California Department of Education
Statewide Assessment Division

California Modified Assessment
Technical Report
Spring 2009 Administration

Submitted February 16, 2010

Educational Testing Service

Contract No. 5417

Table of Contents

Acronyms and Initialisms Used in the *CMA Technical Report* viii

Chapter 1: Introduction **1**
Background **1**
Test Purpose **1**
Test Content **2**
Intended Population **2**
Intended Use and Purpose of Test Scores **2**
Testing Window **3**
Significant Developments in 2009 ..... **3**
Elimination of California Achievement Tests, Sixth Edition Survey (CAT/6 Survey) 3
New Cut Scores and Score Scales 3
Testing in Grades Six through Eight 4
Limitations of the Assessment **4**
Score Interpretation 4
Out-of-Level Testing 4
Score Comparisons 4
Groups and Organizations involved with the STAR Program **4**
State Board of Education (SBE) 4
California Department of Education (CDE) 4
Contractors 5
Overview of the Technical Report **5**
Reference **7**

Chapter 2: An Overview of CMA Processes **8**
Item Development **8**
Item Formats 8
Item Development Specifications 8
Item Banking 8
Item Refresh Rate 8
Test Assembly **9**
Test Length 9
Test Blueprint 9
Content Rules and Item Selection 9
Psychometric Criteria 9
Test Administration **10**
Test Security and Confidentiality 10
Procedures to Maintain Standardization 11
Test Variations and Accommodations **11**
Accommodation Summaries 12
Scores **12**
Aggregation Procedures **12**
Individual Scores 13
Group Scores 13
Equating **14**
Calibration 14
Scaling 14
Equating the Braille Versions of the CMA 15
References **16**
Appendix 2.A—CMA Item and Estimated Time Chart **17**
Appendix 2.B—Reporting Clusters **18**
English–Language Arts 18
Mathematics 18
Science 18
Appendix 2.C—2009 Test Variations and Accommodations **19**
Appendix 2.D—Accommodation Summary Tables **21**

Chapter 3: Item Development **60**
Rules for Item Development **60**
Item Development Specifications 60
Expected Item Ratio 61
Selection of Item Writers **62**
Criteria for Selecting Item Writers 62
Item Writer Training 62

Item Review Process **63**
Contractor Review 63
Content Expert Reviews 64
Statewide Pupil Assessment Review (SPAR) Panel 67
Field Testing **67**
Stand-Alone Field Testing 67
Embedded Field-test Items 68
Data Review Meetings **69**
Item Banking **69**
References **70**
Chapter 4: Test Assembly **71**
Test Length **71**
Rules for Item Selection **71**
Test Blueprint 71
Content Rules and Item Selection 71
Psychometric Criteria 72
Rules for Item Sequence and Layout **73**
Chapter 5: Test Administration **74**
Test Security and Confidentiality **74**
ETS's Office of Testing Integrity (OTI) 74
Test Development 74
Item Review 74
Item Bank 75
Transfer of Forms and Items to the CDE 75
Firewall 75
Printing 76
Test Administration 76
Test Delivery 76
Processing and Scoring 77
Data Management 77
Transfer of Scores via Secure Data Exchange 78
Statistical Analysis 78
Reporting and Posting Results 78
Student Confidentiality 78
Test Results 78
Procedures to Maintain Standardization **79**
Test Administrators 79
Directions for Administration (DFA) 80
District and Test Site Coordinator Manual 80
STAR Management System Manuals 80
Test Booklets 81
Accommodations for Students with Disabilities **81**
Identification 81
Scoring 82
Demographic Data Corrections **82**
Testing Irregularities **82**
Test Administration Incidents **82**
References **83**
Chapter 6: Performance Standards **84**
Background **84**
Standard Setting Procedure **84**
Development of Competencies Lists 85
Standard Setting Methodology **86**
Results **86**
References **88**
Chapter 7: Scoring and Reporting **89**
Procedures for Maintaining and Retrieving Individual Scores **89**
Scoring and Reporting Specifications 89
Scanning and Scoring 90
Types of Scores and Subscores **90**
Raw Score 90
Subscore 90
Scale Score 90
Performance Levels 90

Writing Score 91
Score Verification Procedures 91
Scoring Key Verification Process 91
Monitoring and Quality Control of Writing Scoring 92
Score Verification Process 94
Overview of Score Aggregation Procedures 94
Individual Scores 94
Reports to be Produced and Scores for each Report 97
Types of Score Reports 97
Score Report Contents 97
Score Report Applications 98
Criteria for Interpreting Test Scores 98
Criteria for Interpreting Score Reports 98
Reference 100
Appendix 7.A—ELA for Writing (Grades Four and Seven) Rubrics 101
Grade Four Scoring Rubric 101
Grade Seven Scoring Rubric 104
Appendix 7.B—Scale Score Distributions 107
Appendix 7.C—Demographic Summaries 110
Appendix 7.D—Types of Score Reports 124
Chapter 8: Analyses 127
Samples Used for the Analyses 127
Classical Analyses 128
Multiple-Choice Items 128
Writing Tasks 129
Reliability Analyses 129
Intercorrelations, Reliabilities, and SEMs for Reporting Clusters 131
Subgroup Reliabilities and SEMs 131
Conditional Standard Errors of Measurement 131
Decision Classification Analyses 132
Writing Score Reliability 133
***Prompt and Rater Agreement Summary* 134**
Validity Evidence 135
Purpose of the CMA 136
The Constructs to Be Measured 136
Intended Test Population(s) 137
Validity Evidence Collected 137
Evidence Based on Response Processes 140
Evidence Based on Internal Structure 140
Evidence Based on Consequences of Testing 142
IRT Analyses 142
IRT Model-Data Fit Analyses 142
Model Fit Assessment Results 144
Summaries of Scaled IRT b-values 144
Scaling Results 145
DIF Analyses 145
References 148
Appendix 8.A—Classical Analyses 150
Appendix 8.B—Reliability Analyses 157
Appendix 8.C—Validity Analyses 182
Appendix 8.D—IRT Analyses 190
Appendix 8.E—DIF Analyses 204
Chapter 9: Quality Control 218
Quality Control of Item Development 218
Item Specifications 218
Item Writers 218
Internal Contractor Reviews 218
Assessment Review Panel (ARP) Review 219
Statewide Pupil Assessment Review (SPAR) Panel Review 219
Data Review of Field Tested Items 219
Quality Control of the Item Bank 220
Quality Control of Test Materials 220
Collecting Test Materials 220
Processing Test Materials 221

Quality Control of Scanning **221**
Post-scanning Edits 221
Quality Control of Image Editing **222**
Quality Control of Answer Document Processing and Scoring **222**
Accountability of Answer Documents 222
Processing of Answer Documents 222
Scoring and Reporting Specifications 223
Matching Information on CMA Answer Documents 223
Matching Multiple-choice and Writing Scores for ELA Grades Four and Seven 223
Storing Answer Documents 223
Quality Control of Psychometric Processes **224**
Score Key Verification Procedures 224
Quality Control of Item Analyses, and DIF 224
Score Verification Process 225
Offloads to Test Development 225
Quality Control of Reporting **225**
Excluding Student Scores from Summary Reports 226

Tables

Table 2.1 Subgroup Definitions 13
Table 2.C.1 List of Test Variations and Accommodations 19
Table 2.D.1 Accommodation Summary for ELA, Grade Three 21
Table 2.D.2 Accommodation Summary for ELA, Grade Four 24
Table 2.D.3 Accommodation Summary for ELA, Grade Five 27
Table 2.D.4 Accommodation Summary for ELA, Grade Six 30
Table 2.D.5 Accommodation Summary for ELA, Grade Seven 33
Table 2.D.6 Accommodation Summary for ELA, Grade Eight 36
Table 2.D.7 Accommodation Summary for Mathematics, Grade Three 39
Table 2.D.8 Accommodation Summary for Mathematics, Grade Four 42
Table 2.D.9 Accommodation Summary for Mathematics, Grade Five 45
Table 2.D.10 Accommodation Summary for Mathematics, Grade Six 48
Table 2.D.11 Accommodation Summary for Mathematics, Grade Seven 51
Table 2.D.12 Accommodation Summary for Science, Grade Five 54
Table 2.D.13 Accommodation Summary for Science, Grade Eight 57
Table 3.1 Field-test Percentages for the CMA 62
Table 3.2 CMA ARP Member Qualifications, by Subject and Total 65
Table 3.3 Summary of Items and Forms Presented in 2009 CMA 68
Table 4.1 Target Statistical Specifications for the CMA 73
Table 6.1 Scale Scores Ranges for Performance Levels 87
Table 6.2 Percentage of Examinees in Each Performance Level 87
Table 7.1 Mean and Standard Deviation of Raw and Scale Scores for CMA 95
Table 7.2 Percentage of Examinees in Proficiency Levels for CMA 95
Table 7.3 Subgroup Definitions 96
Table 7.4 Types of CMA Reports 97
Table 7.B.1 Distribution of CMA Scale Scores for English–Language Arts 107
Table 7.B.2 Distribution of CMA Scale Scores for Mathematics 107
Table 7.B.3 Distribution of CMA Scale Scores for Science 108
Table 7.B.4 Distribution of CMA Raw Scores for English–Language Arts Grades Six, Seven, and Eight 108
Table 7.B.5 Distribution of CMA Raw Scores for Mathematics Grades Six and Seven 109
Table 7.B.6 Distribution of CMA Raw Scores for Science Grade Eight 109
Table 7.C.1 Demographic Summary for ELA, Grade Three 110
Table 7.C.2 Demographic Summary for ELA, Grade Four 112
Table 7.C.3 Demographic Summary for ELA, Grade Five 114
Table 7.C.4 Demographic Summary for Mathematics, Grade Three 116
Table 7.C.5 Demographic Summary for Mathematics, Grade Four 118
Table 7.C.6 Demographic Summary for Mathematics, Grade Five 120
Table 7.C.7 Demographic Summary for Science, Grade Five 122
Table 7.D.1 Score Reports Reflecting CMA Results 124
Table 8.1 Summary Statistics for P1 and Calibration Samples 128
Table 8.2 Average and Median Proportion Correct and Point-Biserial 128
Table 8.3 Reliabilities and SEMs for the CMA 130
Table 8.4 Scale Score CSEM at Cut Points 132
Table 8.5 2009 CMA Content Area Correlations (All Valid Scores) 139
Table 8.6 Subgroup Classification for DIF Analyses 146
Table 8.A.1 Item-by-item *p*-value and Point-Biserial for ELA 150
Table 8.A.2 Item-by-item *p*-value and Point-Biserial for Mathematics 151

Table 8.A.3 Item-by-item *p*-value and Point-Biserial for Science 152
Table 8.A.4 Distribution of Essay Scores for ELA Grade Four—Overall and by Subgroup (all %) 153
Table 8.A.5 Distribution of Essay Scores for ELA Grade Seven—Overall and by Subgroup (all %) 154
Table 8.A.6 Mean Scores for ELA Grade Four Essay—Overall and by Subgroup 155
Table 8.A.7 Mean Scores for ELA Grade Seven Essay—Overall and by Subgroup 155
Table 8.A.8 Effect Sizes for ELA Grades Four and Seven Essay—by Subgroup 156
Table 8.B.1 Subscore Reliabilities and Correlations for ELA 157
Table 8.B.2 Subscore Reliabilities and Correlations for Mathematics 157
Table 8.B.3 Subscore Reliabilities and Correlations for Science 157
Table 8.B.4 Reliabilities and SEMs for the CMA by Gender 158
Table 8.B.5 Reliabilities and SEMs for the CMA by Economic Status 158
Table 8.B.6 Reliabilities and SEMs for the CMA by Language Fluency 159
Table 8.B.7 Reliabilities and SEMs for the CMA by Primary Ethnicity 160
Table 8.B.8 Reliabilities and SEMs for the CMA by Primary Ethnicity-for-Not-Economically-Disadvantaged 161
Table 8.B.9 Reliabilities and SEMs for the CMA by Primary Ethnicity-for-Economically-Disadvantaged 162
Table 8.B.10 Reliabilities and SEMs for the CMA by Gender by Economic Status 163
Table 8.B.11 Reliabilities and SEMs for the CMA by Primary Disability 164
Table 8.B.12 Reliabilities and SEMs for the CMA by Primary Disability (continued) 165
Table 8.B.13 Overall Subgroup Reliabilities 166
Table 8.B.14 Overall Subgroup Reliabilities—Primary Ethnicity 166
Table 8.B.15 Overall Subgroup Reliabilities by Primary Ethnicity—Not Economically Disadvantaged 166
Table 8.B.16 Overall Subgroup Reliabilities by Primary Ethnicity—Economically Disadvantaged 167
Table 8.B.17 Overall Subgroup Reliabilities by Gender/Economic Status 167
Table 8.B.18 Overall Subgroup Reliabilities by Primary Disability 167
Table 8.B.19 Overall Subgroup Reliabilities by Primary Disability (continued) 168
Table 8.B.20 Subscore Reliabilities and SEM by Gender 168
Table 8.B.21 Subscore Reliabilities and SEM by Gender—Not Economically Disadvantaged 169
Table 8.B.22 Subscore Reliabilities and SEM by Gender—Economically Disadvantaged 170
Table 8.B.23 Subscore Reliabilities and SEM by Language Fluency 171
Table 8.B.24 Subscore Reliabilities and SEM by Ethnicity 172
Table 8.B.25 Subscore Reliabilities and SEM by Ethnicity—Not Economically Disadvantaged 173
Table 8.B.26 Subscore Reliabilities and SEM by Ethnicity—Economically Disadvantaged 174
Table 8.B.27 Subscore Reliabilities and SEM by Economic Status 175
Table 8.B.28 Subscore Reliabilities and SEM by Disability 176
Table 8.B.29 Subscore Reliabilities and SEM by Disability (continued) 177
Table 8.B.30 Reliability of Classification for ELA, Grade Three 178
Table 8.B.31 Reliability of Classification for ELA, Grade Four (Reading only) 178
Table 8.B.32 Reliability of Classification for ELA, Grade Four (Reading & Writing) 178
Table 8.B.33 Reliability of Classification for ELA, Grade Five 179
Table 8.B.34 Reliability of Classification for Mathematics, Grade Three 179
Table 8.B.35 Reliability of Classification for Mathematics, Grade Four 179
Table 8.B.36 Reliability of Classification for Mathematics, Grade Five 180
Table 8.B.37 Reliability of Classification for Science, Grade Five 180
Table 8.B.38 Inter-Rater Analyses for ELA, Grade Four 180
Table 8.B.39 Inter-Rater Analyses for ELA, Grade Seven 181
Table 8.B.40 Descriptive Statistics for the Ratings by the Two Raters 181
Table 8.B.41 Generalizability Analyses for Grade Four Essay—[(Person: Essay) x Rater)] 181
Table 8.B.42 Generalizability Analyses for Grade Seven Essay—[(Person: Essay) x Rater] 181
Table 8.C.1 2009 CMA Content Area Correlations (All Valid Scores) 182
Table 8.C.2 2009 CMA Content Area Correlations (Gender) 182
Table 8.C.3 2009 CMA Content Area Correlations (Primary Ethnicity) 183
Table 8.C.4 2009 CMA Content Area Correlations (English Fluency) 184
Table 8.C.5 2009 CMA Content Area Correlations (Economic Status) 184
Table 8.C.6 2009 CMA Content Area Correlations (Primary Disability) 185
Table 8.C.7 2009 CMA Content Area Correlations (Raw Score: All Valid Scores) 185
Table 8.C.8 2009 CMA Content Area Correlations (Raw Score: Gender) 186
Table 8.C.9 2009 CMA Content Area Correlations (Raw Score: Primary Ethnicity) 186
Table 8.C.10 2009 CMA Content Area Correlations (Raw Score: English Fluency) 187
Table 8.C.11 2009 CMA Content Area Correlations (Raw Score: Economic Status) 188
Table 8.C.12 2009 CMA Content Area Correlations (Raw Score: Primary Disability) 189
Table 8.D.1 IRT Model Data Fit Distribution for ELA, Grades Three through Eight 190
Table 8.D.2 IRT Model Data Fit Distribution for Mathematics, Grades Three through Seven 190
Table 8.D.3 IRT Model Data Fit Distribution for Science, Grades Five and Eight 190
Table 8.D.4 IRT Model Data Fit Distribution for ELA, Grades Three through Eight (field-test items) 190
Table 8.D.5 IRT Model Data Fit Distribution for Mathematics, Grades Three through Seven (field-test items) 190
Table 8.D.6 IRT Model Data Fit Distribution for Science, Grades Five and Eight (field-test items) 191

Table 8.D.7 IRT *b*-values for ELA, Grade Three 191
Table 8.D.8 IRT *b*-values for ELA, Grade Four 191
Table 8.D.9 IRT *b*-values for ELA, Grade Five 191
Table 8.D.10 IRT *b*-values for Mathematics, Grade Three 191
Table 8.D.11 IRT *b*-values for Mathematics, Grade Four 192
Table 8.D.12 IRT *b*-values for Mathematics, Grade Five 192
Table 8.D.13 IRT *b*-values for Science, Grade Five 192
Table 8.D.14 Distribution of IRT *b*-values for ELA 192
Table 8.D.15 Distribution of IRT *b*-values for ELA (field-test items) 193
Table 8.D.16 Distribution of IRT *b*-values for Mathematics 193
Table 8.D.17 Distribution of IRT *b*-values for Mathematics (field–test items) 194
Table 8.D.18 Distribution of IRT *b*-values for Science 194
Table 8.D.19 Distribution of IRT *b*-values for Science (field-test items) 195
Table 8.D.20 2009 New Conversions for ELA, Grade Three—Regular Standard Form 195
Table 8.D.21 2009 New Conversions for ELA, Grade 3—Braille Version 196
Table 8.D.22 2009 New Conversions for ELA, Grade Four—Multiple Choice Only 197
Table 8.D.23 2009 New Conversions for ELA, Grade Four—Composite Version (multiple choice and essay) 198
Table 8.D.24 2009 New Conversions for ELA, Grade Five 199
Table 8.D.25 2009 New Conversions for Mathematics, Grade Three 200
Table 8.D.26 2009 New Conversions for Mathematics, Grade Four 201
Table 8.D.27 2009 New Conversions for Mathematics, Grade Five 202
Table 8.D.28 2009 New Conversions for Science, Grade Five 203
Table 8.E.1 Operational Items Exhibiting Significant DIF 204
Table 8.E.2 Field Test Items Exhibiting Significant DIF 204
Table 8.E.3 DIF Classifications for ELA, Grade Three Operational Items 205
Table 8.E.4 DIF Classifications for ELA, Grade Four Operational Items 205
Table 8.E.5 DIF Classifications for ELA, Grade Five Operational Items 206
Table 8.E.6 DIF Classifications for ELA, Grade Six Operational Items 206
Table 8.E.7 DIF Classifications for ELA, Grade Seven Operational Items 207
Table 8.E.8 DIF Classifications for ELA, Grade Eight Operational Items 207
Table 8.E.9 DIF Classifications for Mathematics, Grade Three Operational Items 208
Table 8.E.10 DIF Classifications for Mathematics, Grade Four Operational Items 208
Table 8.E.11 DIF Classifications for Mathematics, Grade Five Operational Items 209
Table 8.E.12 DIF Classifications for Mathematics, Grade Six Operational Items 209
Table 8.E.13 DIF Classifications for Mathematics, Grade Seven Operational Items 210
Table 8.E.14 DIF Classifications for Science, Grade Five Operational Items 210
Table 8.E.15 DIF Classifications for Science, Grade Eight Operational Items 211
Table 8.E.16 DIF Classifications for ELA, Grade Three Field-test Items 211
Table 8.E.17 DIF Classifications for ELA, Grade Four Field-test Items 212
Table 8.E.18 DIF Classifications for ELA, Grade Five Field-test Items 212
Table 8.E.19 DIF Classifications for ELA, Grade Six Field-test Items 213
Table 8.E.20 DIF Classifications for ELA, Grade Seven Field-test Items 213
Table 8.E.21 DIF Classifications for ELA, Grade Eight Field-test Items 214
Table 8.E.22 DIF Classifications for Mathematics, Grade Three Field-test Items 214
Table 8.E.23 DIF Classifications for Mathematics, Grade Four Field-test Items 215
Table 8.E.24 DIF Classifications for Mathematics, Grade Five Field-test Items 215
Table 8.E.25 DIF Classifications for Mathematics, Grade Six Field-test Items 216
Table 8.E.26 DIF Classifications for Mathematics, Grade Seven Field-test Items 216
Table 8.E.27 DIF Classifications for Science, Grade Five Field-test Items 217
Table 8.E.28 DIF Classifications for Science, Grade Eight Field-test Items 217

Figures

Figure 3.1 Item and Test Development Cycle 60
Figure 8.1 Decision Accuracy for Achieving a Performance Level 133
Figure 8.2 Decision Consistency for Achieving a Performance Level 133
Figure 8.3 Items from the 2005 History–Social Science Grade Ten Field-Test Calibration 143

Acronyms and Initialisms Used in the *CMA Technical Report*

ADA	Americans with Disabilities Act
ARP	Assessment Review Panel
API	Academic Performance Index
AYP	adequate yearly progress
CAPA	California Alternate Performance Assessment
CAT/6 Survey	California Achievement Tests, Sixth Edition Survey
CDE	California Department of Education
CDS	county/district/school
CI	confidence interval
CMA	California Modified Assessment
CR	constructed response
CSTs	California Standards Tests
CSEMs	conditional standard errors of measurement
DFA	*Directions for Administration*
DIF	differential item functioning
DOK	depth of knowledge
DPLT	designated primary language test
DQS	Data Quality Services
d-study	decision study
EL	English learner
ELA	English–language arts
EM	expectation maximization
EOC	end-of-course
ePEN	Electronic Performance Evaluation Network
ESEA	Elementary and Secondary Education Act
ETS	Educational Testing Service
FIA	final item analysis
g-study	generalizability study
GENASYS	Generalized Analysis System
ICC	item characteristic curve
IEP	individualized education program
I-FEP	initially fluent English proficient
IRT	item response theory
LEA	local education agency
MC	multiple choice
MH DIF	Mantel-Haenszel DIF
NRT	norm-referenced test
NPS	nonpublic, nonsectarian school
NSLP	National School Lunch Program
OTI	Office of Testing Integrity
p-value	item proportion correct
Pt-Rbis	point-biserial correlations
RACF	Random Access Control Facility
R-FEP	reclassified fluent English proficient
SBE	State Board of Education
SD	standard deviation
SEMs	standard errors of measurement
SFTP	secure file transfer protocol
SGID	School and Grade Identification sheet
SKM	score key management
SPAR	Statewide Pupil Assessment Review
SR&T	Scoring, Reporting and Technology
STAR	Standardized Testing and Reporting
STAR TAC	STAR Technical Assistance Center
STS	Standards-based Tests in Spanish

Chapter 1: Introduction

Background

In 1997 and 1998, the California State Board of Education (SBE) adopted rigorous content standards in four major content areas: English–language arts (ELA), mathematics, history–social science, and science. These standards are designed to guide instruction and learning for all students in the state and to bring California students to world-class levels of achievement.

