
-0.6	164.177	263.280	184.499	141.457	202.164	88.681	359.100	293.729	73.746	731.609	387.423	830.912	67.839	668.050
-0.4	203.100	287.441	227.315	167.659	231.257	106.082	408.233	333.349	86.635	676.005	429.648	851.668	81.219	732.166
-0.2	244.972	330.118	286.872	198.898	343.287	136.023	465.868	396.274	108.852	707.003	483.794	873.559	97.809	806.873
0.0	293.670	385.943	319.126	239.029	410.709	189.841	533.622	477.812	142.787	739.737	555.823	870.556	124.056	885.200
0.2	346.976	439.580	343.420	292.462	440.538	245.600	600.001	557.461	183.365	695.848	625.816	829.239	157.737	934.822
0.4	400.554	484.821	459.257	342.268	522.314	288.203	651.412	624.673	227.956	638.400	674.539	732.502	198.423	912.626
0.6	456.925	528.569	584.598	378.634	639.230	359.014	689.450	686.535	283.715	598.842	705.341	595.535	254.258	837.773
0.8	516.856	578.529	685.964	415.335	730.555	459.589	722.541	732.461	360.381	560.855	731.388	516.166	326.243	751.335
1.0	582.017	631.365	705.222	474.076	720.831	571.853	752.295	752.851	458.907	501.592	751.860	482.335	414.709	644.122
1.2	653.201	678.948	674.276	546.051	679.101	698.743	763.921	762.929	577.208	430.341	747.580	376.814	522.959	510.843
1.4	721.062	719.211	732.606	610.830	692.218	809.957	740.024	765.973	711.529	381.381	713.038	265.924	650.014	380.928
1.6	774.872	749.050	888.598	679.462	747.320	904.359	682.393	742.290	852.216	340.703	651.832	192.977	790.379	263.001
1.8	806.015	755.273	965.574	740.155	833.878	982.621	599.031	671.983	978.094	283.058	561.916	135.131	931.154	161.248
2.0	802.506	716.543	821.225	781.980	836.137	1002.492	488.380	554.169	1051.702	216.500	448.215	72.608	1039.280	89.665
2.2	752.449	624.558	545.813	807.837	678.965	940.836	358.701	412.162	1036.801	161.523	326.804	35.204	1067.985	48.594
2.4	652.851	492.217	388.707	796.147	431.498	791.355	233.510	276.524	920.575	118.347	214.796	16.632	989.505	26.068
2.6	516.049	345.620	320.182	717.053	244.876	576.804	134.247	166.596	724.491	82.477	125.859	7.462	815.798	13.468
2.8	367.346	214.548	243.160	574.373	137.481	355.575	68.928	88.984	500.316	52.184	65.864	6.605	592.907	6.992
3.0	234.097	118.306	121.347	402.133	70.397	189.770	32.379	42.710	302.961	31.560	31.091	2.919	378.565	3.808
3.2	133.211	58.437	41.444	243.222	41.162	92.696	14.150	18.779	161.404	19.219	13.335	0.453	212.421	1.792
3.4	67.918	25.995	10.797	127.956	23.678	43.741	5.735	7.413	76.080	10.575	5.230	0.442	105.183	0.806
3.6	30.889	10.274	2.089	59.373	13.743	20.360	2.099	2.555	31.535	4.501	1.863	0.372	45.711	0.315
3.8	12.792	3.625	0.285	24.818	8.980	8.376	0.690	0.785	11.718	1.393	0.601	0.155	17.807	0.093
4.0	5.006	1.167	0.028	9.353	6.818	2.586	0.210	0.193	4.051	0.332	0.185	0.031	6.464	0.019
N	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000	10000

**Table 7.D.4: Predicted Passing Rates on the Standard Setting Forms Versus Actual Passing Rates (Actual Rates from March and May 2003)**

Demographic Category	Mathematics		English Language Arts	
	Predicted	Actual	Predicted	Actual
Total Group	59%	60%	79%	79%
Males	59%	60%	74%	74%
Females	59%	59%	82%	83%
African American	37%	37%	69%	69%
Asian	83%	83%	85%	86%
Hispanic	42%	42%	65%	66%
White	76%	77%	91%	91%
Low 20% Schools	9%	N/A	31%	N/A
High 20% Schools	78%	N/A	92%	N/A
NSLP	42%	43%	63%	64%
English Learners	30%	30%	41%	43%
Special Education	24%	22%	41%	39%

Note: Actual passing rates for the lowest and highest 20% of the schools were not available.

**Figure 7.D.1: Simulated versus Actual Percentages of Males and Females at Various Raw Score Points—English Language Arts MC Items**



## Appendix 7.E: Technical Note on Converting Panelist Bookmarks to Passing scores on the CAHSEE

The purpose of this document is to explain the conversion of the panelist bookmarks on the CAHSEE to the passing scores presented to the California State Board of Education. The Bookmark methodology uses a response probability (RP) of 0.67, which will be explained in this document. Below the terms that are essential for understanding the process are defined, and then the technical details behind the standard are described, specifically when and where conversions and calculations using RP67 occurred.