In order to measure and evaluate student achievement of the content standards, the state instituted the Standardized Testing and Reporting (STAR) Program. This Program, administered annually, was authorized in 1997 by state law (Senate Bill 376). Senate Bill 1448, approved by the Legislature and the Governor in August 2004, reauthorized the STAR Program through January 1, 2011, in grades three through eleven. STAR Program testing in grade two has also been extended to the 2011 school year (spring 2011 administration) after Senate Bill 80 was passed in September 2007.

During its 2009 administration, the STAR Program had four components:

- California Standards based Tests (CSTs), produced for California public schools
- California Modified Assessment (CMA), an assessment of students' achievement of California's content standards for ELA, mathematics, and science, developed for students with an IEP who meet the CMA eligibility criteria approved by the SBE (In 2009, the CMA was administered for ELA in grades three through eight, for mathematics in grades three through seven, and for science in grades five and eight.)
- California Alternate Performance Assessment (CAPA), produced for students with significant cognitive disabilities who have an IEP and are not able to take the CSTs or the CMA
- Standards-based Tests in Spanish (STS), an assessment of students' achievement of California's content standards for Spanish-speaking English learners that is administered as the STAR Program's designated primary language test (DPLT) (In 2009, the STS was administered for reading/language arts in grades two through eleven, for grade-level mathematics in grades two through seven, for end-of-course [EOC] Algebra I in grades seven through eleven, and for EOC Geometry in grades eight through eleven.)

Test Purpose

The CMA is a grade-level assessment for students who have an individualized education program (IEP), are receiving grade-level instruction, and, even with interventions, are not likely to achieve grade-level proficiency within the year covered by the student's IEP. The purpose of the CMA is to allow students to demonstrate their level of performance relative to the content standards in English-language art, mathematics, and science.

In 2009, the CMA tests were administered to students in ELA for grades three through eight and mathematics for grades three through seven and in science for grades five and eight. CMA results for grades three through five are used for in calculating school and district Academic Performance Index (API) and adequate yearly progress (AYP), which applies toward meeting the requirement of the federal Elementary and Secondary Education Act (ESEA) that all students score proficient or above by 2014.

In November 2008, the SBE approved performance levels for the CMA in grades three through five. CMA results for these grade levels are included in the API beginning with the 2008 Base API. As with CAPA results in API reporting, the performance level the student receives on the CMA (far below basic, below basic, basic, proficient, or advanced) was the level that was included in the API calculations.

In the spring of 2009, the CMA was expanded to include ELA in grades six through eight, mathematics in grades six and seven, and science in grade eight. The SBE is scheduled to adopt performance levels for these grade levels in November 2009. Accordingly, CMA results were not available in time for reporting the 2009 Growth API in September 2009.

Test Content

The CMA includes multiple-choice items for all grades and subjects tested and was used to assess skills in three content areas in 2009: ELA, mathematics, and science. Students in grades three through eight were tested in ELA, students in grades three through seven were tested in mathematics, and students in grades five and eight were tested in science. Students who took the CMA for ELA test in grades four and seven also took a writing test.

Intended Population

The CMA tests are designed for students with an IEP who cannot achieve grade level proficiency on the CSTs with or without accommodations and/or modifications and who meet eligibility criteria adopted by the SBE. The decision to administer the CMA is made by a special-education student's IEP team. In addition, to be eligible to take the CMA, the student must have scored at the below basic or far below basic performance level on a previously administered CST.

Parents may submit written requests to have their children exempted from taking any of the tests within the STAR Program. For grade four and grade seven ELA tests, parents can submit a written request to have their child exempted from taking the essay portion of the ELA tests. Only students with a written parent request may be exempted from taking the tests.

Intended Use and Purpose of Test Scores

The results for tests within the STAR Program are used for three primary purposes, described as follows (excerpted from California *Education Code* Section 60602, http://www.leginfo.ca.gov/cgi-bin/displaycode?section=edc&group=60001-61000&file=60600-60603):

"60602. (a) (1) First and foremost, provide information on the academic status and progress of individual pupils to those pupils, their parents, and their teachers. This information should be designed to assist in the improvement of teaching and learning in California public classrooms. The Legislature recognizes that, in addition to statewide assessments that will occur as specified in this chapter, school districts will conduct additional ongoing pupil diagnostic assessment and provide information regarding pupil performance based on those assessments on a regular basis to parents or guardians and schools. The legislature further recognizes that local diagnostic assessment is a primary mechanism through which academic strengths and weaknesses are identified."

"60602. (a) (4) Provide information to pupils, parents or guardians, teachers, schools, and school districts on a timely basis so that the information can be used to further the development of the pupil and to improve the educational program."

"60602. (c) It is the intent of the Legislature that parents, classroom teachers, other educators, governing board members of school districts, and the public be involved, in an active and ongoing basis, in the design and implementation of the statewide pupil assessment program and the development of assessment instruments."

"60602. (d) It is the intent of the Legislature, insofar as is practically feasible and following the completion of annual testing, that the content, test structure, and test items in the assessments that are part of the Standardized Testing and Reporting Program become open and transparent to teachers, parents, and pupils, to assist all the stakeholders in working together to demonstrate improvement in pupil academic achievement. A planned change in annual test content, format, or design, should be made available to educators and the public well before the beginning of the school year in which the change will be implemented."

In addition, STAR program assessments are used to provide data for state and federal accountability purposes.

Testing Window

The CMA tests are administered at different times of the year, depending on the progression of the school year within each particular school district. Specifically, schools must administer the CSTs, CMA, CAPA, and STS tests within a 21-day window, which begins 10 days before and ends 10 days after the day on which 85 percent of the instructional year is completed. Multiple-choice CMA tests may be administered over multiple days. For grade three, the CDE and ETS strongly recommend that no more than one test part be administered in a day. (California *Code of Regulations, Title 5,* Education, Division 1, Chapter 2, Subchapter 3.75, Article 2, § 855; http://www.cde.ca.gov/ta/tg/sr/starregs0207cln.doc).

Significant Developments in 2009

Elimination of California Achievement Tests, Sixth Edition Survey (CAT/6 Survey)

The CAT/6 Survey, the STAR Program's designated norm-referenced test (NRT) administered in grades three and seven, was eliminated from the spring 2009 administration. As a result of the Budget Act of 2008, the California *Education Code* Section 60640 was amended by Assembly Bill 519 to remove the NRT from subsequent test administrations of the STAR Program. The elimination of the NRT reduced testing time for students in grades three and seven.

New Cut Scores and Score Scales

A standard setting was held in the fall of 2008 to establish new cut scores for the below basic, basic, proficient, and advanced performance levels for the ELA and mathematics tests in grades three through five and for the science test in grade five. A new scale for reporting CMA test results was developed and used to report scores after the spring 2009 administration.

ELA scores for students in grades four and seven included the results for the California Writing Modified Assessment if the student took this component of the ELA test. In grade four, the result was included in the scale score and reporting cluster results. In grade seven, the result was included with the percent correct score reported for ELA and also reported separately as a numerical score on the CMA Student Report.

Testing in Grades Six through Eight

The first operational tests for ELA in grades six through eight, ,grades six to seven in mathematics, and for science in grade eight were administered. Scores on these tests were reported as a percent correct of the whole content area. The CMA for ELA (Grade 7) included a Writing component. Cut scores and scale scores for reporting on these tests will be set during the standard setting scheduled for fall 2009.

Limitations of the Assessment

Score Interpretation

A school district may use CMA results to help make decisions about student placement, promotion, retention, or other considerations related to student achievement. However, it is important to remember that a single test can provide only limited information. Other relevant information should be considered as well. It is advisable for parents to evaluate their child's strengths and weaknesses in the relevant topics by reviewing classroom work and progress reports in addition to the child's CMA results (CDE, 2009). It is also important to note that, statistically, a student's score in a content area probably contains a standard measurement error and could vary somewhat if the student was retested.

Out-of-Level Testing

Testing below a student's grade is not allowed for the CMA or any test in the STAR Program; all students are required to take the test for the grade in which they are enrolled. Districts are advised to review all IEPs to ensure that any provision for testing below a student's grade level has been removed. The student's IEP team makes the decision annually by evaluating the student's progress on multiple measures. The IEP team must specify annually the CMA content area(s) the student is assigned to take

Score Comparisons

When comparing results for the CMA, compare results only within a given year, content area, and grade. No direct comparisons should be made between grades or between content areas because the scale scores vary over grades and content areas. The user may compare results for the same subject and grade, within a school, between schools, or between a school and its district, or its county. In addition, scores obtained in 2009 should not be compared with those obtained in previous years because scale scores were used to report scores in 2009.

When comparing percent correct scores reported for the CMA tests in grades six through eight, the user is limited to making comparisons within the same content area and grade. No direct comparisons should be made between grades or between content areas.

Groups and Organizations involved with the STAR Program

State Board of Education (SBE)

The State Board of Education (SBE) is the governing and policy-making body of the California Departmetn of Education. The SBE sets K–12 education policy in the areas of standards, instructional materials, assessment, and accountability. The SBE adopts textbooks for kindergarten through grade eight, adopts regulations to implement legislation, and has the authority to grant waivers of the *Education Code*.

California Department of Education (CDE)

The CDE oversees California's public school system, which is responsible for the education of more than seven million children and young adults in more than 9,000 schools. The CDE's mission is to provide leadership, assistance, oversight, and resources so that every

child in California has access to an educational system that meets world-class standards. As part of its mission to promote district and school accountability for improving student achievement as defined by the SBE, the CDE oversees the development and administration of the STAR Program.

Contractors

Educational Testing Service (ETS)

The CDE and the SBE contracted with Educational Testing Service (ETS) to develop and administer the STAR Program. As the prime contractor, ETS has overall responsibility for working with the CDE to improve its overall assessment system and to coordinate the work of ETS and its subcontractor Pearson. Activities directly conducted by ETS include overall management of the program activities; development of all test items; construction and production of test booklets and related test materials; support and training provided to counties, school districts, and independently testing charter schools; implementation and maintenance of the STAR Management System for orders of materials and pre-identification services; and completion of all psychometric activities.

Pearson

ETS also monitors and manages the work of Pearson, subcontractor to ETS for the STAR Program. Activities conducted by Pearson include production of all scannable test materials; packaging, distribution, and collection of test materials to school districts and independently testing charter schools; scanning and scoring of all responses, including performance scoring of the writing responses; and production of all score reports and data files of test results.

Overview of the Technical Report

This technical report addresses the characteristics of the CMA administered in spring 2009. The technical report contains eight additional chapters as follows:

- Chapter 2 presents a conceptual overview of processes involved in a testing cycle for a CMA. This includes test construction, test administration, generation of test scores, and dissemination of score reports. Information about the distributions of scores aggregated by subgroups based on demographics and the use of special services is also included in this chapter.
- Chapter 3 describes the procedures followed during the development of valid CMA items; the chapter explains the process of field-testing new items and the review of items by contractors and content experts.
- Chapter 4 details the content and psychometric criteria applicable to the construction of CMA tests for 2009.
- Chapter 5 presents the processes involved in the actual administration of the 2009 CMA tests with an emphasis on efforts made to ensure standardization f the tests. It also includes a detailed section that describes the procedures that were followed by ETS to ensure test security.
- Chapter 6 describes the standard setting process previously conducted for newly introduced CMA. In addition, results describing students' proficiency classifications and their reported scores are also provided.
- Chapter 7 details the types of scores and score reports that are produced at the end of each administration of the CMA.
- Chapter 8 summarizes the results of the test and item-level analyses performed during the spring 2009 administration of the tests. These include the classical item analyses,

the reliability analyses that include assessments of test reliability and the consistency and accuracy of the CMA proficiency-level classifications, and the procedures designed to ensure the validity of CMA score uses and interpretations. Also discussed in this chapter are the, item response theory (IRT) and model-fit analyses, as well as documentation of the equating, along with CMA conversion tables. Finally, the chapter summarizes the results of analyses investigating the differential item functioning (DIF) for each CMA.

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

Each chapter contains summary tables in the body of the text. However, extended appendixes that give more detailed information are provided at the end of the relevant chapters.

Reference

California Department of Education. (2009). *Interpreting 2009 STAR program test results*. Sacramento, CA. http://www.cde.ca.gov/ta/tg/sr/documents/star09intrprslt.pdf.

Chapter 2: An Overview of CMA Processes

This chapter provides an overview of the processes involved in a typical test development and administration cycle for a CMA test. Also described are the specifications maintained by ETS to carry out each of those processes. The chapter is organized to provide a brief description of each process followed by a summary of the associated specifications. More details about the specifications and the analyses associated with each process are described in other chapters that are referenced below.

Item Development

Item Formats

All CMA contain three-option multiple-choice items. The CMA for ELA in grades four and seven include a writing component that contains one Writing task.

Item Development Specifications

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

The item specifications describe the characteristics of the items that should be written to measure each content standard. The item specifications help ensure that the items in the CMA measure the content standards in the same way. To do this, the item specifications provide detailed information to item writers that are developing items for the CMA.

The items selected for the CMA undergo an extensive item review process that is designed to provide the best standards-based tests possible. Details about the item development specifications, the item review process, the item utilization plan, and the rules for arranging items on the forms are presented in Chapter 3, starting on page 60.

Item Banking

The newly developed items are placed in the item bank along with the corresponding information obtained at the review sessions. Items that are accepted by the content experts are updated to a “field-test ready” status; items that are rejected are updated to a “rejected before use” status. ETS then delivers the items to the CDE by means of a delivery of the STAR electronic item bank.

Items are then field tested to obtain information about item performance and to obtain statistics that can be used to assemble operational forms. ETS then prepares the items and the associated statistics for another round of review by content experts. Subsequent updates to items are based on the operational use of the items. However, only the latest content of the item is retained in the bank at any time, along with the administration data from every administration that has included the item.

Further details on item banking are presented on page 69 in Chapter 3.

Item Refresh Rate

The Item Utilization Plan assumes that each year, 50 percent of items on an operational form are refreshed; these items remain in the item bank for future use.

Test Assembly

Test Length

The number of items in CMA and the estimated time to complete a test is presented in Appendix 2.A on page 17. In 2009, the CMA tests for grades three through five ELA consist of 48 operational items and nine field test items. The CMA for ELA administered in grades six through eight consist of 54 operational items and nine field test items. The CMA for grades three through five mathematics and grade five science are composed of 48 operational items and nine field test items each. The CMA for grades six and seven mathematics and for grade eight science consist of 54 operational items and nine field test items. The considerations in deciding upon the test length are described on page 71 in Chapter 4.

Test Blueprint

ETS selects all CMA test items to conform with the SBE-approved California content standards and test blueprints. The content blueprints for the CMA can be found on the CDE "STAR CMA Blueprints" Web page at http://www.cde.ca.gov/ta/tg/sr/cmablueprints.asp.

Although the test blueprints specify the number of items at the individual standard level, scores on the CMA items in grades three through five are grouped into subcontent areas, referred to as "reporting clusters." These data were not available for the CMA tests in grades six through eight. For each CMA reporting cluster, the percentage of questions correctly answered is reported on a student's score report. A description of the CMA reporting clusters for grades three through five and the standards that compose each are provided in Appendix 2.B, which begins on page 18.

Content Rules and Item Selection

When developing a new test for a given grade and content area, test developers follow a number of rules. First and foremost, they select items that meet the blueprint for that grade and content area. Using an electronic item bank, assessment specialists begin by identifying a number of linking items. These are items that appeared in the previous year's operational administration and they are used to equate the test forms administered each year. After the linking items are approved, assessment specialists populate the rest of the test form. In 2009, the CMA test forms used in 2008 were re-used and a new score scale was developed, so no linking back to the previously used score scales was required.

Another consideration is the difficulty of each item. Test developers strive to ensure that there are some easy and some hard items and that there are a number of items in the middle range of difficulty. The detailed rules are presented in Chapter 4, which begins on page 71.

Psychometric Criteria

For the CMA, the test developers and psychometricians strive to accomplish three goals while developing a test:

1. The test must have desired precision of measurement at all ability levels.
2. The test score must be valid and reliable for the intended population and for the various subgroups of test takers.
3. The test forms must be comparable across years of administration to ensure the generalizability of scores over time.

In order to achieve these goals, a set of rules is developed that outlines the desired psychometric properties of each CMA. Such rules are referred to as statistical targets. Two types of assembly targets are developed for each CMA: the total test target and reporting cluster targets. These targets are provided to test developers before a test construction cycle begins. The test developers and psychometricians work together to design the tests to these targets.

These specifications were developed from the analyses of field-test forms administered in fall 2007 for ELA and mathematics in grades three through five and for science in grade five. They were used to develop the spring 2008 operational forms, which were re-used in the spring of 2009. The target values for ELA in grades six through eight, mathematics in grades six and seven, and science in grade eight were developed using data from a fall 2008 field test and were used to develop the spring 2009 operational ELA, mathematics, and science forms for these grades. The target values for all grades and subject areas are presented in Table 4.1, on page 73.

The items in test forms are organized and sequenced differently according to the requirements of the content area. Further details on the arrangement of items during test assembly are also described in Chapter 4, which begins on page 71.

Test Administration

It is of utmost priority to ETS to administer the CMA in an appropriate, consistent, confidential, and standardized manner.

Test Security and Confidentiality

All tests within the STAR Program are secure documents. For the CMA administration, every person having access to test materials maintains the security and confidentiality of the tests. ETS's Code of Ethics requires that all test information, including tangible materials (such as test booklets, test questions, test results), confidential files, processes, and activities are kept secure. To ensure security for all the tests that ETS develops or handles, ETS maintains an Office of Testing Integrity (OTI). A detailed description of the OTI and its mission is presented in Chapter 5, on page 74.

In its pursuit of enforcing secure practices, ETS and the OTI strive to safeguard the various processes involved in a test development and administration cycle. Those processes are listed below. The practices related to each process are discussed in detail in Chapter 3, starting on page 60.

- Test development process
- Item and data review process
- Item banking
- Transfer of forms and Items to the CDE
- Security of electronic files via firewall
- Printing and publishing
- Test administration
- Test delivery
- Processing and scoring
- Data management
- Transfer of scores via secure data exchange
- Statistical analysis

- Reporting and posting results
- Student confidentiality
- Student test results

Procedures to Maintain Standardization

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

Test Administrators

The CMA tests are administered in conjunction with the other tests that comprise the STAR Program. In that respect, ETS employs personnel who facilitate various processes involved in the standardization of an administration cycle. Staff at school districts that are central to the processes include district STAR coordinators, test examiners, proctors, and scribes. The responsibilities for each of the staff members are included in the *STAR District and Test Site Coordinator Manua*l (CDE, 2009), which is presented in more detail on page 80 in Chapter 5.

Test Directions

ETS maintains a series of instructions compiled in detailed manuals that are available to the test administrators. Such documents include, but are not limited to, the following:

***Directions for Administration* (*DFA*s)**—Manuals used by test examiners to administer the CMA to students to be followed exactly so that all students have an equal opportunity to demonstrate their academic achievement

District and Test Site Coordinator Manual—Test administration procedures

STAR Management System manuals—Instructions for the Web-based modules that allow district STAR coordinators to set up test administrations, order materials, and submit and correct student Pre-ID data; every module has its own user manual with detailed instructions on how to use the STAR Management System

Test Variations and Accommodations

All public school students participate in the STAR Program, including students with disabilities and English learners. Most students with disabilities and English learners take the CMA under standard conditions. Some students with disabilities and English learners, however, may need assistance when taking the CMA. This assistance takes the form of test variations or accommodations. All students in these categories may have test administration directions simplified or clarified. In addition, all eligible students may have test variations, as defined on the next page, if these variations are regularly used in the classroom. They also must be allowed to use the accommodations that are specified in each student's individualized education program (IEP) or Section 504 plan. These accommodations must match the one(s) used for classroom work throughout the year.

Accommodations change the way the test is given but do not change what is tested. The purpose of test variations, and accommodations is to enable the students to take the CMA, not to give them an advantage over other students or to improve their scores. Test administration variations and accommodations do not result in changes to the students' scores for API or AYP calculations.

Test variations, and accommodations for the statewide assessments, including the STAR Program, are defined as follows:

Category 1: Test Variations—Eligible students may have test variations if regularly used in the classroom. For example, students may take a test in a group smaller than the regular testing group or take the test individually. They also may use special lighting, adaptive furniture, or magnifying equipment.

Category 2: Accommodations—Eligible students are permitted to take the CMA with accommodations if specified in their IEP or Section 504 plan for use on the CMA or for use during classroom instruction and assessment. Examples of accommodations are large-print or braille versions of the CMA or providing more than one day for a test designed for a single sitting.

Appendix 2.C presents a complete list of test variations and accommodations for the 2009 administration.

Accommodation Summaries

The percentage of students using various testing accommodations during the 2009 administration of CMA are presented in Appendix 2.D. The data are organized into three sections within each table. The first section presents the percentages of students using each accommodation in the total testing population. The next section presents the results for students in special education and not in special education. The final section presents the results for various categories based on English-language proficiency. The information within each section is presented for the relevant grades. Most accommodations are common across CMA tests, although the CMA for grades four and seven ELA also include accommodations related to the Writing task. Additional accommodations are included for CMA for mathematics that involved the use of calculators, arithmetic tables, and manipulatives.

Scores

The CMA total raw scores equal the sum of examinees' scores on the multiple-choice test items. In grades four and seven, the total ELA raw score equals the sum of examinees' scores on both the multiple-choice items and the Writing task. The Writing score is reported on a four-point scale with possible scores of 0, 1, 2, 3, and 4.

Raw scores on each CMA in grades three through five are transformed to three-digit scale scores using the equating process described starting on page 14. CMA results are reported through the use of these scale scores; the scores range from 150 to 600 for each test. Also reported are performance levels obtained by categorizing the scale score into one of the following categories: far below basic, below basic, basic, proficient, or advanced. The state's target is for all students to score at the proficient or advanced level.

In addition to total scale scores, CMA performance on various reporting clusters is reported for grades three through five. The subscore or reporting cluster score is obtained by summing an examinee's scores on the items in each reporting cluster. That information is reported in terms of a percent correct score.

Detailed descriptions of CMA scores are described in Chapter 7, which starts on page 89.

Aggregation Procedures

In order to provide meaningful results to the stakeholders, CMA scores for a given grade and content area are aggregated at the school, independent charter school, district, county, and state levels. The aggregated scores are generated both for individual scores as well as

group scores. The following section presents the types of aggregation performed on CMA scores.

Individual Scores

Table 7.1 and Table 7.2 in Chapter 7 provide summary statistics describing student performance on each CMA. Included in the tables are the means and standard deviations of student scores expressed in terms of both raw scores and scale scores. The raw score means and standard deviations also are expressed as percentages of the total raw score points in each test. Finally the tables show the percentages of students in each performance category.

Group Scores

Statistics summarizing student performance by content area and grade for selected groups of students are provided in Table 7.C.1 through Table 7.C.7 in Appendix 7.C, which starts on page 110. In the tables, students are grouped by demographic characteristics, including gender, ethnicity, English proficiency, economic status, and primary disability. The tables show the numbers of valid cases in each group, scale score means and standard deviations as well as the mean percent correct for each reporting cluster for each demographic group. Table 2.1 defines the demographic groups included in the tables. Students' economic status is determined by considering the education level of their parents and whether or not they were eligible for the National School Lunch Program (NSLP).

Table 2.1 Subgroup Definitions

Subgroup	Definition
Gender	Male Female
Ethnicity	American Indian or Alaska Native Asian – Chinese – Japanese – Korean – Vietnamese – Asian Indian – Laotian – Cambodian – Other Asian Pacific Islander – Native Hawaiian – Guamanian – Samoan – Tahitian – Other Pacific Islander Filipino Hispanic or Latino African American White (not Hispanic)
English Language Fluency	English-language fluency Initially fluent English proficient English learner Reclassified fluent English proficient

Subgroup	Definition
Economic Status	Not economically disadvantaged Economically disadvantaged
Special Services	No special services Special services

Equating

The 2008 CMA tests were re-used in 2009 and a new scale for the CMA test was developed using 2009 test data. Therefore there was no equating of the CMA tests to previously used scales.

Calibration

To obtain item calibrations, a proprietary version of PARSCALE program is used. The estimation process is constrained by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588) and by setting the lower asymptote for all multiple-choice items to zero. The resulting estimation is equivalent to the Rasch model for multiple-choice items and the Rasch partial credit model for polytomously scored items, which was used to obtain calibrations for the essays in the CMA for ELA tests administered in grades four and seven. This approach is in keeping with previous CMA calibration procedures accomplished using the WINSTEPS program (Linacre, 2000). The PARSCALE calibrations are run in two stages, following procedures used with other ETS testing programs. In the first stage, estimation imposed normal constraints on the updated prior ability 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 (EM) cycle with no constraints. For both stages, the metric of the scale is controlled by the constant discrimination parameter.

Scaling

The item calibrations for CMA tests in grades three through five were used to generate raw score to theta scoring tables. The thetas in these tables then were linearly transformed to a three-digit score scale that ranged from 150 to 600. Since the basic and proficiency cut scores were required to be equal to 300 and 350, respectively, the following formula was used to make this transformation:

$$ScaleScore = \left(350 - \hat{\theta}_{pro} \times \left(\frac{350-300}{\hat{\theta}_{pro} - \hat{\theta}_{bas}} \right) \right) + \left(\frac{350-300}{\hat{\theta}_{pro} - \hat{\theta}_{bas}} \right) \times \hat{\theta} \tag{2.1}$$

where,

$\hat{\theta}$ represents student ability

$\hat{\theta}_{pro}$ represents theta cut-score for proficient on spring 2009 base scale

$\hat{\theta}_{bas}$ represents theta cut-score for basic on spring 2009 base scale

Complete raw-to-scale score conversion tables for the 2009 CMA are presented in Table 8.D.20 through Table 8.D.28, starting on page 195. The raw scores and corresponding unrounded converted scale scores are listed in those tables. The scale scores defining the cut scores for all performance levels are presented in Table 6.1 on page 87.

Equating the Braille Versions of the CMA

In some cases, it is not possible to translate all of the items contained in a CMA into braille. In 2009, this was the case for the CMA for ELA (Grade 3). One item in this test could not be translated into braille so it did not appear in the translated test.

This situation requires that a new scoring table be developed for the shortened grade three test. To obtain this table, the shortened test is equated to the full-length test using the IRT equating methods. The braille test is equated to a reference form using a common-item nonequivalent groups design and methods based on item response theory. The “base” or “reference” calibrations were those estimated for the full length, operational CMA grade three ELA test. This process ensures that the scaled cut scores established for the full-length test are used to classify students who take the shorter test.

References

California Department of Education. (2009). *2009 STAR district and test site coordinator manual.* Sacramento, CA. http://www.startest.org/pdfs/STAR.coord_man.2009.pdf.

Educational Testing Service. (2002). *ETS standards for quality and fairness*. Office of Testing Integrity, Princeton, NJ: Educational Testing Service.

Linacre, J. M. (2000). *WINSTEPS: Rasch measurement* (Version 3.23). Chicago, IL: MESA Press.

Appendix 2.A—CMA Item and Estimated Time Chart

California Modified Assessment	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
	Total No. of Items	Time	Total No. of Item	Time	Total No of Items	Time	Total No. of Items	Time	Total No. of Items	Time	Total No. of Items	Time
English–Language Arts		**180**		**135**		**135**		**165**		**165**		**165**
Part 1		45		45		45		55		55		55
Part 2	57	45	57	45	57	45	63–	55	63	55	63	55
Part 3		45		45		45		55		55		55
Part 4 – only grade 3		45	–	–	–	–	–	–	–	–	–	–
Writing Application	–	–	1	70[1]	–	–	–	–	1	70[2]	–	–
Mathematics		**140**		**105**		**105**		**120**		**120**		–
Part 1		35	57	35	57	35	63	40	63	40	N/A	–
Part 2	57	35		35		35		40		40		–
Part 3		35		35		35		40		40		–
Part 4 – only grade 3		35	–	–	–	–	–	–	–	–	–	–
Science		–		–		**120**		–		–		**135**
Part 1	–		–		57	40	–		–		63	45
Part 2						40						45
Part 3						40						45

[1] The writing test in grade four is given on a separate date from the multiple-choice tests. Writing test times are not included in the estimated time for ELA multiple-choice tests.

[2] The writing test in grade seven is given on a separate date from the multiple-choice tests. Writing test times are not included in the estimated time for ELA multiple-choice tests.

Appendix 2.B—Reporting Clusters

English–Language Arts

English–Language Arts Modified Standards Assessment (Grade Three)

Multiple Choice	
Vocabulary	14 items
Reading for Understanding	17 items
Language	17 items

English–Language Arts Modified Standards Assessment (Grade Four)

Multiple Choice	
Vocabulary	11 items
Reading for Understanding	16 items
Language	21 items
Writing	
Writing Applications	4 points (1 task)

English–Language Arts Modified Standards Assessment (Grade Five)

Multiple Choice	
Vocabulary	8 items
Reading for Understanding	18 items
Language	22 items

Mathematics

Mathematics Modified Standards Assessment (Grade Three)

Multiple Choice	
Number Sense	24 items
Algebra and Data Analysis	13 items
Measurement and Geometry	11 items

Mathematics Modified Standards Assessment (Grade Four)

Multiple Choice	
Number Sense	23 items
Algebra and Data Analysis	15 items
Measurement and Geometry	10 items

Mathematics Modified Standards Assessment (Grade Five)

Multiple Choice	
Number Sense	21 items
Algebra and Data Analysis	17 items
Measurement and Geometry	10 items

Science

Science Modified Standards Assessment (Grade Five)

Multiple Choice	
Physical Sciences	16 items
Life Sciences	16 items
Earth Sciences	16 items

Appendix 2.C—2009 Test Variations and Accommodations

Table 2.C.1 List of Test Variations and Accommodations

Test Variation (1) or Accommodation (2)	Category
Test administration directions that are simplified or clarified (does not apply to test questions)	ALL
Student marks in test booklet (other than responses) including highlighting	ALL For grade 3 marks must be removed to avoid scanning interference or transcribe
Test students in a small group setting	ALL
Extra time on a test within a testing day	ALL
Test individual student separately, provided that a test examiner directly supervises the student	1
Visual magnifying equipment	1
Audio amplification equipment	1
Noise buffers (e.g., individual carrel or study enclosure)	1
Special lighting or acoustics; special or adaptive furniture	1
Colored overlay, mask, or other means to maintain visual attention	1
Manually Coded English or American Sign Language to present directions for administration (does not apply to test questions)	1
Student marks responses in test booklet and responses are transferred to a scorable answer document by an employee of the school, district, or nonpublic, nonsectarian school	2
Responses dictated (orally, or in Manually Coded English or American Sign Language) to a scribe for selected-response items (multiple-choice questions)	2
Word processing software with spell and grammar check tools turned off for the essay responses (writing portion of the test)	2
Essay responses dictated orally or in Manually Coded English to a scribe, audio recorder, or speech-to-text converter and the student provides all spelling and language conventions	2
Assistive device that does not interfere with the independent work of the student on the multiple-choice and/or essay responses (writing portion of the test)	2
Braille transcriptions provided by the test contractor	2
Large-print versions Test items enlarged if font larger than required on large-print versions	2
Test over more than one day for a test or test part to be administered in a single sitting	2
Supervised breaks within a section of the test	2
Administration of the test at the most beneficial time of day to the student	2
Test administered at home or in hospital by a test examiner	2

Test Variation (1) or Accommodation (2)	Category
Manually Coded English or American Sign Language to present test questions	2 Math and Science
	2 ELA (including Writing)
Answer options read aloud to student	2
Test questions read aloud to student or used audio CD presentation	2 Math and Science
	2 ELA (including Writing)
Calculator on the mathematics tests	2 Grades 5–7
Calculator on the science tests	2
Math manipulatives on the mathematics tests	2
Math manipulatives on the science tests	2
Unlisted Accommodation	Check with CDE prior to use

Test Variation (1) = Eligible students may have testing variations if regularly used in the classroom.

Accommodation (2) = Eligible students shall be permitted to take the examination/test with accommodations if specified in the eligible student's IEP or Section 504 plan for use on the examination, standardized testing, or for use during classroom instruction and assessment.

Appendix 2.D—Accommodation Summary Tables

Table 2.D.1 Accommodation Summary for ELA, Grade Three

Accommodation Summary for ELA, Grade Three		
All Tested	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	122	1.08%
C: Dictated responses to a scribe	28	0.25%
F: Used non-interfering assistive device	53	0.47%
G: Used braille test	7	0.06%
H: Used large-print test	49	0.43%
J: Tested over more than one day	590	5.21%
K: Had supervised breaks	1,633	14.42%
L: Most beneficial time of day	903	7.98%
M: Administered at home or in a hospital	2	0.02%
O: Examiner presented with MCE or ASL	22	0.19%
P: Used audio presentation (CD)	42	0.37%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	418	3.69%
Y: Leave blank	21	0.19%
Z: Examiner read test questions aloud	1,855	16.39%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	3,379	29.85%
English Learner Test Variation A	21	0.19%
English Learner Test Variation B	31	0.27%
English Learner Test Variation C	53	0.47%
Any Accommodation or EL Variation	3,811	33.66%
No Accommodation or EL Variation	7,510	66.34%
English-Only Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	56	0.91%
C: Dictated responses to a scribe	20	0.33%
F: Used non-interfering assistive device	36	0.59%
G: Used braille test	2	0.03%
H: Used large-print test	28	0.46%
J: Tested over more than one day	330	5.36%
K: Had supervised breaks	884	14.37%
L: Most beneficial time of day	501	8.14%
M: Administered at home or in a hospital	2	0.03%
O: Examiner presented with MCE or ASL	16	0.26%
P: Used audio presentation (CD)	19	0.31%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	234	3.80%
Y: Leave blank	7	0.11%
Z: Examiner read test questions aloud	987	16.04%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,834	29.81%
English Learner Test Variation A	1	0.02%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	4	0.07%
Any Accommodation or EL Variation	2,051	33.33%
No Accommodation or EL Variation	4,102	66.67%

Accommodation Summary for ELA, Grade Three		
Initially Fluent English Proficient (I-FEP) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	1	0.69%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.69%
J: Tested over more than one day	7	4.83%
K: Had supervised breaks	20	13.79%
L: Most beneficial time of day	13	8.97%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.69%
P: Used audio presentation (CD)	1	0.69%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	6	4.14%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	19	13.10%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	37	25.52%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	45	31.03%
No Accommodation or EL Variation	100	68.97%
English Learner (EL) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	60	1.28%
C: Dictated responses to a scribe	8	0.17%
F: Used non-interfering assistive device	17	0.36%
G: Used braille test	5	0.11%
H: Used large-print test	19	0.41%
J: Tested over more than one day	238	5.08%
K: Had supervised breaks	667	14.25%
L: Most beneficial time of day	358	7.65%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	4	0.09%
P: Used audio presentation (CD)	22	0.47%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	168	3.59%
Y: Leave blank	9	0.19%
Z: Examiner read test questions aloud	815	17.41%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,418	30.29%
English Learner Test Variation A	17	0.36%
English Learner Test Variation B	29	0.62%
English Learner Test Variation C	49	1.05%
Any Accommodation or EL Variation	1,604	34.26%
No Accommodation or EL Variation	3,078	65.74%

Accommodation Summary for ELA, Grade Three		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	1	2.70%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	0	0.00%
K: Had supervised breaks	4	10.81%
L: Most beneficial time of day	1	2.70%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	0	0.00%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	2	5.41%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	6	16.22%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	6	16.22%
No Accommodation or EL Variation	31	83.78%

Table 2.D.2 Accommodation Summary for ELA, Grade Four

Accommodation Summary for ELA, Grade Four		
All Tested	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	575	3.36%
C: Dictated responses to a scribe	25	0.15%
F: Used non-interfering assistive device	77	0.45%
G: Used braille test	3	0.02%
H: Used large-print test	65	0.38%
J: Tested over more than one day	761	4.44%
K: Had supervised breaks	2,489	14.54%
L: Most beneficial time of day	1,295	7.56%
M: Administered at home or in a hospital	6	0.04%
O: Examiner presented with MCE or ASL	32	0.19%
P: Used audio presentation (CD)	99	0.58%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	655	3.83%
Y: Leave blank	29	0.17%
Z: Examiner read test questions aloud	2,717	15.87%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	5,373	31.38%
English Learner Test Variation A	24	0.14%
English Learner Test Variation B	30	0.18%
English Learner Test Variation C	52	0.30%
Any Accommodation or EL Variation	5,841	34.12%
No Accommodation or EL Variation	11,280	65.88%
English-Only Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	361	3.78%
C: Dictated responses to a scribe	18	0.19%
F: Used non-interfering assistive device	43	0.45%
G: Used braille test	1	0.01%
H: Used large-print test	43	0.45%
J: Tested over more than one day	435	4.55%
K: Had supervised breaks	1,357	14.19%
L: Most beneficial time of day	716	7.49%
M: Administered at home or in a hospital	5	0.05%
O: Examiner presented with MCE or ASL	23	0.24%
P: Used audio presentation (CD)	42	0.44%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	349	3.65%
Y: Leave blank	12	0.13%
Z: Examiner read test questions aloud	1,450	15.17%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	2,957	30.93%
English Learner Test Variation A	3	0.03%
English Learner Test Variation B	1	0.01%
English Learner Test Variation C	2	0.02%
Any Accommodation or EL Variation	3,197	33.44%
No Accommodation or EL Variation	6,363	66.56%

Accommodation Summary for ELA, Grade Four		
Initially Fluent English Proficient (I-FEP) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	4	1.18%
C: Dictated responses to a scribe	1	0.29%
F: Used non-interfering assistive device	2	0.59%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	17	5.00%
K: Had supervised breaks	46	13.53%
L: Most beneficial time of day	28	8.24%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	7	2.06%
P: Used audio presentation (CD)	5	1.47%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	17	5.00%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	44	12.94%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	93	27.35%
English Learner Test Variation A	1	0.29%
English Learner Test Variation B	2	0.59%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	102	30.00%
No Accommodation or EL Variation	238	70.00%
English Learner (EL) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	197	2.78%
C: Dictated responses to a scribe	6	0.08%
F: Used non-interfering assistive device	32	0.45%
G: Used braille test	1	0.01%
H: Used large-print test	21	0.30%
J: Tested over more than one day	302	4.27%
K: Had supervised breaks	1,064	15.03%
L: Most beneficial time of day	540	7.63%
M: Administered at home or in a hospital	1	0.01%
O: Examiner presented with MCE or ASL	2	0.03%
P: Used audio presentation (CD)	52	0.73%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	280	3.96%
Y: Leave blank	17	0.24%
Z: Examiner read test questions aloud	1,206	17.04%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	2,275	32.14%
English Learner Test Variation A	20	0.28%
English Learner Test Variation B	27	0.38%
English Learner Test Variation C	50	0.71%
Any Accommodation or EL Variation	2,492	35.21%
No Accommodation or EL Variation	4,586	64.79%

Accommodation Summary for ELA, Grade Four		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	7	7.00%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	1	1.00%
H: Used large-print test	1	1.00%
J: Tested over more than one day	5	5.00%
K: Had supervised breaks	11	11.00%
L: Most beneficial time of day	5	5.00%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	7	7.00%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	13	13.00%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	36	36.00%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	37	37.00%
No Accommodation or EL Variation	63	63.00%