### Definition of Terms

**Bookmark:** This term describes the point that the panelists believe separates sufficient performance from insufficient performance. When we use this term in this text, we intend for the reader to think only in terms of the panelist placement *before* any statistical conversion has occurred. The Bookmark placement will be referenced in terms of the item number or the page number indicated.

**Passing score:** This term describes the point that the panelists indicated separates sufficient performance from insufficient performance. However, it differs from the Bookmark in that it references the point *after* the statistical conversion has occurred. When discussing passing scores, we no longer talk about page or item numbers, but about b-values, scale scores, or raw scores.

**Raw Score:** The raw score is the number of points a student received on the test. It does not take into consideration which items a student answered correctly, or how difficult those items are, but only the number answered correctly or the points received on a writing task.

**Standard Error of Judgment (SEJ):** The SEJ is a measure of variance in the panelists' bookmark placements. It is calculated as the standard error of the mean (the standard deviation of the ratings divided by the square root of the number of panelists) with a minor adjustment for the correlation of judgments.

**IRT model:** An item response theory (IRT) model describes examinee performance with one number (scale score) and the properties of the items administered with other numbers (item parameters). There are several types of IRT models, but for the CAHSEE we speak only of the Rasch, or one-parameter model. The IRT model used on the CAHSEE focuses on difficulty (b-value) as the property that differentiates items. We use the IRT model to relate examinee performance to items and test scores.

**b-value:** This is a psychometric term used in the IRT model to describe the difficulty of an item. Each item has its own b-value that is based on how students performed on that item relative to other items. It is also called the "b-parameter" or the "difficulty parameter." When an examinee's scale score equals the item's b-value, that examinee has a 50% probability of correctly answering that item.

**Scale Score:** Scale scores are 3-digit numbers that reflect an examinee's test performance. Scale scores are nonlinear conversions of raw scores that take test difficulty into account. Scale scores are comparable from one test form to another. The scale scores on the CAHSEE range from 250–350. The current passing score is set at 350.

**Standard Error of Measurement (SEM):** The SEM is the standard deviation of an individual's observed score from repeated administrations of a test under identical conditions. Because such data generally cannot be collected, it is usually estimated from group data. The final SEM represents the difference between an observed score and the corresponding true score.

**Response Probability (RP):** The response probability (RP) is defined in standard setting as the degree of mastery required of students. The Bookmark procedure uses an RP of 0.67, also referred to as RP67. An item's RP67 value equals the scale score at which a student has a 0.67, or two-thirds, probability of answering the item correctly (for MC items) or of reaching at least that score point (for writing tasks).

### **Overview of the Process**

Several factors should be considered in any discussion of converting bookmarks to passing scores. First, one needs to understand how the test materials given to panelists were produced. Secondly, one needs to know what instructions were given to panelists about placing their bookmark. And, thirdly, one needs to know how these bookmarks were used.

Panelists were given an ordered item booklet containing all the information about the item needed to complete the bookmark task. The items were ordered by their difficulty parameter. For the MC items, each page of the ordered item booklet showed the item, along with any passage or graphic, and the possible responses. For the writing task, the initial prompt, the scoring rubric, and a sample student response for each possible score was provided.

During the standard setting meeting, panelists were told to place their bookmarks in the ordered-item booklet on the first page after the “hardest item for which two-thirds of those students with sufficient knowledge and skills to just pass the CAHSEE would be able to answer correctly.” They recorded the *page* number of their bookmark, and an ETS staff member entered that number. For the mathematics test, this page number matched the item number. The ELA ordered item booklet contained passages that covered more than one page and also included rubrics and sample student responses for the writing task, so we referenced the page number back to an item number before providing any feedback to the panelists. During the feedback stages of rounds 2 and 3, ETS gave panelists feedback about the group’s median, high, and low bookmarks using item number terminology.

However, it is important to realize that even though the feedback was given in terms of item numbers, at other stages of the standard setting process we converted the item numbers to a raw score using the IRT model that included a response probability (RP) value of 0.67. That is, ETS calculated the point on the scale at which students would have a 67% probability of answering correctly the item indicated by the bookmark. The RP67 conversions were done at three stages: in creating the ordered item booklet, in providing consequence feedback to the panelist; and in calculating the median bookmark recommended by the panelists. The use of RP67 at these three stages maintained necessary consistency between the instructions given to panelists, the results of the panelists’ deliberations, and the raw score cut points on the test.

### **Creating the Ordered-Item Booklet**

When using the Rasch model, the RP value has no effect on the ordering of the MC items. The RP67 value is determined by adding a constant to each item’s b-value. For writing tasks ETS determined the b-value for each CR score point at which a student at a given scale score would have a two-thirds probability of receiving that number of score points or more. Then difficulty values for both the MC items and the CR points were available on the same RP67 scale. These values were sorted from easiest (lowest RP67 value) to hardest (highest) to create the ordered item booklet.