Table 2.D.3 Accommodation Summary for ELA, Grade Five

Accommodation Summary for ELA, Grade Five		
All Tested	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	518	2.91%
C: Dictated responses to a scribe	19	0.11%
F: Used non-interfering assistive device	72	0.40%
G: Used braille test	1	0.01%
H: Used large-print test	48	0.27%
J: Tested over more than one day	809	4.54%
K: Had supervised breaks	2,611	14.65%
L: Most beneficial time of day	1,394	7.82%
M: Administered at home or in a hospital	15	0.08%
O: Examiner presented with MCE or ASL	19	0.11%
P: Used audio presentation (CD)	76	0.43%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	577	3.24%
Y: Leave blank	48	0.27%
Z: Examiner read test questions aloud	2,691	15.10%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	5,318	29.84%
English Learner Test Variation A	9	0.05%
English Learner Test Variation B	31	0.17%
English Learner Test Variation C	45	0.25%
Any Accommodation or EL Variation	5,854	32.85%
No Accommodation or EL Variation	11,967	67.15%
English-Only Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	325	3.32%
C: Dictated responses to a scribe	14	0.14%
F: Used non-interfering assistive device	45	0.46%
G: Used braille test	0	0.00%
H: Used large-print test	35	0.36%
J: Tested over more than one day	453	4.63%
K: Had supervised breaks	1,367	13.98%
L: Most beneficial time of day	749	7.66%
M: Administered at home or in a hospital	11	0.11%
O: Examiner presented with MCE or ASL	13	0.13%
P: Used audio presentation (CD)	44	0.45%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	322	3.29%
Y: Leave blank	20	0.20%
Z: Examiner read test questions aloud	1,385	14.17%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	2,862	29.28%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	3	0.03%
English Learner Test Variation C	4	0.04%
Any Accommodation or EL Variation	3,126	31.98%
No Accommodation or EL Variation	6,650	68.02%

Accommodation Summary for ELA, Grade Five		
Initially Fluent English Proficient (I-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	9	2.24%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	4	1.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	15	3.74%
K: Had supervised breaks	47	11.72%
L: Most beneficial time of day	30	7.48%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.25%
P: Used audio presentation (CD)	1	0.25%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	5	1.25%
Y: Leave blank	1	0.25%
Z: Examiner read test questions aloud	52	12.97%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	105	26.18%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	108	26.93%
No Accommodation or EL Variation	293	73.07%
English Learner (EL) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	181	2.43%
C: Dictated responses to a scribe	4	0.05%
F: Used non-interfering assistive device	23	0.31%
G: Used braille test	0	0.00%
H: Used large-print test	12	0.16%
J: Tested over more than one day	333	4.47%
K: Had supervised breaks	1,161	15.59%
L: Most beneficial time of day	604	8.11%
M: Administered at home or in a hospital	4	0.05%
O: Examiner presented with MCE or ASL	5	0.07%
P: Used audio presentation (CD)	31	0.42%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	239	3.21%
Y: Leave blank	27	0.36%
Z: Examiner read test questions aloud	1,221	16.40%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	2,290	30.76%
English Learner Test Variation A	9	0.12%
English Learner Test Variation B	28	0.38%
English Learner Test Variation C	41	0.55%
Any Accommodation or EL Variation	2,556	34.33%
No Accommodation or EL Variation	4,889	65.67%

Accommodation Summary for ELA, Grade Five		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	1	0.64%
C: Dictated responses to a scribe	1	0.64%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	1	0.64%
H: Used large-print test	0	0.00%
J: Tested over more than one day	7	4.49%
K: Had supervised breaks	29	18.59%
L: Most beneficial time of day	8	5.13%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	9	5.77%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	27	17.31%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	47	30.13%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	50	32.05%
No Accommodation or EL Variation	106	67.95%

Table 2.D.4 Accommodation Summary for ELA, Grade Six

Accommodation Summary for ELA, Grade Six		
All Tested	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	286	1.80%
C: Dictated responses to a scribe	12	0.08%
F: Used non-interfering assistive device	35	0.22%
G: Used braille test	0	0.00%
H: Used large-print test	60	0.38%
J: Tested over more than one day	547	3.44%
K: Had supervised breaks	1,837	11.55%
L: Most beneficial time of day	893	5.62%
M: Administered at home or in a hospital	13	0.08%
O: Examiner presented with MCE or ASL	10	0.06%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	473	2.97%
Y: Leave blank	33	0.21%
Z: Examiner read test questions aloud	1,900	11.95%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	3,626	22.81%
English Learner Test Variation A	12	0.08%
English Learner Test Variation B	11	0.07%
English Learner Test Variation C	10	0.06%
Any Accommodation or EL Variation	4,032	25.36%
No Accommodation or EL Variation	11,868	74.64%
English-Only Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	166	1.93%
C: Dictated responses to a scribe	11	0.13%
F: Used non-interfering assistive device	23	0.27%
G: Used braille test	0	0.00%
H: Used large-print test	32	0.37%
J: Tested over more than one day	321	3.72%
K: Had supervised breaks	932	10.81%
L: Most beneficial time of day	481	5.58%
M: Administered at home or in a hospital	9	0.10%
O: Examiner presented with MCE or ASL	4	0.05%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	249	2.89%
Y: Leave blank	15	0.17%
Z: Examiner read test questions aloud	960	11.14%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	1,898	22.02%
English Learner Test Variation A	1	0.01%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	2,115	24.53%
No Accommodation or EL Variation	6,506	75.47%

Accommodation Summary for ELA, Grade Six		
Initially Fluent English Proficient (I-FEP) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	6	2.06%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	1	0.34%
G: Used braille test	0	0.00%
H: Used large-print test	4	1.37%
J: Tested over more than one day	11	3.78%
K: Had supervised breaks	33	11.34%
L: Most beneficial time of day	20	6.87%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	4	1.37%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	7	2.41%
Y: Leave blank	1	0.34%
Z: Examiner read test questions aloud	35	12.03%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	61	20.96%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	70	24.05%
No Accommodation or EL Variation	221	75.95%
English Learner (EL) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	108	1.60%
C: Dictated responses to a scribe	1	0.01%
F: Used non-interfering assistive device	10	0.15%
G: Used braille test	0	0.00%
H: Used large-print test	23	0.34%
J: Tested over more than one day	210	3.12%
K: Had supervised breaks	851	12.62%
L: Most beneficial time of day	382	5.67%
M: Administered at home or in a hospital	4	0.06%
O: Examiner presented with MCE or ASL	2	0.03%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	215	3.19%
Y: Leave blank	17	0.25%
Z: Examiner read test questions aloud	883	13.10%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,628	24.15%
English Learner Test Variation A	11	0.16%
English Learner Test Variation B	11	0.16%
English Learner Test Variation C	10	0.15%
Any Accommodation or EL Variation	1,804	26.76%
No Accommodation or EL Variation	4,937	73.24%

Accommodation Summary for ELA, Grade Six		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	5	2.29%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	5	2.29%
K: Had supervised breaks	19	8.72%
L: Most beneficial time of day	9	4.13%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	1	0.46%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	19	8.72%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	34	15.60%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	37	16.97%
No Accommodation or EL Variation	181	83.03%

Table 2.D.5 Accommodation Summary for ELA, Grade Seven

Accommodation Summary for ELA, Grade Seven		
All Tested	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	160	1.23%
C: Dictated responses to a scribe	10	0.08%
F: Used non-interfering assistive device	19	0.15%
G: Used braille test	0	0.00%
H: Used large-print test	32	0.25%
J: Tested over more than one day	298	2.28%
K: Had supervised breaks	1,065	8.15%
L: Most beneficial time of day	254	1.94%
M: Administered at home or in a hospital	11	0.08%
O: Examiner presented with MCE or ASL	14	0.11%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	273	2.09%
Y: Leave blank	20	0.15%
Z: Examiner read test questions aloud	856	6.55%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	1,825	13.97%
English Learner Test Variation A	12	0.09%
English Learner Test Variation B	16	0.12%
English Learner Test Variation C	8	0.06%
Any Accommodation or EL Variation	2,123	16.26%
No Accommodation or EL Variation	10,937	83.74%
English-Only Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	110	1.49%
C: Dictated responses to a scribe	5	0.07%
F: Used non-interfering assistive device	12	0.16%
G: Used braille test	0	0.00%
H: Used large-print test	18	0.24%
J: Tested over more than one day	193	2.62%
K: Had supervised breaks	613	8.31%
L: Most beneficial time of day	172	2.33%
M: Administered at home or in a hospital	6	0.08%
O: Examiner presented with MCE or ASL	14	0.19%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	146	1.98%
Y: Leave blank	11	0.15%
Z: Examiner read test questions aloud	460	6.23%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,047	14.19%
English Learner Test Variation A	1	0.01%
English Learner Test Variation B	3	0.04%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	1,199	16.25%
No Accommodation or EL Variation	6,181	83.75%

Accommodation Summary for ELA, Grade Seven		
Initially Fluent English Proficient (I-FEP) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	3	1.12%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	7	2.61%
K: Had supervised breaks	23	8.58%
L: Most beneficial time of day	1	0.37%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	4	1.49%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	7	2.61%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	24	8.96%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	30	11.19%
No Accommodation or EL Variation	238	88.81%
English Learner (EL) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	41	0.80%
C: Dictated responses to a scribe	5	0.10%
F: Used non-interfering assistive device	7	0.14%
G: Used braille test	0	0.00%
H: Used large-print test	13	0.25%
J: Tested over more than one day	87	1.70%
K: Had supervised breaks	413	8.08%
L: Most beneficial time of day	79	1.55%
M: Administered at home or in a hospital	4	0.08%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	117	2.29%
Y: Leave blank	9	0.18%
Z: Examiner read test questions aloud	367	7.18%
Accommodation is in Section 504 plan	1	0.02%
Accommodation is in IEP	706	13.81%
English Learner Test Variation A	11	0.22%
English Learner Test Variation B	13	0.25%
English Learner Test Variation C	8	0.16%
Any Accommodation or EL Variation	844	16.51%
No Accommodation or EL Variation	4,267	83.49%

Accommodation Summary for ELA, Grade Seven		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	5	1.92%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.38%
J: Tested over more than one day	10	3.85%
K: Had supervised breaks	16	6.15%
L: Most beneficial time of day	2	0.77%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	5	1.92%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	18	6.92%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	40	15.38%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	42	16.15%
No Accommodation or EL Variation	218	83.85%

Table 2.D.6 Accommodation Summary for ELA, Grade Eight

Accommodation Summary for ELA, Grade Eight		
All Tested	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	151	1.21%
C: Dictated responses to a scribe	8	0.06%
F: Used non-interfering assistive device	12	0.10%
G: Used braille test	0	0.00%
H: Used large-print test	19	0.15%
J: Tested over more than one day	228	1.83%
K: Had supervised breaks	932	7.50%
L: Most beneficial time of day	254	2.04%
M: Administered at home or in a hospital	7	0.06%
O: Examiner presented with MCE or ASL	4	0.03%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	283	2.28%
Y: Leave blank	46	0.37%
Z: Examiner read test questions aloud	646	5.20%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,709	13.74%
English Learner Test Variation A	12	0.10%
English Learner Test Variation B	4	0.03%
English Learner Test Variation C	6	0.05%
Any Accommodation or EL Variation	1,938	15.59%
No Accommodation or EL Variation	10,496	84.41%
English-Only Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	99	1.39%
C: Dictated responses to a scribe	8	0.11%
F: Used non-interfering assistive device	10	0.14%
G: Used braille test	0	0.00%
H: Used large-print test	10	0.14%
J: Tested over more than one day	132	1.85%
K: Had supervised breaks	522	7.31%
L: Most beneficial time of day	132	1.85%
M: Administered at home or in a hospital	4	0.06%
O: Examiner presented with MCE or ASL	3	0.04%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	147	2.06%
Y: Leave blank	26	0.36%
Z: Examiner read test questions aloud	336	4.71%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	964	13.50%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	3	0.04%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	1,075	15.06%
No Accommodation or EL Variation	6,064	84.94%

Accommodation Summary for ELA, Grade Eight		
Initially Fluent English Proficient (I-FEP) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	2	0.59%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.29%
J: Tested over more than one day	15	4.42%
K: Had supervised breaks	31	9.14%
L: Most beneficial time of day	13	3.83%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.29%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	8	2.36%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	11	3.24%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	43	12.68%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	48	14.16%
No Accommodation or EL Variation	291	85.84%
English Learner (EL) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	46	0.98%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	2	0.04%
G: Used braille test	0	0.00%
H: Used large-print test	8	0.17%
J: Tested over more than one day	76	1.62%
K: Had supervised breaks	357	7.61%
L: Most beneficial time of day	104	2.22%
M: Administered at home or in a hospital	3	0.06%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	127	2.71%
Y: Leave blank	19	0.40%
Z: Examiner read test questions aloud	276	5.88%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	667	14.22%
English Learner Test Variation A	12	0.26%
English Learner Test Variation B	1	0.02%
English Learner Test Variation C	6	0.13%
Any Accommodation or EL Variation	775	16.52%
No Accommodation or EL Variation	3,917	83.48%

Accommodation Summary for ELA, Grade Eight		
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	4	1.71%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	4	1.71%
K: Had supervised breaks	17	7.26%
L: Most beneficial time of day	4	1.71%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	1	0.43%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	18	7.69%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	28	11.97%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
Any Accommodation or EL Variation	33	14.10%
No Accommodation or EL Variation	201	85.90%

Table 2.D.7 Accommodation Summary for Mathematics, Grade Three

Accommodation Summary for Mathematics, Grade Three		
All Tested	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	110	0.98%
C: Dictated responses to a scribe	26	0.23%
F: Used non-interfering assistive device	52	0.46%
G: Used braille test	7	0.06%
H: Used large-print test	51	0.45%
J: Tested over more than one day	564	5.02%
K: Had supervised breaks	1,525	13.58%
L: Most beneficial time of day	851	7.58%
M: Administered at home or in a hospital	2	0.02%
O: Examiner presented with MCE or ASL	41	0.37%
P: Used audio presentation (CD)	67	0.60%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	147	1.31%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	396	3.53%
Y: Leave blank	14	0.12%
Z: Examiner read test questions aloud	3,269	29.12%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	4,200	37.41%
English Learner Test Variation A	21	0.19%
English Learner Test Variation B	31	0.28%
English Learner Test Variation C	48	0.43%
English Learner Test Variation D	10	0.09%
Any Accommodation or EL Variation	4,726	42.09%
No Accommodation or EL Variation	6,501	57.91%
English-Only Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	47	0.77%
C: Dictated responses to a scribe	20	0.33%
F: Used non-interfering assistive device	32	0.52%
G: Used braille test	2	0.03%
H: Used large-print test	28	0.46%
J: Tested over more than one day	318	5.22%
K: Had supervised breaks	812	13.32%
L: Most beneficial time of day	469	7.69%
M: Administered at home or in a hospital	2	0.03%
O: Examiner presented with MCE or ASL	25	0.41%
P: Used audio presentation (CD)	33	0.54%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	91	1.49%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	223	3.66%
Y: Leave blank	4	0.07%
Z: Examiner read test questions aloud	1,750	28.70%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	2,291	37.58%

Accommodation Summary for Mathematics, Grade Three		
English Learner Test Variation A	3	0.05%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	3	0.05%
English Learner Test Variation D	1	0.02%
Any Accommodation or EL Variation	2,556	41.92%
No Accommodation or EL Variation	3,541	58.08%
Initially Fluent English Proficient (I-FEP) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	1	0.69%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.69%
J: Tested over more than one day	9	6.25%
K: Had supervised breaks	19	13.19%
L: Most beneficial time of day	14	9.72%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	2	1.39%
P: Used audio presentation (CD)	2	1.39%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	7	4.86%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	34	23.61%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	50	34.72%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	60	41.67%
No Accommodation or EL Variation	84	58.33%
English Learner (EL) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	56	1.20%
C: Dictated responses to a scribe	6	0.13%
F: Used non-interfering assistive device	20	0.43%
G: Used braille test	5	0.11%
H: Used large-print test	19	0.41%
J: Tested over more than one day	220	4.72%
K: Had supervised breaks	636	13.64%
L: Most beneficial time of day	335	7.19%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	13	0.28%
P: Used audio presentation (CD)	30	0.64%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	52	1.12%

Accommodation Summary for Mathematics, Grade Three		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	156	3.35%
Y: Leave blank	6	0.13%
Z: Examiner read test questions aloud	1,408	30.20%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,748	37.49%
English Learner Test Variation A	16	0.34%
English Learner Test Variation B	30	0.64%
English Learner Test Variation C	45	0.97%
English Learner Test Variation D	9	0.19%
Any Accommodation or EL Variation	1,974	42.34%
No Accommodation or EL Variation	2,688	57.66%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 3**	**Pct. of Total**
B: Marked in test booklet	1	2.70%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	0	0.00%
K: Had supervised breaks	4	10.81%
L: Most beneficial time of day	1	2.70%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	2	5.41%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	0	0.00%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	7	18.92%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	11	29.73%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	11	29.73%
No Accommodation or EL Variation	26	70.27%

Table 2.D.8 Accommodation Summary for Mathematics, Grade Four

Accommodation Summary for Mathematics, Grade Four		
All Tested	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	486	3.37%
C: Dictated responses to a scribe	24	0.17%
F: Used non-interfering assistive device	61	0.42%
G: Used braille test	3	0.02%
H: Used large-print test	56	0.39%
J: Tested over more than one day	659	4.57%
K: Had supervised breaks	2,060	14.29%
L: Most beneficial time of day	1,107	7.68%
M: Administered at home or in a hospital	5	0.03%
O: Examiner presented with MCE or ASL	47	0.33%
P: Used audio presentation (CD)	127	0.88%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	160	1.11%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	555	3.85%
Y: Leave blank	24	0.17%
Z: Examiner read test questions aloud	3,973	27.56%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	5,582	38.72%
English Learner Test Variation A	22	0.15%
English Learner Test Variation B	23	0.16%
English Learner Test Variation C	44	0.31%
English Learner Test Variation D	1	0.01%
Any Accommodation or EL Variation	6,071	42.12%
No Accommodation or EL Variation	8,344	57.88%
English-Only Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	303	3.77%
C: Dictated responses to a scribe	18	0.22%
F: Used non-interfering assistive device	35	0.44%
G: Used braille test	1	0.01%
H: Used large-print test	35	0.44%
J: Tested over more than one day	371	4.61%
K: Had supervised breaks	1,120	13.92%
L: Most beneficial time of day	611	7.60%
M: Administered at home or in a hospital	4	0.05%
O: Examiner presented with MCE or ASL	31	0.39%
P: Used audio presentation (CD)	54	0.67%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	76	0.94%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	291	3.62%
Y: Leave blank	14	0.17%
Z: Examiner read test questions aloud	2,174	27.03%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	3,095	38.48%

Accommodation Summary for Mathematics, Grade Four		
English Learner Test Variation A	3	0.04%
English Learner Test Variation B	1	0.01%
English Learner Test Variation C	3	0.04%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	3,339	41.51%
No Accommodation or EL Variation	4,705	58.49%
Initially Fluent English Proficient (I-FEP) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	4	1.50%
C: Dictated responses to a scribe	1	0.37%
F: Used non-interfering assistive device	1	0.37%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.37%
J: Tested over more than one day	15	5.62%
K: Had supervised breaks	31	11.61%
L: Most beneficial time of day	21	7.87%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	7	2.62%
P: Used audio presentation (CD)	5	1.87%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	4	1.50%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	11	4.12%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	62	23.22%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	96	35.96%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	2	0.75%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	106	39.70%
No Accommodation or EL Variation	161	60.30%
English Learner (EL) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	168	2.81%
C: Dictated responses to a scribe	5	0.08%
F: Used non-interfering assistive device	25	0.42%
G: Used braille test	1	0.02%
H: Used large-print test	18	0.30%
J: Tested over more than one day	268	4.48%
K: Had supervised breaks	892	14.90%
L: Most beneficial time of day	468	7.82%
M: Administered at home or in a hospital	1	0.02%
O: Examiner presented with MCE or ASL	9	0.15%
P: Used audio presentation (CD)	68	1.14%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	77	1.29%

Accommodation Summary for Mathematics, Grade Four		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	245	4.09%
Y: Leave blank	10	0.17%
Z: Examiner read test questions aloud	1,716	28.66%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	2,347	39.20%
English Learner Test Variation A	19	0.32%
English Learner Test Variation B	20	0.33%
English Learner Test Variation C	41	0.68%
English Learner Test Variation D	1	0.02%
Any Accommodation or EL Variation	2,580	43.09%
No Accommodation or EL Variation	3,408	56.91%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 4**	**Pct. of Total**
B: Marked in test booklet	7	8.97%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	1	1.28%
H: Used large-print test	2	2.56%
J: Tested over more than one day	5	6.41%
K: Had supervised breaks	10	12.82%
L: Most beneficial time of day	3	3.85%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	3	3.85%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	6	7.69%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	15	19.23%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	35	44.87%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	36	46.15%
No Accommodation or EL Variation	42	53.85%

Table 2.D.9 Accommodation Summary for Mathematics, Grade Five

Accommodation Summary for Mathematics, Grade Five		
All Tested	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	474	3.03%
C: Dictated responses to a scribe	20	0.13%
F: Used non-interfering assistive device	63	0.40%
G: Used braille test	1	0.01%
H: Used large-print test	53	0.34%
J: Tested over more than one day	716	4.57%
K: Had supervised breaks	2,316	14.79%
L: Most beneficial time of day	1,269	8.11%
M: Administered at home or in a hospital	14	0.09%
O: Examiner presented with MCE or ASL	34	0.22%
P: Used audio presentation (CD)	112	0.72%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	148	0.95%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	561	3.58%
Y: Leave blank	28	0.18%
Z: Examiner read test questions aloud	3,917	25.02%
Accommodation is in Section 504 plan	2	0.01%
Accommodation is in IEP	5,687	36.32%
English Learner Test Variation A	9	0.06%
English Learner Test Variation B	30	0.19%
English Learner Test Variation C	42	0.27%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	6,247	39.90%
No Accommodation or EL Variation	9,409	60.10%
English-Only Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	312	3.56%
C: Dictated responses to a scribe	14	0.16%
F: Used non-interfering assistive device	39	0.45%
G: Used braille test	0	0.00%
H: Used large-print test	38	0.43%
J: Tested over more than one day	404	4.61%
K: Had supervised breaks	1,240	14.16%
L: Most beneficial time of day	695	7.94%
M: Administered at home or in a hospital	10	0.11%
O: Examiner presented with MCE or ASL	20	0.23%
P: Used audio presentation (CD)	70	0.80%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	75	0.86%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	321	3.67%
Y: Leave blank	16	0.18%
Z: Examiner read test questions aloud	2,042	23.32%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	3,108	35.49%

Accommodation Summary for Mathematics, Grade Five		
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	3	0.03%
English Learner Test Variation C	3	0.03%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	3,382	38.62%
No Accommodation or EL Variation	5,375	61.38%
Initially Fluent English Proficient (I-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	7	2.08%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	4	1.19%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	14	4.17%
K: Had supervised breaks	44	13.10%
L: Most beneficial time of day	26	7.74%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	2	0.60%
P: Used audio presentation (CD)	3	0.89%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	2	0.60%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	4	1.19%
Y: Leave blank	1	0.30%
Z: Examiner read test questions aloud	63	18.75%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	107	31.85%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	111	33.04%
No Accommodation or EL Variation	225	66.96%
English Learner (EL) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	153	2.39%
C: Dictated responses to a scribe	5	0.08%
F: Used non-interfering assistive device	20	0.31%
G: Used braille test	0	0.00%
H: Used large-print test	14	0.22%
J: Tested over more than one day	290	4.54%
K: Had supervised breaks	997	15.60%
L: Most beneficial time of day	539	8.44%
M: Administered at home or in a hospital	4	0.06%
O: Examiner presented with MCE or ASL	11	0.17%
P: Used audio presentation (CD)	36	0.56%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	69	1.08%

Accommodation Summary for Mathematics, Grade Five		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	225	3.52%
Y: Leave blank	11	0.17%
Z: Examiner read test questions aloud	1,768	27.67%
Accommodation is in Section 504 plan	2	0.03%
Accommodation is in IEP	2,405	37.64%
English Learner Test Variation A	9	0.14%
English Learner Test Variation B	27	0.42%
English Learner Test Variation C	39	0.61%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	2,680	41.95%
No Accommodation or EL Variation	3,709	58.05%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	0	0.00%
C: Dictated responses to a scribe	1	0.76%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	1	0.76%
H: Used large-print test	0	0.00%
J: Tested over more than one day	7	5.30%
K: Had supervised breaks	28	21.21%
L: Most beneficial time of day	7	5.30%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.76%
P: Used audio presentation (CD)	3	2.27%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	2	1.52%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	9	6.82%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	36	27.27%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	52	39.39%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	58	43.94%
No Accommodation or EL Variation	74	56.06%

Table 2.D.10 Accommodation Summary for Mathematics, Grade Six

Accommodation Summary for Mathematics, Grade Six		
All Tested	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	264	1.73%
C: Dictated responses to a scribe	10	0.07%
F: Used non-interfering assistive device	31	0.20%
G: Used braille test	0	0.00%
H: Used large-print test	53	0.35%
J: Tested over more than one day	501	3.28%
K: Had supervised breaks	1,711	11.20%
L: Most beneficial time of day	860	5.63%
M: Administered at home or in a hospital	10	0.07%
O: Examiner presented with MCE or ASL	21	0.14%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	93	0.61%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	478	3.13%
Y: Leave blank	25	0.16%
Z: Examiner read test questions aloud	2,774	18.16%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	4,071	26.65%
English Learner Test Variation A	14	0.09%
English Learner Test Variation B	7	0.05%
English Learner Test Variation C	11	0.07%
English Learner Test Variation D	1	0.01%
Any Accommodation or EL Variation	4,516	29.56%
No Accommodation or EL Variation	10,762	70.44%
English-Only Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	154	1.80%
C: Dictated responses to a scribe	9	0.11%
F: Used non-interfering assistive device	21	0.25%
G: Used braille test	0	0.00%
H: Used large-print test	29	0.34%
J: Tested over more than one day	301	3.52%
K: Had supervised breaks	901	10.54%
L: Most beneficial time of day	476	5.57%
M: Administered at home or in a hospital	7	0.08%
O: Examiner presented with MCE or ASL	10	0.12%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	54	0.63%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	259	3.03%
Y: Leave blank	14	0.16%
Z: Examiner read test questions aloud	1,453	17.00%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	2,208	25.83%

Accommodation Summary for Mathematics, Grade Six		
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	1	0.01%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	2,442	28.57%
No Accommodation or EL Variation	6,106	71.43%
Initially Fluent English Proficient (I-FEP) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	6	2.08%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	1	0.35%
G: Used braille test	0	0.00%
H: Used large-print test	4	1.39%
J: Tested over more than one day	10	3.47%
K: Had supervised breaks	28	9.72%
L: Most beneficial time of day	14	4.86%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	6	2.08%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	1	0.35%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	11	3.82%
Y: Leave blank	1	0.35%
Z: Examiner read test questions aloud	39	13.54%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	64	22.22%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	74	25.69%
No Accommodation or EL Variation	214	74.31%
English Learner (EL) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	99	1.60%
C: Dictated responses to a scribe	1	0.02%
F: Used non-interfering assistive device	8	0.13%
G: Used braille test	0	0.00%
H: Used large-print test	19	0.31%
J: Tested over more than one day	183	2.95%
K: Had supervised breaks	757	12.22%
L: Most beneficial time of day	359	5.79%
M: Administered at home or in a hospital	3	0.05%
O: Examiner presented with MCE or ASL	5	0.08%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	37	0.60%

Accommodation Summary for Mathematics, Grade Six		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	204	3.29%
Y: Leave blank	10	0.16%
Z: Examiner read test questions aloud	1,238	19.98%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,743	28.13%
English Learner Test Variation A	13	0.21%
English Learner Test Variation B	7	0.11%
English Learner Test Variation C	10	0.16%
English Learner Test Variation D	1	0.02%
Any Accommodation or EL Variation	1,934	31.21%
No Accommodation or EL Variation	4,262	68.79%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 6**	**Pct. of Total**
B: Marked in test booklet	4	1.86%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	7	3.26%
K: Had supervised breaks	20	9.30%
L: Most beneficial time of day	10	4.65%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	1	0.47%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	3	1.40%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	39	18.14%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	46	21.40%
English Learner Test Variation A	1	0.47%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	55	25.58%
No Accommodation or EL Variation	160	74.42%

Table 2.D.11 Accommodation Summary for Mathematics, Grade Seven

Accommodation Summary for Mathematics, Grade Seven		
All Tested	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	175	1.33%
C: Dictated responses to a scribe	10	0.08%
F: Used non-interfering assistive device	18	0.14%
G: Used braille test	0	0.00%
H: Used large-print test	30	0.23%
J: Tested over more than one day	273	2.08%
K: Had supervised breaks	1,080	8.22%
L: Most beneficial time of day	251	1.91%
M: Administered at home or in a hospital	13	0.10%
O: Examiner presented with MCE or ASL	19	0.14%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	21	0.16%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	275	2.09%
Y: Leave blank	14	0.11%
Z: Examiner read test questions aloud	1,199	9.12%
Accommodation is in Section 504 plan	2	0.02%
Accommodation is in IEP	2,109	16.04%
English Learner Test Variation A	10	0.08%
English Learner Test Variation B	22	0.17%
English Learner Test Variation C	8	0.06%
English Learner Test Variation D	7	0.05%
Any Accommodation or EL Variation	2,431	18.49%
No Accommodation or EL Variation	10,714	81.51%
English-Only Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	120	1.58%
C: Dictated responses to a scribe	5	0.07%
F: Used non-interfering assistive device	12	0.16%
G: Used braille test	0	0.00%
H: Used large-print test	18	0.24%
J: Tested over more than one day	181	2.38%
K: Had supervised breaks	646	8.48%
L: Most beneficial time of day	173	2.27%
M: Administered at home or in a hospital	8	0.11%
O: Examiner presented with MCE or ASL	14	0.18%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	14	0.18%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	153	2.01%
Y: Leave blank	8	0.11%
Z: Examiner read test questions aloud	674	8.85%
Accommodation is in Section 504 plan	1	0.01%
Accommodation is in IEP	1,247	16.38%

Accommodation Summary for Mathematics, Grade Seven		
English Learner Test Variation A	1	0.01%
English Learner Test Variation B	3	0.04%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	1,407	18.48%
No Accommodation or EL Variation	6,208	81.52%
Initially Fluent English Proficient (I-FEP) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	3	1.11%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	6	2.21%
K: Had supervised breaks	22	8.12%
L: Most beneficial time of day	1	0.37%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.37%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	1	0.37%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	4	1.48%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	11	4.06%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	30	11.07%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	36	13.28%
No Accommodation or EL Variation	235	86.72%
English Learner (EL) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	47	0.95%
C: Dictated responses to a scribe	5	0.10%
F: Used non-interfering assistive device	6	0.12%
G: Used braille test	0	0.00%
H: Used large-print test	11	0.22%
J: Tested over more than one day	75	1.52%
K: Had supervised breaks	397	8.03%
L: Most beneficial time of day	76	1.54%
M: Administered at home or in a hospital	4	0.08%
O: Examiner presented with MCE or ASL	4	0.08%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	3	0.06%

Accommodation Summary for Mathematics, Grade Seven		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	112	2.26%
Y: Leave blank	6	0.12%
Z: Examiner read test questions aloud	486	9.83%
Accommodation is in Section 504 plan	1	0.02%
Accommodation is in IEP	780	15.77%
English Learner Test Variation A	9	0.18%
English Learner Test Variation B	19	0.38%
English Learner Test Variation C	8	0.16%
English Learner Test Variation D	7	0.14%
Any Accommodation or EL Variation	934	18.88%
No Accommodation or EL Variation	4,012	81.12%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 7**	**Pct. of Total**
B: Marked in test booklet	4	1.47%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.37%
J: Tested over more than one day	10	3.66%
K: Had supervised breaks	15	5.49%
L: Most beneficial time of day	1	0.37%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	3	1.10%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	6	2.20%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	24	8.79%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	45	16.48%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	47	17.22%
No Accommodation or EL Variation	226	82.78%

Table 2.D.12 Accommodation Summary for Science, Grade Five

Accommodation Summary for Science, Grade Five		
All Tested	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	502	3.01%
C: Dictated responses to a scribe	19	0.11%
F: Used non-interfering assistive device	63	0.38%
G: Used braille test	1	0.01%
H: Used large-print test	50	0.30%
J: Tested over more than one day	712	4.27%
K: Had supervised breaks	2,290	13.73%
L: Most beneficial time of day	1,270	7.61%
M: Administered at home or in a hospital	14	0.08%
O: Examiner presented with MCE or ASL	32	0.19%
P: Used audio presentation (CD)	113	0.68%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	71	0.43%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	562	3.37%
Y: Leave blank	30	0.18%
Z: Examiner read test questions aloud	4,301	25.78%
Accommodation is in Section 504 plan	2	0.01%
Accommodation is in IEP	6,071	36.39%
English Learner Test Variation A	9	0.05%
English Learner Test Variation B	27	0.16%
English Learner Test Variation C	38	0.23%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	6,656	39.90%
No Accommodation or EL Variation	10,025	60.10%
English-Only Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	321	3.47%
C: Dictated responses to a scribe	14	0.15%
F: Used non-interfering assistive device	38	0.41%
G: Used braille test	0	0.00%
H: Used large-print test	35	0.38%
J: Tested over more than one day	394	4.26%
K: Had supervised breaks	1,177	12.73%
L: Most beneficial time of day	669	7.24%
M: Administered at home or in a hospital	10	0.11%
O: Examiner presented with MCE or ASL	19	0.21%
P: Used audio presentation (CD)	71	0.77%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	40	0.43%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	303	3.28%
Y: Leave blank	19	0.21%
Z: Examiner read test questions aloud	2,258	24.43%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	3,245	35.11%

Accommodation Summary for Science, Grade Five		
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	3	0.03%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	3,541	38.31%
No Accommodation or EL Variation	5,702	61.69%
Initially Fluent English Proficient (I-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	8	2.19%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	4	1.10%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	14	3.84%
K: Had supervised breaks	41	11.23%
L: Most beneficial time of day	25	6.85%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	2	0.55%
P: Used audio presentation (CD)	1	0.27%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	5	1.37%
Y: Leave blank	1	0.27%
Z: Examiner read test questions aloud	82	22.47%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	123	33.70%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	126	34.52%
No Accommodation or EL Variation	239	65.48%
English Learner (EL) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	171	2.49%
C: Dictated responses to a scribe	4	0.06%
F: Used non-interfering assistive device	21	0.31%
G: Used braille test	0	0.00%
H: Used large-print test	14	0.20%
J: Tested over more than one day	297	4.32%
K: Had supervised breaks	1,035	15.04%
L: Most beneficial time of day	566	8.23%
M: Administered at home or in a hospital	4	0.06%
O: Examiner presented with MCE or ASL	10	0.15%
P: Used audio presentation (CD)	38	0.55%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	29	0.42%

Accommodation Summary for Science, Grade Five		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	242	3.52%
Y: Leave blank	10	0.15%
Z: Examiner read test questions aloud	1,909	27.74%
Accommodation is in Section 504 plan	2	0.03%
Accommodation is in IEP	2,628	38.19%
English Learner Test Variation A	9	0.13%
English Learner Test Variation B	27	0.39%
English Learner Test Variation C	35	0.51%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	2,908	42.26%
No Accommodation or EL Variation	3,973	57.74%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 5**	**Pct. of Total**
B: Marked in test booklet	0	0.00%
C: Dictated responses to a scribe	1	0.68%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	1	0.68%
H: Used large-print test	0	0.00%
J: Tested over more than one day	6	4.05%
K: Had supervised breaks	29	19.59%
L: Most beneficial time of day	7	4.73%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.68%
P: Used audio presentation (CD)	3	2.03%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	2	1.35%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	10	6.76%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	43	29.05%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	59	39.86%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	64	43.24%
No Accommodation or EL Variation	84	56.76%

Table 2.D.13 Accommodation Summary for Science, Grade Eight

Accommodation Summary for Science, Grade Eight		
All Tested	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	146	1.28%
C: Dictated responses to a scribe	9	0.08%
F: Used non-interfering assistive device	12	0.11%
G: Used braille test	0	0.00%
H: Used large-print test	18	0.16%
J: Tested over more than one day	207	1.82%
K: Had supervised breaks	817	7.18%
L: Most beneficial time of day	240	2.11%
M: Administered at home or in a hospital	6	0.05%
O: Examiner presented with MCE or ASL	14	0.12%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	1	0.01%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	251	2.20%
Y: Leave blank	27	0.24%
Z: Examiner read test questions aloud	1,088	9.56%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,840	16.16%
English Learner Test Variation A	11	0.10%
English Learner Test Variation B	4	0.04%
English Learner Test Variation C	5	0.04%
English Learner Test Variation D	7	0.06%
Any Accommodation or EL Variation	2,066	18.15%
No Accommodation or EL Variation	9,318	81.85%
English-Only Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	94	1.42%
C: Dictated responses to a scribe	9	0.14%
F: Used non-interfering assistive device	10	0.15%
G: Used braille test	0	0.00%
H: Used large-print test	9	0.14%
J: Tested over more than one day	119	1.80%
K: Had supervised breaks	459	6.96%
L: Most beneficial time of day	126	1.91%
M: Administered at home or in a hospital	3	0.05%
O: Examiner presented with MCE or ASL	12	0.18%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	132	2.00%
Y: Leave blank	16	0.24%
Z: Examiner read test questions aloud	591	8.96%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	1,031	15.62%

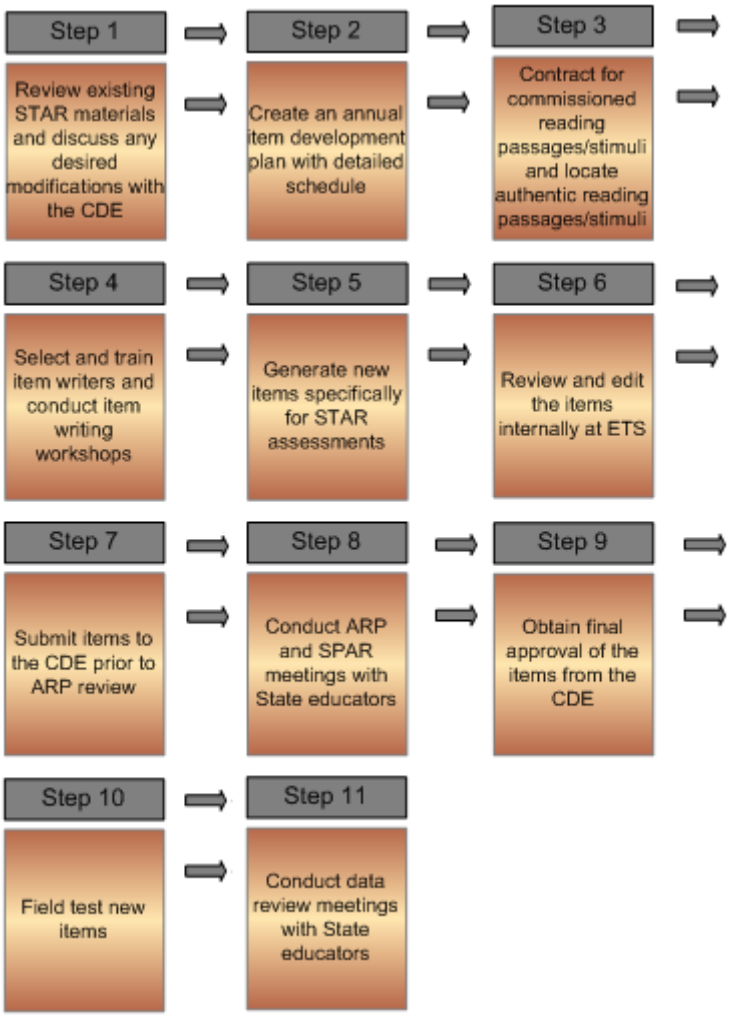
Accommodation Summary for Science, Grade Eight		
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	2	0.03%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	1,138	17.25%
No Accommodation or EL Variation	5,461	82.75%
Initially Fluent English Proficient (I-FEP) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	2	0.71%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	1	0.35%
J: Tested over more than one day	16	5.65%
K: Had supervised breaks	28	9.89%
L: Most beneficial time of day	12	4.24%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.35%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	8	2.83%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	21	7.42%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	48	16.96%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	51	18.02%
No Accommodation or EL Variation	232	81.98%
English Learner (EL) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	47	1.11%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	2	0.05%
G: Used braille test	0	0.00%
H: Used large-print test	8	0.19%
J: Tested over more than one day	67	1.58%
K: Had supervised breaks	310	7.29%
L: Most beneficial time of day	97	2.28%
M: Administered at home or in a hospital	3	0.07%
O: Examiner presented with MCE or ASL	0	0.00%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	1	0.02%

Accommodation Summary for Science, Grade Eight		
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	110	2.59%
Y: Leave blank	11	0.26%
Z: Examiner read test questions aloud	445	10.47%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	724	17.03%
English Learner Test Variation A	11	0.26%
English Learner Test Variation B	2	0.05%
English Learner Test Variation C	5	0.12%
English Learner Test Variation D	7	0.16%
Any Accommodation or EL Variation	836	19.66%
No Accommodation or EL Variation	3,416	80.34%
Reclassified Fluent English Proficient (R-FEP) Students	**Grade 8**	**Pct. of Total**
B: Marked in test booklet	3	1.35%
C: Dictated responses to a scribe	0	0.00%
F: Used non-interfering assistive device	0	0.00%
G: Used braille test	0	0.00%
H: Used large-print test	0	0.00%
J: Tested over more than one day	4	1.80%
K: Had supervised breaks	15	6.76%
L: Most beneficial time of day	4	1.80%
M: Administered at home or in a hospital	0	0.00%
O: Examiner presented with MCE or ASL	1	0.45%
P: Used audio presentation (CD)	0	0.00%
Q: Used a calculator	0	0.00%
R: Used an arithmetic table	0	0.00%
S: Used math manipulatives	0	0.00%
V: Used interfering assistive device	0	0.00%
X: Used an unlisted accommodation	1	0.45%
Y: Leave blank	0	0.00%
Z: Examiner read test questions aloud	24	10.81%
Accommodation is in Section 504 plan	0	0.00%
Accommodation is in IEP	29	13.06%
English Learner Test Variation A	0	0.00%
English Learner Test Variation B	0	0.00%
English Learner Test Variation C	0	0.00%
English Learner Test Variation D	0	0.00%
Any Accommodation or EL Variation	33	14.86%
No Accommodation or EL Variation	189	85.14%

Chapter 3: Item Development

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

Figure 3.1 Item and Test Development Cycle



Rules for Item Development

ETS maintains item development specifications for the CMA and has developed an Item Utilization Plan to guide the development of the items for each content area. Item writing emphasis is determined in consultation with the CDE.