### **Providing Consequence Data**

At the beginning of Round 3, ETS facilitators told each of the four panels what the median bookmark of the panel was as well as the highest and lowest bookmarks set by individual panelists. They then gave them consequence data for the median. That is, they told them what percentage of students would be expected to pass that section of the CAHSEE given the panel’s current median bookmark. To

calculate the predicted passing rate ETS used the census data from 10<sup>th</sup>-graders taking the test for the first time during the March and May 2003 administration of the CAHSEE. For the mathematics test, ETS found the point on the scale that corresponded to the score at which a student had a two-thirds probability of answering correctly the item indicated by the panel's median bookmark. ELA was a little more complicated as we first had to apply the appropriate weighting to the writing task to bring the 76-item ordered-item booklet back to the full 90-point test. Each point value on the writing task were multiplied by 4.5, and then ETS followed the same procedures as in mathematics to calculate the IRT passing score. The appropriate point on each of the scales (mathematics and ELA) was found and then a percentage distribution table of student scores was used to determine the percentage of students likely to score at or above the median passing score.

One difference between the calculation made during the workshop and the calculations done for the preliminary report was in the denominator used to determine the percentage of students. During the workshop, the facilitators presented the number of students scoring at or above the passing score as a percentage of all students enrolled at the time the test was administered. This calculation basically assumes that all students who were enrolled but who did not take the test would have scored a zero. For purposes of creating the tables in this report, ETS calculated the number of students scoring at or above the passing score as a percentage of all students who took the test. This method assumes that the distribution of scores of the students who were enrolled but did not take the test would have been the same as the distribution of scores as the test takers.

#### **Calculating the Final Passing Score**

Again, in calculating the final median passing score that appeared in the preliminary report, ETS converted the median bookmark to the passing score using the IRT model that included RP67. This was achieved by first locating the item number associated with the median bookmark for the panel. Then, using the IRT model, ETS determined the scale score a student would need to reach to have a two-thirds probability of correctly answering the bookmarked item. ETS then converted that scale score back to a raw score. So the steps are as follows:

1. Locate the item indicated by the median bookmark.
2. Calculate the scale score associated with having a two-thirds probability of successfully answering that item (i.e., the RP67 value).
3. Convert that scale score back to a raw score.

#### **Examples**

It is easier to begin with an example for mathematics, since the calculation is more straightforward for this test. The reader can follow along with this example using Table 7.E.1. The median bookmark of the educators' panel was at item 44. That is, the panelists recommended that to pass the CAHSEE a student needed a two-third probability of correctly answering the 44<sup>th</sup> easiest item. The scale score corresponding to RP67 for this item was 358. The raw score corresponding to that scale score was 53. ETS then calculated the number of students who took the CAHSEE in March or May 2003 who would be predicted to obtain a raw score of 53 or higher and divided that number by the number of test takers. This value came to 52%. (If we had used the enrollment figure as the denominator, the value would have been 48%.)

**Table 7.E.1: Sample Mathematics Conversion Table**

Median bookmark	Bookmark	Scale score at RP67	Raw score out of 80
	Item 34	352	50
	...	...	...
→	Item 44	358	53
	...	...	...
	Item 50	365	56

For ELA (Table 7.E.2), the median bookmark of the educators' panel was at item 37 out of the 76 items that were in the ordered item booklet (72 MC items plus score categories of one to four for the CR item). ETS then converted the number 37 to a scale score a student would need to reach to have a two-thirds probability of correctly answering that item. Finally, ETS converted that scale score back to a raw score with RP67 on the full 90-point test (where the CR item was weighted 18 points and each MC item was weighted one point). This raw score was 58. ETS then calculated the percentage of students who would be predicted to achieve a raw score of 58 or higher as described above. Using test-takers as the denominator, 71% of first-time test takers were predicted to pass as 10<sup>th</sup>-graders, while using enrollment as the denominator resulted in the lower pass rate of 64%.

**Table 7.E.2: Sample ELA Conversion Table**

Median bookmark	Bookmark	Scale score at RP67	Raw score out of 90
	Item 18	346	50
	...	...	...
→	Item 37	362	58
	...	...	...
	Item 62	384	68

## References

- Hambleton, R. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G.J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives*, (pp. 89-116). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.
- Mislevy, R. (1984). Estimating latent distributions. *Psychometrika*, 49, 359-381.
- Muraki, E. (2000). *RESGEN: Item Response Generator* (Version 4.0). Unpublished software program.
- Muraki, E., & Bock, R. D. (1995). *PARSCALE: Parameter scaling of rating data* (Version 2.2). Chicago, IL: Scientific Software, Inc.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7, 201-210.