Item Development Specifications

The item development specifications describe the characteristics of the items that should be written to measure each content standard. These specifications help ensure that all items developed for CMA measure the content standards consistently. To do this, the item specifications provide detailed information to writers who develop items for the CMA. The specifications include the following:

- A full statement of each academic content standard, as defined by the SBE (CDE, 2009)
- A description of each content strand

- The expected depth of knowledge (DOK) measured by items written for each standard (coded as 1, 2, 3, or 4; items assigned a DOK of 1 are the least cognitively complex, items assigned a DOK of 3 being the most cognitively complex, and the code of 4 only applies to some writing tasks)
- The homogeneity of the construct measured by each standard
- A description of the kinds of item stems appropriate for multiple-choice items used to assess each standard
- A description of the kinds of distractors that are appropriate for multiple-choice items assessing each standard
- A description of appropriate stimuli (such as charts, tables, graphs, or other illustrations) for mathematics and science items
- The content limits for the standard (such as one or two variables, maximum place values of numbers) for mathematics and science items
- A description of appropriate reading passages, if applicable, for ELA items
- A description of specific kinds of items to be avoided, if any (for example, items with any negative connotation in the stem, e.g. "Which of the following is NOT…")

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

- The acceptable ranges for passage length
- The expected distribution of passages by genre
- Guidelines for readability and cognitive load, using standards agreed to by the CDE and ETS
- Expected use of illustrations
- The target number of items that should follow each reading passage and each writing passage
- Writing passages for grades four and seven to have a readability level appropriate to their respective grades
- A list of topics to be avoided

Expected Item Ratio

ETS has developed the Item Utilization Plan to continue the development of CMA items. The plan includes strategies for developing items that will permit coverage of all appropriate standards for all tests in each content area and at each grade level. ETS test development staff uses this plan to determine the number of items to develop for content standard within a given subject and grade.

The Item Utilization Plan assumes that each year beginning after the first two operational administrations, 35 percent of items on an operational form will be refreshed each year; these items remain in the item bank for future use. An additional five percent of the operational items are likely to become unusable because of normal attrition. In addition, the plan notes that there is a need to focus development on what are called "critical" standards, which are standards that are difficult to measure well.

For all content areas except science, it is assumed that at least 75 percent of all field-tested items are expected to have acceptable field test statistics and become candidates for use in operational tests. For science, it is expected that 60 percent of the items will achieve this status.

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

Table 3.1 Field-test Percentages for the CMA

Content Area	Percentage of Operational Form to Be Field Tested per Grade	Number of Items to Be Field Tested per Grade
English–Language Arts, grades 3 to 5	113%	54
English–Language Arts, grades 6 to 8	67%	36
Mathematics, grades 3 to 5	113%	54
Mathematics, grades 6 and 7	67%	36
Science, grade 5	113%	54
Science, grade 8	100%	54

Selection of Item Writers

Criteria for Selecting Item Writers

The items for each CMA test are developed by item writers that have a thorough understanding of the California content standards. Applicants for item writing are screened by senior ETS content staff. Only those with strong content and teaching backgrounds are approved for inclusion in the training program for item writers. Because most of the participants are current or former California educators, they are particularly knowledgeable about the standards assessed on the CMA. All item writers meet the following minimum qualifications:

- Possession of a Bachelor's degree in the relevant content area or in the field of education with special focus on a particular content of interest; an advanced degree in the relevant content area is desirable
- Previous experience in writing items for standards-based assessments, including knowledge of the many considerations that are important when developing items to match state-specific standards
- Previous experience in writing items in the content areas covered by CMA grades and/or courses
- Familiarity, understanding, and support of the California content standards
- Current or previous teaching experience in California, when possible
- Knowledge about the abilities of the students taking these tests

Item Writer Training

ETS holds training sessions for item writers. At these sessions ETS test development specialists train attendees in the basics of item writing. They also review items that participants create during the training, offering feedback in both group and individual settings. In 2009, item writer training was conducted over two days in Newport Beach, California. Out of 50 total participants,12 were trained to write items for the CMA. An effort was made to evenly distribute the 12 participants across the three CMA content areas.

Item Review Process

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

Contractor Review

Once the items have been written, ETS employs a series of internal reviews. The reviews establish the criteria used to judge the quality of the item content and are designed to ensure that each item measures what it is intended to measure. The internal reviews also examine the overall quality of the test items before they are prepared for presentation to the CDE and the Assessment Review Panels (ARPs). Because of the complexities involved in producing defensible items for high-stakes programs such as the STAR Program, it is essential that many experienced individuals review each item before it is brought to the CDE, the ARPs, and Statewide Pupil Assessment Review (SPAR) panels.

The ETS review process for the CMA includes the following:

1. Internal content review
2. Internal editorial review
3. Internal sensitivity review

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

1. Internal Content Review

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

- Relevance of each item to the purpose of the test
- Match of each item to the item specifications, including depth of knowledge
- Match of each item to the principles of quality item writing
- Match of each item to the identified standard or standards
- Difficulty of the item
- Accuracy of the content of the item
- Readability of the item or passage
- Grade-level appropriateness of the item
- Appropriateness of any illustrations, graphs, figures, or other illustrations

Each item is classified with a code for the standard that it is intended to measure. The assessment specialists also check all items against their classification codes, both to evaluate the correctness of the classification and to ensure that the task posed by the item is relevant to the outcome it is intended to measure. The reviewers may accept the item and classification as it was written, they may suggest revisions, or they may recommend that the item be discarded. These steps occur prior to the CDE's review.

2. Internal Editorial Review

After the content-area assessment specialists review each item, a group of specially trained editors review each item in preparation for review by the CDE and the ARPs. The editors

check items for clarity, correctness of language, appropriateness of language for the grade level assessed, adherence to the style guidelines, and conformity with acceptable item-writing practices.

3. Internal Sensitivity Review
ETS assessment specialists who are specially trained to identify and eliminate questions that contain content or wording that could be construed to be offensive to or biased against members of specific ethnic, racial, or gender groups, conduct the next level of review. These trained staff members review each item before the CDE and ARP reviews.

The review process promotes a general awareness of and responsiveness to the following:

- Cultural diversity
- Diversity of background, cultural tradition, and viewpoints to be found in the test-taking populations
- Changing roles and attitudes toward various groups
- Role of language in setting and changing attitudes toward various groups
- Contributions of diverse groups (including ethnic and minority groups, individuals with disabilities, and women) to the history and culture of the United States and the achievements of individuals within these groups

Content Expert Reviews

Assessment Review Panels (ARPs)
ETS is responsible for working with ARPs as items are developed for the CMA. The ARPs are advisory panels to the CDE and ETS on areas related to item development for the CMA. The ARPs are responsible for reviewing all newly developed items for alignment to the California content standards. The ARPs also review the items for accuracy of content, clarity of phrasing, and quality. ETS provides the ARPs with the opportunity to review the items with the applicable field-test statistics and to make recommendations for the use of items in subsequent test forms. In their examination of test items, the ARPs may raise concerns related to age/grade appropriateness, and gender, racial, ethnic, and socioeconomic bias.

Composition of ARPs
The ARPs are comprised of current and former teachers, resource specialists, administrators, curricular experts, and other education professionals. Current school staff members must meet minimum qualifications to serve on the CMA ARPs, including:

- Three or more years of general teaching experience in grades kindergarten through grade twelve and in the relevant content areas (English–language arts, mathematics, or science);
- Possession of a bachelor's or higher degree in a grade or subject area related to English–language arts, mathematics, or science;
- Knowledge and experience with the California content standards for English–language arts, mathematics, or science;
- Special education credential;
- Experience with more than one type of disability; and
- Three to five years as a teacher or school administrator with a special education credential.

School administrators, district/county content/program specialists, or university educators serving on the CMA ARPs must meet the following qualifications:

- Three or more years of experience as a school administrator, district/county content/ program specialist, or university instructor in a grade-specific area or area related to English–language arts, mathematics, or science
- Bachelor's or higher degree in a grade-specific or subject area related to English–language arts, mathematics, or science
- Knowledge of and experience with the California content standards in English–language arts, mathematics, or science

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

Current ARP members are recruited through an application process. Recommendations are solicited from school districts and county offices of education as well as from CDE and SBE staff. Applications are received and reviewed throughout the year. They are reviewed by the ETS assessment directors, who confirm that the applicant's qualifications meet the specified criteria. Applications that meet the criteria are forwarded to CDE and SBE staff for further review and final approval. Upon approval, the applicant is notified that he or she has been selected to serve on the ARP committee.

Table 3.2 shows the educational qualifications, present occupation, and credentials of the current CMA ARP members.

Table 3.2 CMA ARP Member Qualifications, by Subject and Total

CMA	ELA	Math	Science	Total
Total	**19**	**9**	**11**	**39**
Occupation (Members may teach multiple levels.)				
Teacher or Program Specialist, Elementary/Middle School	14	5	5	24
Teacher or Program Specialist, High School	1	2	2	5
Teacher or Program Specialist, K–12	1	0	1	2
University Personnel	0	1	1	2
Other District Personnel (e.g., Director of Special Services, etc.)	4	2	1	7
Highest Degree Earned				
Bachelor's Degree	5	2	3	10
Master's Degree	12	7	7	26
Doctorate	0	0	1	1
Credential (Members may hold multiple credentials.)				
Elementary Teaching (multiple subjects)	10	8	2	20
Secondary Teaching (single subject)	2	0	7	9
Special Education	4	7	4	15
Reading Specialist	4	3	0	7
English Learner (CLAD, BCLAD)	4	5	1	10
Administrative	7	1	0	8
Other	3	0	1	4
None (teaching at the university level)	0	0	0	0

ARP Meetings for Review of CMA Items

The ETS content-area assessment specialists facilitate the CMA ARP meetings. Each meeting begins with a brief training session on how to review items. ETS provides this training, which consists of the following topics:

- Overview of the purpose and scope of the CMA
- Overview of the CMA test design specifications and blueprints
- Analysis of the CMA item specifications
- Overview of criteria for evaluating multiple-choice test items and for reviewing constructed-response writing tasks
- Review and evaluation of items for bias and sensitivity issues

The criteria for evaluating multiple-choice items and constructed-response writing tasks included:

- Overall technical quality
- Match to the California content standards
- Match to the construct being assessed by the standard
- Difficulty range
- Clarity
- Correctness of the answer
- Plausibility of the distracters
- Bias and sensitivity factors

Criteria also include more global issues, including—for ELA—the appropriateness, difficulty, and readability of reading passages. The ARPs also are trained on how to make recommendations on item revision. Guidelines for reviewing items are provided by ETS and approved by the CDE. The set of guidelines for reviewing items is summarized below.

Does the item:

- Have one and only one clearly correct answer?
- Measure the content standard?
- Match the test item specifications?
- Align with the construct being measured?
- Test worthwhile concepts or information?
- Reflect good and current teaching practices?
- Have a stem that gives the student a full sense of what the item is asking?
- Avoid unnecessary wordiness?
- Use response options that relate to the stem in the same way?
- Use response options that are plausible and have reasonable misconceptions and errors?
- Avoid having one response option that is markedly different from the others?
- Avoid clues to students, such as absolutes or words repeated in both the stem and options?
- Reflect content that is free of bias against any person or group?

Is the stimulus, if any, for the item:

- Required in order to answer the item?
- Likely to be interesting to students?

- Clearly and correctly labeled?
- Providing all the information needed to answer the item?

As the first step of the item review process, ARP members review a set of items independently and record their individual comments. The next step in the review process is for the group to discuss each item. The content-area assessment specialists facilitate the discussion and record all recommendations. Those recommendations are recorded in a master item review booklet. Item review binders and other item evaluation materials also identify potential bias and sensitivity factors the ARP will consider as a part of its item reviews.

Depending on CDE approval and the numbers of items still to be reviewed, some ARPs are divided further into smaller groups. The science ARP, for example, divides into content area and grade-level groups. These smaller groups are also facilitated by the content-area assessment specialists.

ETS staff maintains the minutes summarizing the review process, and then forwards copies of the minutes to the CDE, emphasizing in particular the recommendations of the panel members.

Statewide Pupil Assessment Review (SPAR) Panel

The SPAR panel is responsible for reviewing and approving a single achievement test to be used statewide for testing students in California public schools, grades two through eleven. At the SPAR panel meetings, all new items are presented in binders for review. The SPAR panel representatives ensure that the test items conform to the requirements of *Education Code* Section 60614. The constructed-response writing tasks are also presented for review. If the SPAR panel rejects specific items and/or constructed-response writing tasks, the items and/or tasks are replaced with other items and/or tasks. For the SPAR panel meeting, the item development coordinator or an ETS content specialist, requested in advance by the CDE, attends the opening session and remains in a nearby location or near a telephone to be available to respond to any questions during the course of the meeting.

Field Testing

The primary purpose of field testing is to obtain information about item performance and to obtain statistics that can be used to assemble operational forms.

Stand-Alone Field Testing

For each new CMA launched, a pool of items is initially constructed by administering the newly developed items in a stand-alone field test. In a stand-alone field testing, examinees are recruited to take tests outside of the usual testing circumstances, and the test results are typically not used for instructional or accountability purposes (Schmeiser & Welch, 2006).

ETS field tested writing prompts in the fall of 2008 to identify prompts that could be used in the operational ELA tests for grades four and seven. Six prompts were field-tested in each of the two grades. Following reviews of field test results by the CDE, ETS, and the ELA ARP, two prompts for grade four and two prompts for grade seven were selected for use in subsequent operational administrations. Two different writing prompts from the same genre are used in a given year. One prompt is taken by the majority of test takers, the other is administered in programs not in session during the first administration.

Items developed for ELA in grades six to eight, mathematics in grades six through seven and science in grade eight were field-tested in a stand-alone format in the fall of 2008. All forms were built to meet content specifications as outlined in the test blueprint. Each field-test form contained a set of common items embedded in each form; these items were used to link the forms. Following reviews of field test results by the CDE, ETS, and the ARP, the items were selected for use in future operational forms.

Embedded Field-test Items

Although a stand-alone field test is useful for developing a new test because it can produce a large pool of quality items, an embedded field testing is generally preferred because the items being field tested are scattered throughout the operational test. Variables such as test-taker motivation and test security are the same in embedded field testing as they will be when the field-tested items are later administered operationally. Such field testing involves distributing the items being field-tested within each operational test form. Different forms will contain the same operational items and different field test items.

Allocation of Students to Field-test Items

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

Number of Forms and Sample Sizes

Nine field-test items were embedded in each CMA operational form administered in 2009; the sets of nine field-test items differ between forms. The number of forms varied with subject area and grade level.

Table 3.3 shows the number of forms administered for each CMA in 2009 and the numbers of examinees included in samples used for the field test or “final” item analyses (FIA) of these forms. The field-test samples constitute approximately 80 to 90 percent of the entire population tested.

Table 3.3 Summary of Items and Forms Presented in 2009 CMA

		Operational			Field Test	
Subject	**CMA**	**No. Items**	**No. Examinees FIA Sample**	**No. Forms**	**No. Items**	**No. Examinees FIA Sample**
English–Language Arts	**3***	48	11,321	6	9	1,815–2,100
	4**	48	17,121	5	9	2,847–4,011
	5	48	17,821	6	9	2,858–3,256
	6	54	15,900	4	9	3,934–4,033
	7**	54	13,060	4	9	3,252–3,284
	8	54	12,434	6	9	2,028–2,102
* Standard form ** Multiple choice only						
Mathematics	**3**	48	11,227	6	9	1,802–2,090
	4	48	14,415	5	9	2,407–3,400
	5	48	15,656	6	9	2,532–2,866
	6	54	15,278	4	9	3,784–3,890
	7	54	13,145	4	9	3,260–3,315
Science	**5**	48	16,681	6	9	2,679–3,048
	8	54	11,384	6	9	1,876–1,920

Data Review Meetings

Once items have been field tested, ETS prepares the items and the associated statistics for review by the ARPs. ETS assessment specialists facilitate the data review sessions with qualified psychometric staff for technical assistance. Upon completion of the meeting, ETS provides the SBE and the CDE with summaries of the recommendations based on the field-test analyses and ARP reviews that are relevant to future form construction of the CMA. All final decisions on acceptance of items rest with the CDE in consultation with the SBE.

At data review meetings, the ARP members discuss items that have poor statistics and do not meet the psychometric criteria for item quality. The CDE defines the criteria for acceptable or unacceptable item statistics. These criteria ensure that the item (1) has an appropriate level of difficulty for the target population, (2) discriminates well between examinees that differ in ability, and (3) conforms well to the statistical model underlying the measurement of the intended constructs. The panel members also use the results of analyses for differential item functioning (DIF) to make judgments about the appropriateness of items for various subgroups. The panelists respond to questions such as:

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

The panelists recommend whether to accept or reject each item for inclusion in the STAR item bank.

Item Banking

Once the ARP review is complete, the items are placed in the item bank along with their corresponding review information. Items that are accepted by the ARP are updated to a “field-test ready” status; items that are rejected are updated to a “rejected before use” status. ETS then delivers the items to the CDE by means of a delivery of the STAR electronic item bank. Subsequent updates to items are based on field-test and operational use of the items. However, only the latest content of the item is in the bank at any time, along with the administration data from every administration that has included the item.

After field-test or operational use, items may be rejected if they do not meet statistical specifications; such items are updated with a status of “rejected for statistical reasons” and remain unavailable in the bank. These statistics are obtained by the research group at ETS who carefully evaluates each item for its level of difficulty and discrimination as well as conformance to the IRT Rasch model. Researchers also determine if the item functions similarly for various subgroups of interest.

All unavailable items are clearly marked with an availability indicator of “Unavailable,” a reason for rejection reason as described above, and cause alerts so they are not inadvertently included in subsequent test forms. Status and availability of an item are updated programmatically as items are presented for review, accepted or rejected, placed on a form for field testing, presented for statistical review, used operationally, or released. All rejection and release indications are monitored and controlled through ETS’s assessment development processes.

References

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

Educational Testing Service (2002). *ETS standards for quality and fairness*. Office of Testing Integrity, Educational Testing Service, Princeton, NJ.

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

Schmeiser, C.B., & Welch, C.J. (2006). Test development. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport, CT: American Council on Education and Praeger Publishers.

Chapter 4: Test Assembly

The CMA is constructed to measure students' performance relative to California's content standards approved by the SBE. The CMA tests are also constructed to meet professional standards for validity and reliability. For each CMA, the content standards and psychometric attributes are used as the basis for assembling the test forms.

Test Length

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

There are 48 items in the CMA for English–language arts (ELA) for grades three through five, in the mathematics tests for grades three through five, and in the science test for grade five. There are 54 items in the ELA tests for grades six through eight, in the mathematics tests for grades six and seven, and in the science test for grade eight. For more details on the distribution of items, see Appendix 2.A on page 17.

Rules for Item Selection

Test Blueprint

ETS selects all CMA test items to conform to the SBE-approved California content standards and test blueprints. The content blueprints for the CMA can be found on the CDE "STAR CMA Blueprints" Web page at http://www.cde.ca.gov/ta/tg/sr/cmablueprints.asp.

Although the test blueprints call for the number of items at the individual standard level, scores on the CMA items are grouped into subcontent areas, referred to as "reporting clusters." For each CMA reporting cluster, the percentage of questions correctly answered is reported on a student's score report. A description of the CMA reporting clusters and the standards that compose each reporting cluster is provided in Appendix 2.B, which starts on page 18.

Content Rules and Item Selection

When developing a new test for a given grade and content area, test developers follow a number of rules. First and foremost, they select items that meet the blueprint for that grade and content area. Using an electronic item bank, assessment specialists begin by identifying a number of linking items. These are items that appeared in the previous year's operational administration and they are used to equate the test forms administered each year. Linking items are selected to proportionally represent the full blueprint. For example, if 25 percent of all of the items in a test are in the first reporting cluster, then 25 percent of the linking items should come from that cluster. The linking items are selected for their strong match to the content and are reviewed to ensure that they meet set psychometric criteria.

After the linking items are approved, assessment specialists populate the rest of the test form. Their first consideration is the strength of the content and the match of each item to the standard. In selecting items, team members also try to ensure that they include a variety of formats and content and that at least some of the items include graphics for visual interest.

Another consideration is the difficulty of each item. Test developers strive to ensure that there are some easy and some hard items and that there are a number of items in the middle range of difficulty. If items do not meet all content and psychometric criteria, staff reviews the other available items to determine if there are other selections that could improve the match of the test to all of the requirements. If such match is not attainable, the content team works in conjunction with psychometricians and the CDE to determine which combination of items will best serve the needs of the students taking the test. Chapter 3 on page 60 contains further information about this process.

Psychometric Criteria

For the CMA, the test developers and psychometricians strive to accomplish three goals while developing a test:

1. The test must have desired precision of measurement at all ability levels.
2. The test score must be valid and reliable for the intended population and for the various subgroups of test takers.
3. The test forms must be comparable across years of administration to ensure the generalizability of scores over time.

In order to achieve these goals, a set of rules is developed that outlines the desired psychometric properties of each CMA. Such rules are referred to as statistical targets. Three types of assembly targets are developed for each CMA: the total test target, the linking block target, and (reporting) cluster targets. These targets are provided to test developers before a test construction cycle begins.

The total test target or primary statistical targets used for assembling CMA for the 2009 STAR administration were the test information function based on the item response theory (IRT) item parameters and an average point biserial correlation. When using the IRT Rasch model, the target information function makes it possible to choose items to produce a test that has the desired precision of measurement at all ability levels. The target mean and standard deviation of item difficulty (*b*-values) consistent with the information curves were also provided to test development staff to help with the test construction process. The point biserial correlation is a measure of how well the items discriminate among test takers that differ in their ability, and it is related to the overall reliability of the test.

The minimum target value for an item point-biserial was set at 0.14 for each test. This value approximates a biserial correlation of 0.20. The target *b*-value range approximates a minimum proportion-correct value (*p*-value) of 0.33 and a maximum *p*-value of 0.95 for each test.

These specifications were developed from the analyses of field-test forms administered in fall 2007 for ELA and mathematics in grades three through five and for science in grade five. They were used to develop the spring 2008 operational forms, which were re-used in the spring of 2009. The target values for ELA in grades six through eight, mathematics in grades six and seven, and science in grade eight were developed using data from a fall 2008 field test. These target values were used to develop the spring 2009 operational ELA, mathematics, and science forms for these grades. The target values for all grades and subject areas are presented in Table 4.1, on the next page.

Table 4.1 Target Statistical Specifications for the CMA

Subject	Grade	Point Biserial		B-Values		P-value	
		Mean	Minimum	Mean	S.D.	Minimum	Maximum
English–Language Arts	3	.42–.48	0.14	–0.47	0.50	0.33	0.95
	4	.42–.48	0.14	–0.43	0.50	0.33	0.95
	5	.42–.48	0.14	–0.44	0.50	0.33	0.95
	6	.42–.48	0.14	–0.40	0.50	0.33	0.95
	7	.42–.48	0.14	–0.40	0.50	0.33	0.95
	8	.42–.48	0.14	–0.40	0.50	0.33	0.95
Mathematics	3	.42–.48	0.14	–0.45	0.50	0.33	0.95
	4	.42–.48	0.14	–0.44	0.50	0.33	0.95
	5	.42–.48	0.14	–0.47	0.50	0.33	0.95
	6	.42–.48	0.14	–0.44	0.50	0.33	0.95
	7	.42–.48	0.14	–0.44	0.50	0.33	0.95
Science	5	42–.48	0.14	–0.42	0.50	0.33	0.95
	8	42–.48	0.14	–0.48	0.50	0.33	0.95

Rules for Item Sequence and Layout

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

- **English–Language Arts—**Because the ELA test is primarily passage-dependent, items are sequenced with their associated reading passages. Passages are sequenced according to genre and interest level; test developers work to place high-interest pieces (typically narrative selections) near lower-interest pieces (typically functional or technical writing). Stand-alone items are placed throughout the form, where appropriate.
- **Mathematics—**The items on the CMA mathematics test form are grouped by reporting category. All Number Sense items appear first, followed by Algebra and Functions items, Measurement and Geometry items, and lastly, by Statistics, Data Analysis, and Probability items. These reporting categories make up the three reporting clusters for the mathematics CMA tests. The Number Sense cluster consists of all items solely from the Number Sense reporting category. The Algebra and Data Analysis cluster consists of all items from both the Algebra and Functions reporting category as well as the Statistics, Data Analysis, and Probability reporting clusters. The final cluster, Measurement and Geometry, consists of all of the items in the Measurement and Geometry reporting category.
- **Science—**The science test for grade five is sequenced according to reporting cluster; that is, all items from a single reporting cluster are presented together and then all of the items from the next reporting cluster are presented. For the grade eight science test, the test is sequenced by contents standard and does not yet have associated reporting clusters.

Chapter 5: Test Administration

Test Security and Confidentiality

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

ETS's Office of Testing Integrity (OTI)

The OTI is a division of ETS that provides quality assurance services for all testing programs administered by ETS and resides in the ETS legal department. The Quality Assurance division of ETS publishes and maintains *ETS Standards for Quality and Fairness,* which supports the OTI's goals and activities. The purposes of the *ETS Standards for Quality and Fairness* are to help ETS design, develop, and deliver technically sound, fair, and useful products and services, and to help the public and auditors evaluate those products and services.

The OTI's mission is to

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

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

Test Development

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

- Only authorized individuals have access to test content at any step in the test development, item review, and data analysis processes.
- Test developers keep all hard copy test content, computer disk copies, art, film, proofs, and plates in locked storage when not in use.
- ETS shreds working copies of secure content as soon as they are no longer needed in the test development process.
- Test developers take further security measures when test materials are to be shared outside of ETS; this is achieved by using registered and/or secure mail, using express delivery methods, and actively tracking records of dispatch and receipt of the materials.

Item Review

ETS enforces security measures at ARP meetings to protect the integrity of meeting materials using the following guidelines:

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

- Meeting materials are strictly managed before, during, and after the review meetings.
- Meeting participants are supervised at all times during the meetings.
- Use of electronic devices is strictly prohibited in the meeting rooms.

Item Bank

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

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

- Electronic forms of test content, documentation, and item banks are backed up, and the backups are kept offsite.
- The offsite backup files are kept in secure storage with access limited to authorized personnel only.
- To prevent unauthorized electronic access to the item bank, state-of-the-art network security measures are used.

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

Transfer of Forms and Items to the CDE

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

Firewall

A firewall is software that prevents unauthorized entry to files, e-mail, and other organization-specific programs. ETS data exchange and internal e-mail remain within the ETS firewall at all ETS locations, ranging from Princeton, New Jersey, to San Antonio, Texas, to Sacramento, California.

All electronic applications included in the STAR Management System (CDE, 2009a) remain protected by the ETS firewall software at all times. Due to the sensitive nature of the student information processed by the STAR Management System, the firewall plays a significant role in maintaining an assurance of confidentiality in the users of this information. (It should be noted that the STAR Management System neither stores nor processes tests or student test results.)

Printing

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

Test Administration

Pearson receives testing materials from printers, packages them, and sends them to school districts. After testing, the school districts return materials to Pearson for scoring. During these events, Pearson takes extraordinary measures to protect the testing materials. Pearson's customized Oracle business applications verify that inventory controls are in place from receipt of materials to packaging. The reputable carriers used by Pearson provide a specialized handling and delivery service that maintains test security and meets the CMA program schedule. The carriers provide inside delivery directly to the district STAR coordinators or authorized recipients of the assessment materials.

Test Delivery

Test security requires accounting for all secure materials before, during, and after each test administration. The district STAR coordinators are, therefore, required to keep all test materials in central locked storage except during actual test administration times. Test site coordinators are responsible for accounting for and returning all secure materials to the district STAR coordinator, who is responsible for returning them to the STAR Scoring and Processing Center. The following measures are in place to ensure security of STAR testing materials:

- District STAR coordinators are required to sign and submit a "STAR Test (including field tests) Security Agreement for District and Test Site Coordinators" form to the STAR Technical Assistance Center before ETS can ship any testing materials to the district.
- Test site coordinators have to sign and submit a "STAR Test (including field tests) Security Agreement for District and Test Site Coordinators" form to the district STAR coordinator before any testing materials can be delivered to the school/test site.
- Anyone having access to the test materials must sign and submit a "STAR Test (including field tests) Security Affidavit for Test Examiners, Proctors, Scribes, and Any Other Person Having Access to STAR Tests" form to the test site coordinator before receiving access to any testing materials.
- It is the responsibility of each person participating in the STAR Program to report immediately any violation or suspected violation of test security or confidentiality. The test site coordinator is responsible for immediately reporting any security violation to the district STAR coordinator. The district STAR coordinator must contact the CDE immediately; the coordinator will be asked to follow up with a written explanation of the violation or suspected violation.

Processing and Scoring

An environment that promotes the security of the test prompts, student responses, data, and employees throughout a project is of utmost concern to Pearson. Pearson requires the following standard safeguards for security at their sites:

- There is controlled access to the facility.
- No test materials may leave the facility during the project without the permission of a person or persons designated by the CDE.
- All scoring personnel must sign a nondisclosure and confidentiality form in which they agree not to use or divulge any information concerning tests, scoring guides, or individual student responses.
- All staff must wear Pearson identification badges at all times in Pearson facilities.

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

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

All answer documents and test booklets are destroyed after October 31 each year.

Data Management

Pearson provides overall security for assessment materials through its limited-access facilities and through its secure data processing capabilities. Pearson enforces stringent procedures to prevent unauthorized attempts to access their facilities. Entrances are monitored by security personnel and a computerized badge-reading system is utilized. Upon entering the facilities, all Pearson employees are required to display identification badges that must be worn at all times while in the facility. Visitors must sign in and out. While they are at the facility, they are assigned a visitor badge and escorted by Pearson personnel. Access to the Data Center is further controlled by the computerized badge-reading system that allows entrance only to those employees who possess the proper authorization.

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

For mainframe processes, Pearson utilizes Random Access Control Facility (RACF) to limit and control access to all data files (test and production), source code, object code, databases, and tables. RACF controls who is authorized to alter, update, or even read the files. All attempts to access files on the mainframe by unauthorized users are logged and monitored. In addition, Pearson uses ChangeMan, a mainframe configuration management tool, to control versions of the software and data files. ChangeMan provides another level of security, combined with RACF, to place the correct tested version of code into production. Unapproved changes are not implemented without prior review and approval.

Transfer of Scores via Secure Data Exchange

After scoring is completed, Pearson sends scored data files to ETS and follows secure data exchange procedures. ETS and Pearson have implemented procedures and systems to provide efficient coordination of secure data exchange. This includes the established SFTP site that is used for secure data transfers between ETS and Pearson. These well-established procedures provide timely, efficient, and secure transfer of data. Access to the STAR data files is limited to appropriate personnel with direct project responsibilities.

Statistical Analysis

The Scoring, Reporting and Technology (SR&T) area at ETS retrieves the Pearson data files from the SFTP site and loads them in a database. The Data Quality Services (DQS) area at ETS extracts the data from the database and performs quality control procedures before passing files to the ETS Statistical Analysis group. The Statistical Analysis group then keeps the files on secure servers and adheres to the ETS Code of Ethics to prevent any unauthorized access.

Reporting and Posting Results

After statistical analysis has been completed on student data, the files flow in three different directions. Paper reports, some with individual student results and others with summary results, are produced. Encrypted files of summary results are also sent to the CDE by means of SFTP. Any summary results that have fewer than ten students are not reported. The statistics based on the results are also entered into the item bank.

Student Confidentiality

To meet ESEA and state requirements, school districts must collect demographic data about students. This includes information about students' ethnicity, parent education, disabilities, whether the student qualifies for the National School Lunch Program (NSLP), and so forth (CDE, 2009b). In addition, students may reveal other information about themselves through the essays they write. ETS takes precautions to prevent any of this information from becoming public or being used for anything other than testing purposes. These procedures are applied to all documents where this demographic information may appear, including the following:

- Pre-ID files
- Reports
- Essays

Test Results

ETS also has security measures for files and reports that show students' scores and performance levels. ETS is committed to safeguarding the information in its possession from unauthorized access, disclosure, modification, or destruction. ETS has strict information security policies in order to protect the confidentiality of ETS and client data. ETS staff's access to production databases is limited to personnel with a business need to access the data. User IDs for production systems must be person-specific or for systems use only.

ETS has implemented network controls for routers, gateways, switches, firewalls, network tier management, and network connectivity. Routers, gateways, and switches represent points of access between networks. However, these do not contain mass storage or represent points of vulnerability, particularly to unauthorized access or denial of service. Routers, switches, firewalls, and gateways may possess little in the way of logical access.

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

Access to the ETS Computer Processing Center is controlled by employee and visitor identification badges. The Center is secured by doors that can only be unlocked by the badges of personnel who have functional responsibilities within its secure perimeter. Authorized personnel accompany visitors to the Data Center at all times. Extensive smoke detection and alarm systems, as well as a pre-action fire-control system, are installed in the Center.

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

- Scoring
- Transfer of scores by means of secure data exchange
- Reporting
- Analysis and reporting of erasure marks
- Internet postings
- Storage

In addition to protecting the confidentiality of testing materials, ETS's Code of Ethics further prohibits ETS employees from financial misuse, conflicts of interest, and unauthorized appropriation of ETS's property and resources. Specific rules are also given to ETS employees and their immediate families who may take a test developed by ETS, such as a STAR examination. The ETS Office of Testing Integrity verifies that these standards are followed throughout ETS. It does this, in part, by conducting periodic onsite security audits of departments, with followup reports containing recommendations for improvement.

Procedures to Maintain Standardization

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

Test Administrators

The CMA is administered in conjunction with the other tests that comprise the STAR Program. In that respect, ETS employs personnel who facilitate various processes involved in the standardization of an administration cycle.

The responsibilities for district and test site staff members are included in the *STAR District and Test Site Coordinator Manua*l (CDE, 2009c). This manual is described in the next section.

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

District STAR Coordinator
Each local education agency[1] (LEA) designates a district STAR coordinator who is responsible for ensuring the proper and consistent administration of the STAR tests. They are also responsible for securing testing materials upon receipt, distributing testing materials to schools, tracking the materials, training and answering questions from district staff and test site coordinators, receiving scorable and nonscorable materials from schools after an administration, and returning the materials to the STAR contractor for processing.

Test Examiner
CMA tests are administered by test examiners who may be assisted by test proctors and scribes. A test examiner is an employee of a school district or an employee of a nonpublic, nonsectarian school (NPS) who has been trained to administer the tests and has signed a STAR Test Security Affidavit. Test examiners must follow the directions in the *California Modified Assessment Directions for Administration* (*DFA*) (CDE, 2009d) exactly.

Test Proctor
A test proctor is an employee of the school district or a person, assigned by an NPS to implement the IEP of a student, who has received training designed to prepare him or her to assist the test examiner in the administration of tests within the STAR Program (*5 CCR* Section 850 [r]). Test proctors must sign STAR Test Security Affidavits (*5 CCR* Section 859 [c]).

Scribe
A scribe is an employee of the school district or a person, assigned by an NPS to implement the IEP of a student, who is required to transcribe a student's responses to the format required by the test. A student's parent or guardian is not eligible to serve as a scribe (*5 CCR* Section 850 [m]). Scribes must sign STAR Test Security Affidavits (*5 CCR* Section 859 [c]).

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

District and Test Site Coordinator Manual
Test administration procedures are to be followed exactly so that all students have an equal opportunity to demonstrate their academic achievement. The *STAR District and Test Site Coordinator Manual* contributes to this goal by providing information about the responsibilities of district and test site coordinators, as well as those of the other staff involved in the administration cycle. However, the manual is not intended as a substitute for the *California Code of Regulations*, *Title 5,* Education (*5 CCR*) or to detail all of the coordinator's responsibilities.

STAR Management System Manuals
The STAR Management System is a series of secure, Web-based modules that allow district STAR coordinators to set up test administrations, order materials, and submit and correct student Pre-ID data. Every module has its own user manual with detailed instructions on how to use the STAR Management System. The modules of the STAR Management System are as follows:

---

[1] Local education agencies include public school districts, statewide benefit charter schools, state board-authorized charter schools, county of education programs, and charter schools testing independently from their home district.

- **Test Administration Setup—**This module allows school districts to determine and calculate dates for scheduling the test administration for school districts, to verify contact information of those school districts, and to update the school district's shipping information. (CDE, 2009e)
- **Order Management—**This module allows school districts to enter quantities of testing materials for schools. Its manual includes guidelines for determining which materials to order. (CDE, 2009f)
- **Pre-ID—**This module allows school districts to enter or upload student information including demographics and to identify the test(s) the student will take. This information is printed on student test booklets or answer documents or on labels that can be affixed to test booklets or answer documents. Its manual includes the CDE's Pre-ID layout. (CDE, 2009b)
- **Extended Data Corrections—**This module allows school districts to correct the data that were submitted during Pre-ID up to seven days prior to the end of the school district's selected testing window. (CDE, 2009g)

Test Booklets

For each grade-level and end-of-course test, multiple versions of test booklets are administered. The versions differ only in terms of the field-test items they contain. These versions are spiraled and packaged, consecutively and are distributed at the student level—that is, each classroom or group of test takers receives at least one of each version of the test.

The test booklets, along with answer documents and other supporting materials, are packaged by school or group, depending on how the district STAR coordinator ordered the materials. All materials are sent to the district STAR coordinator for proper distribution within the LEA. Special formats of test booklets are also available for test takers who require accommodations to participate in testing. These special formats include audio CDs (grades three through five only), large-print test materials, and braille test materials (grades three through five only).

Accommodations for Students with Disabilities

All students participate in the STAR Program, including students with disabilities and English learners. ETS policy states that reasonable testing accommodations be provided to candidates with documented disabilities that are identified in the Americans with Disabilities Act (ADA). The ADA mandates that test accommodations be individualized, meaning that no single type of test accommodation may be adequate or appropriate for all individuals with any given type of disability. ADA authorizes that test takers with disabilities may be tested under standard conditions if ETS determines that only minor adjustments to the testing environment are required (e.g., wheelchair access, large-print test book, a sign language interpreter for spoken directions.)

The purpose of test variations and accommodations is to enable students to take the CMA, not to give them an advantage over other students or to improve their scores.

Identification

Most students with disabilities and English learners take the CMA under standard conditions. However, some students with disabilities and English learners may need assistance when taking the CMA; the assistance takes the form of test variations or accommodations (see Appendix 2.D in Chapter 2 for details). During the test, these students may use the special services specified in their IEP or Section 504 plan. If students

use accommodations for the CMA, test examiners are responsible for marking that event on the students' test booklets, answer documents, or Writing Response Booklet. Because the CMA was developed with modifications built into the test, modifications are not allowed. Students who require additional modifications take the content-area CST with modifications.

In the event that a student injures a hand or arm prior to the writing test, is willing and able to sit for the examination, but unable to write, the school completes a Section 504 plan for the student. The Section 504 plan identifies which accommodations the student will use in completion of his or her writing test (CDE, 2009b).

Scoring

In November 2008, the SBE approved performance levels for the CMA in grades three through five. CMA results for these grade levels are included in the API and AYP beginning with the 2008 Base. The addition of CMA into the API does not change the API test weights; the same test weights and calculation rules used for the CST also apply to the CMA

Demographic Data Corrections

After reviewing student data, some school districts may discover demographic data that are incorrect. The Demographics Data Corrections module of the STAR Management System gives school districts the means to correct these data within a specified availability window. Districts may correct data to: (1) Have the school district's API/AYP recalculated; (2) Rescore uncoded or miscoded CST end-of-course mathematics and/or science tests; (3) Obtain a corrected data CD-ROM for school district records; or (4) Match unmatched records for grades four and seven writing and multiple-choice tests (CDE, 2009h).

Testing Irregularities

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

The district STAR coordinator is responsible for immediately notifying the CDE of any irregularities that occur before, during, or after testing. The test examiner is responsible for immediately notifying the district STAR coordinator of any security breaches or testing irregularities that occur in the administration of the test. Once the district STAR coordinator and the CDE have determined that an irregularity has occurred, the CDE instructs the district STAR coordinator on how and when to identify the irregularity on the student test booklet or answer document. The information and procedures to assist in identifying irregularities and notifying the CDE are provided in the *STAR District and Test Site Coordinator Manual* (CDE, 2009c).

Test Administration Incidents

A test administration incident is any event that occurs before, during, or after test administrations that does not conform to the instructions stated in the *DFAs* and the *STAR District and Test Site Coordinator Manual* (CDE, 2009c). These events include test administration errors, disruptions, and student cheating. Test administration incidents generally do not affect test results. These administration incidents are not reported to the CDE or the STAR Program testing contractor. The STAR test site coordinator should immediately notify the district STAR coordinator of any test administration incidents that occur. It is recommended by the CDE that districts and schools maintain records of these incidents.

References

California Department of Education. (2009a). *2009 STAR Management System.* http://www.startest.org/sms.html.

California Department of Education. (2009b). *2009 STAR Pre-ID instructions manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.pre-id_manual.2009.pdf.

California Department of Education. (2009c). *2009 STAR district and test site coordinator manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.coord_man.2009.pdf.

California Department of Education. (2009d). *2009 California Modified Assessment directions for administration.* Sacramento, CA. http://www.startest.org/pdfs/CMA.grade-5_dfa.2009.pdf.

California Department of Education. (2009e). *2009 STAR Test Administration Setup manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.test_admin_setup.2009.pdf.

California Department of Education. (2009f). *2009 STAR Order Management manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.order_mgmt.2009.pdf.

California Department of Education. (2009g). *2009 STAR Extended Data Corrections manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.xdc_manual.2009.pdf.

California Department of Education. (2009h). *2009 STAR Demographic Data Corrections manual*. Sacramento, CA. http://www.startest.org/pdfs/STAR.data_corrections_manual.2009.pdf.

Chapter 6: Performance Standards

Background

The CMA tests for ELA and mathematics for grades three through five and science for grade five became part of the STAR Program in spring 2008. Five performance standards were developed in September and October 2008 and adopted by the SBE for the 2008 administration of those tests.

In spring 2009, the CMA tests for ELA in grades six through eight, mathematics in grades six and seven, and science in grade eight were introduced. The performance standards for those tests were adopted in the fall of 2009 and will be reported beginning with the 2010 operational administration.

The CMA tests for high school phase 1 (ELA for grade nine, Life Science for grade ten, and end-of-course Algebra I) will be introduced in spring 2010. Performance standards for high school phase 1 will be established for these tests in fall 2010 and reported operationally starting in the 2011 operational administration. The CMA tests for high school phase 2 (ELA for grades ten and eleven and end-of-course Geometry) will be introduced in spring 2011. Performance standards for high school phase 2 will be established in fall 2011 and reported starting in the 2012 operational administration.

The performance standards for the CMA were defined by the SBE as (1) far below basic, (2) below basic, (3) basic, (4) proficient, and (5) advanced. Performance standards are developed from a general description of the performance standard (policy level descriptors) and competencies lists, which operationally define each level. Cut scores numerically define the performance standards. The state target is to have all student achieve proficient or above by 2014.

ETS employed carefully designed standard-setting procedures to facilitate the development of performance standards for each CMA test. All standard settings were conducted using the Bookmark method and one form of each test. One panel of educators was convened to set the standards for each test. After the standard-setting, ETS met with representatives of the CDE to review the preliminary results and provided to the CDE an executive summary of the procedure and tables that showed the panel-recommended cut scores and impact data. The final cut scores were approved by the SBE in November 2008. Please see the technical report for the standard-setting (ETS, 2009) to obtain more information.

This chapter describes the process of developing performance standards to the CMA operational tests grades three through five.

Standard Setting Procedure

The process of standard setting is designed to identify a “cut score” or minimum test score that is required to qualify a student for each performance standard. The process generally requires that a panel of subject matter experts and others with relevant perspectives (such as teachers, school administrators) be assembled. The panelists for the CMA standard setting were selected based on their knowledge of the subject matter assessed, experience in the education of students who take the CMA in grades three through five, familiarity with the California content standards, an understanding of CMA, and an appreciation of the consequences of setting these cut scores. Panelists were recruited to be representative of the educators of the state’s CMA-eligible students (ETS, 2008).

For each test, three cut scores were developed in order to report the four levels; far below basic was defined as chance-level performance. The standard setting processes employed for CMA required panelists to follow the steps listed below:

1. Prior to attending the workshop, all panelists received a pre-workshop assignment. The task was to review, on their own, the content standards upon which the test items are based and take notes on their own expectations in the content area. This allowed the panelists to understand how their perceptions may relate to the complexity of the content standards.
2. At the start of the workshop, panelists received training which included the purpose of standard setting and their role in the work; the meaning of a "cut score" and "impact data"; and specific training and practice in the Bookmark method. Impact data include the percentage of examinees assessed in a previous test administration of the test that would fall into each level, given the panelists' judgments of cut scores.
3. Panelists next became familiar with the difficulty level of the items by taking the actual test, and then assessing and discussing the demands of the test items.
4. Panelists then reviewed the draft list of competencies as a group, noting the increasing demands of each subsequent level. In this step, they began to visualize the knowledge and skills of students in each performance level.
5. Panelists identified characteristics of a "borderline" test taker or "target student." This student is defined as one who possesses just enough knowledge of the content to move over the border separating a performance level from the performance level below it.
6. After training in the method was complete and confirmed through an evaluation questionnaire, panelists made individual judgments. Working in small groups, they discussed feedback related to other panelists' judgments and feedback based on student performance data (impact data). Panelists revised their judgments during the process if they wished.

The final recommended cut scores were based on the median of panelists' judgments[1] at the end of three rounds. For the CMA, the cut scores recommended by the panelists and the recommendation of the State Superintendent of Public Instruction are presented for public comment at regional public hearings. Comments and recommendations are then presented to the SBE for adoption.

Development of Competencies Lists

Prior to the CMA standard-setting workshop, ETS facilitated a meeting in which a subset of the standard-setting panelists were assembled to develop a list of competencies based on the California content standards and California policy level descriptors. For each content area, one panel of educators was assembled for each grade to identify and discuss the competencies required of students in the CMA for each performance level (below basic, basic, proficient, and advanced). The lists were used to facilitate the discussion and construction of the target student definitions during the standard setting workshop.

[1] In the bookmark method, the panel recommendation is calculated by taking the median of the small group (table) medians.

Standard Setting Methodology

Bookmark Method

The Bookmark method for setting passing scores was introduced in 1999 and has been used widely across the United States (Lewis, et al., 1999; Mitzel, et al., 2001). In California, the Bookmark method has been used in standard settings for most of the STAR tests.

This method requires panelists to work through a test booklet in which the items have been re-ordered from easiest to hardest based on student performance data [2]. Panelists are asked to place a bookmark in the ordered booklet to demarcate each performance standard. The bookmarks are placed with the assumption that the borderline students will perform successfully at a given performance level with a probability of at least 0.67. Conversely, these students are expected to perform successfully on the items behind the bookmark with a probability of less than 0.67 (Huynh, 1998).

The cut scores are derived by averaging the corresponding bookmarks placed for the various performance standards across panelists.

Results

The cut scores obtained as a result of the standard-setting process are on the number-correct or raw score scale; the scores are then translated to a score scale that ranges between 150 and 600. The cut score for the basic performance level is 300 for every grade and content area; this means that a student must earn a score of 300 or higher to achieve a basic classification. The cut score for the proficient performance level is 350 for every grade and content area; this means that a student must earn a score of 350 or higher to achieve a proficient classification.

The cut scores for the other performance standards usually vary by grade and subject area. They are derived using procedures based on item response theory (IRT). The raw cut scores for a given test are mapped to IRT *thetas* ($\hat{\theta}_S$) using the test characteristic function[3] and then transformed to the scale score metric using the following equation:

$$Scale\ Cut\ Score = (350 - \theta_{proficient} \times \left(\frac{350-300}{\theta_{proficient} - \theta_{basic}} \right)) + \left(\frac{350-300}{\theta_{proficient} - \theta_{basic}} \right) \times \hat{\theta} \qquad (6.1)$$

where,

$\theta_{proficient}$ represents the theta corresponding to the cut-score for proficient, and

θ_{basic} represents the theta corresponding to the cut score for basic.

The scale scores ranges for each performance level are presented in Table 6.1. The cut score for each performance level is the lower bound of each scale score range. The scale score ranges do not change from year to year, because each grade and content area is scaled independently rather than using a continuous scale from the lowest to the highest grade.

---

[2] An "item map" accompanies the ordered test booklet. The "map" includes information on the content measured by each question, information about each question's difficulty, the correct answer for each question, and where each question was located in the test booklet before the questions were reordered by difficulty.

[3] An IRT test characteristic curve is the sum of item characteristic curves (ICC), where an ICC represents the probability of correctly responding to an item conditioned on examinee ability.

Table 6.1 Scale Scores Ranges for Performance Levels

Subject	CMA	Far Below Basic	Below Basic	Basic	Proficient	Advanced
English–Language Arts	**3**	150 – 234	235 – 299	300 – 349	350 – 396	397 – 600
	4	150 – 249	250 – 299	300 – 349	350 – 406	407 – 600
	5	150 – 219	220 – 299	300 – 349	350 – 401	402 – 600
Mathematics	**3**	150 – 229	230 – 299	300 – 349	350 – 424	425 – 600
	4	150 – 228	229 – 299	300 – 349	350 – 430	431 – 600
	5	150 – 233	234 – 299	300 – 349	350 – 421	422 – 600
Science	**5**	150 – 249	250 – 299	300 – 349	350 – 402	403 – 600

Table 6.2 presents the percentages of examinees meeting each performance standard in 2009.

Table 6.2 Percentage of Examinees in Each Performance Level

Subject	CMA[1]	Far Below Basic	Below Basic	Basic	Proficient	Advanced
English–Language Arts	**3**[2]	16%	23%	21%	21%	20%
	4	17%	27%	25%	19%	13%
	5	6%	33%	27%	21%	14%
Mathematics	**3**	9%	34%	25%	24%	8%
	4	10%	30%	24%	25%	11%
	5	9%	30%	25%	25%	12%
Science	**5**	6%	20%	31%	29%	14%

[1] Proficiency levels are not available for grades six through eight during the 2009 operational administration.

[2] Standard operational form only

The numbers in the summary table may not match exactly the results reported on the CDE Web site because of slight differences in the samples used to compute the statistics. The P1 data file was used for the analyses in this chapter. This file contained more than 99 percent of the entire test-taking population and approximately 99.5 percent of the student records used in the August 18, 2009, reporting of STAR results.

References

Educational Testing Service. (2009). *Technical report on the standard setting workshop for the California Modified Assessment, February 6, 2009,* (California Department of Education Contract Number 5417). Princeton, NJ: Educational Testing Service.

Huynh, H. (1998). On score locations of binary and partial credit items and their applications to item mapping and criterion-referenced interpretation. *Journal of Educational and Behavioral Statistics, 23*(19), 35–56.

Lewis, D. M., Green, D. R., Mitzel, H. C., Baum, K., & Patz, R. J. (1999). *The bookmark standard setting procedure: Methodology and recent implications*. Manuscript under review.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives,* (pp. 249–281). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Chapter 7: Scoring and Reporting

ETS conforms to high standards of quality and fairness (ETS, 2002) when scoring tests and reporting scores. Such standards dictate that ETS provides accurate and understandable assessment results to the intended recipients. It is also ETS's mission to provide appropriate guidelines for score interpretation and cautions about the limitations in the meaning and use of the test scores. Finally, attempts are made to ensure sufficient data are collected for the major subgroups of students. Such data helps ETS to conduct analyses needed to ensure that the assessments are equitable for various groups of test takers.

Procedures for Maintaining and Retrieving Individual Scores

Items for all CMA tests, except for the writing tasks in grades four and seven, are multiple-choice. Students are presented with a question and asked to select the correct answer from among four possible choices. In grade three, students mark their answer choices in the test booklet. In the other grades, students mark their answer choice on an answer document. All multiple-choice questions are machine scored. Responses to the writing task are scored by trained raters.

In order to score and report CMA results, ETS follows an established set of written procedures. These specifications are presented in the next sections.

Scoring and Reporting Specifications

ETS develops standardized scoring procedures and specifications so that test materials are processed and scored accurately. These documents include the following:

- General Reporting Specifications—Provides the calculation rules for the information presented on STAR summary reports and defines the appropriate codes to use when a student does not take or complete a test or when a score will not be reported
- Score Key and Score Conversions—Define file formats and information that is provided for scoring and the process of converting raw scores to scale scores
- Form Planner Specifications—Describes, in detail, the contents of files that contain keys required for scoring
- Aggregation Rules—Describes how and when a school's results are aggregated at the school, district, county, and state level
- "What If" List—Provides a variety of anomalous scenarios that may occur when test materials are returned by districts to Pearson and defines the action(s) to be taken in response
- Edit Specifications—Describes edits, defaults, and solutions to errors encountered while data are being captured as answer documents are processed
- Reporting Cluster Names and Item Numbers—Identifies the reporting clusters for each test and the number of items in each cluster
- CST and CMA Matching Criteria—Describes the criteria necessary to ensure that, for students who take both CST and CMA tests, all results are reported to a single student data record by matching specific demographic fields on test booklets
- Matching Criteria for Multiple-choice and Writing Answer Documents—Describe the method used to match students' writing and multiple-choice results.

The scoring specifications are reviewed and revised by the CDE, ETS, and Pearson each year. After a version that all parties agree to is finalized, the CDE issues a formal approval of the scoring and reporting specifications.

Scanning and Scoring

Answer documents are scanned and scored by Pearson in accord with the scoring specifications that have been approved by the CDE. Answer documents are designed to produce a single complete record for each student. This record includes demographic data and scanned responses for each student; once computed, the scored responses and the total test scores for a student are also merged into the same record. All scores must comply with the ETS scoring specifications. Pearson has quality control checks in place to ensure the quality and accuracy of scanning, and the transfer of scores into the database of student records.

Each school district must return scorable and non-scorable materials within five working days after the last day for each test administration period. For the CMA for Writing test materials, school districts return the writing booklets within two working days after the makeup day for test administration.

Types of Scores and Subscores

Raw Score

For all of the tests, except ELA grades four and seven, the total raw score equals the sum of examinees' scores on the multiple-choice test items. In grades four and seven, the total ELA raw score equals the sum of examinees' scores on both the multiple choice items and the writing task.

Subscore

The CMA items that measure students' progress toward achieving proficiency in California's content standards in ELA and mathematics (grades three through five) and science (grade five) are aggregated into groups, referred to as "reporting clusters." A subscore is obtained by summing an examinee's scores on the items in each reporting cluster. A description of the CMA reporting clusters for grades three through five is provided in Appendix 2.B on page 18.

Scale Score

Raw scores obtained on the ELA and mathematics tests in grades three to five and on the science test in grade five are transformed to three-digit scale scores that are described in Chapter 2 on page 14. Scale scores range from 150 to 600 on each CMA. The raw scores of these examinees cannot be meaningfully compared, because these scores are affected by the difficulty of the test taken as well as the ability of the examinee.

Performance Levels

The CMA performance for each student is categorized into one of the following performance levels:

- far below basic
- below basic
- basic
- proficient
- advanced

For the ELA and mathematics tests in grades three through five and the science test in grade five, the cut score for the basic performance level is 300; this means that a student must achieve a score of 300 or higher to achieve a basic classification. The cut score for the proficient performance level is 350 for every grade and content area; this means that a

student must earn a score of 350 or higher to achieve a proficient classification. The cut scores for the other performance levels usually vary by grade and content area.

Writing Score

Examinees' responses to the essay associated with the grade four and grade seven ELA tests are rated on a 0–4 scale by a single reader. A score of zero indicates that the student attempted the Writing task but either did not provide a response, refused to provide a response, or responded to a Writing task from an earlier test administration. The rubrics used to assign the non-zero scores to the writing tasks for grades four and seven are presented in Appendix 7.A, which starts on page 101.

It is important to note that the writing scores are not related to the performance levels used to show overall student performance on the CMA for ELA test. For example, a writing score of 4 is not related to the advanced performance level, nor does a writing score of 3 relate to the proficient performance level. Students' ELA performance levels are decided in light of their performance on all of the items in the ELA test—in grades four and seven, this total may include the Writing component.[1]

Score Verification Procedures

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

Scoring Key Verification Process

Scoring keys, provided in the form planners, are produced by ETS and verified thoroughly by performing various quality control checks. The form planners contain the information about an assembled test form including scoring keys, test name, administration year, subscore identification, and the standards and statistics associated with each item. Various checks typically are performed before keys are finalized, as listed below:

- The form planners are checked for accuracy against the Form Planner Specification document and the Score Key and Score Conversion document before the keys are loaded into the score key management system (SKM) at ETS.
- The printed lists of the scoring keys are checked again once the keys have been loaded into the SKM system.
- The sequence of items in the form planners are matched with their sequence in the actual test booklets.
- The demarcations of various sections in the actual test book are checked against the list of demarcations provided by ETS test development staff.
- Scoring is verified internally at Pearson. ETS independently generates scores and verifies Pearson's scoring of the data by comparing the two results. Any discrepancies are then resolved.
- The entire scoring system is tested using a test deck that includes typical and extremely atypical responses vectors.

[1] Parents/guardians may choose not to let their child attempt the Writing task. In these cases, the 48-item multiple-choice score counts as the ELA score in grade four, and the 54-item multiple-choice score counts as the ELA score in grade seven.

- Classical item analyses are run on an early sample of data to provide an additional check of the keys. Although rare, if an item is found to be problematic, a followup process is carried out for it to be excluded from further analyses.

Monitoring and Quality Control of Writing Scoring

The CMA for ELA at grades four and seven students respond to one of two writing tasks each year. One task is administered to the majority of test takers; the other is administered to students in schools, tracks or programs not in session during the first administration.

Students' responses to the writing task are read by a single reader; their writing score is based on that reviewer's rating. In addition, ten percent of students' responses also are read by a second reader to provide data that can be used to assess the accuracy and reliability of the writing scores. The score from the second reader does not count toward the student's writing test score. Presented in the next sections are details on the process employed by Pearson to score the writing tasks.

Scoring System

All student responses are scanned into the ePEN™ system. Scorers view assigned responses on a computer at one of Pearson's regional scoring centers. The screen does *not* display the student's name or background information; the scorer sees only the student response.

Scorer Training

Individuals who are selected to serve as scorers must be college graduates who possess at least a Bachelor of Arts degree. Each prospective scorer is required to participate in extensive computer-based training and is provided with the following kinds of information:

- General information about the Electronic Performance Evaluation Network (ePEN™)
- Background information about the STAR Program
- Information about the STAR writing tasks
- Explanations of STAR scoring rubrics and scoring principles
- Sets of prescored annotated training papers (The training papers include anchor and practice papers. Anchor papers provide samples of student writing that represent each of the four points in the scoring rubric.[2] Practice papers include samples of student writing that demonstrate the "high" and "low" end of each score point.)

Scorer Qualification

Once the training is complete, the potential scorers complete the Qualification Sets (three sets of scored papers consisting of ten papers per set) before being eligible to score. On at least two of the three sets, scorers must demonstrate exact agreement or adjacent agreement with the scores assigned to seventy percent of the papers on each of these sets.

Scorers continue to qualify throughout the scoring process. Before each operational scoring session, each scorer scores a Calibration Set of three to four papers. The scores on these sets have been previously agreed upon by scoring directors, in conjunction with other personnel. The sets are given to scorers to ensure that the accuracy of their scoring does not drift. These sets "calibrate" the scorers. The scorers that cannot be calibrated during this process do not qualify for operational scoring.

---

[2] The samples of student responses are identified at the rangefinding sessions. During these sessions, the content experts at Pearson select sample responses that represent each of the four score points and illustrate the different ways of responding to the topic.

Scoring Supervisors and Directors

Scoring supervisors monitor and mentor scorers during operational scoring. If a Writing response is difficult to score, the scorers elevate the response to the scoring supervisor to be scored. In these cases, the scoring supervisor's score is doubled to produce the final Writing score.

Scorers with a history of achieving the highest accuracy on the Qualification Sets and the highest level of scoring consistency and validity statistics during operational scoring are selected as scoring supervisors. Approximately ten scorers are assigned to one scoring supervisor; this ratio allows scoring supervisors to work closely with each scorer. The ePEN™ system also allows scoring supervisors to continuously monitor each response scored and the score point assigned to ensure accuracy. All scoring supervisors participate in a two-day training session that provides the same training that qualify scorers. If a scoring supervisor does not achieve the accuracy required on the Qualification Sets, he or she is not allowed to be a supervisor. In addition, all supervisors received extensive training on how the ePEN™ system works, how to best manage scorers, and how to maintain accuracy as scoring continues.

Scoring directors are responsible for overseeing the scoring of the grade level to which they are assigned. They provide leadership for the scoring supervisors. They also help to manage the scorers and they are ultimately responsible for maintaining the highest accuracy possible during STAR scoring. Scoring directors represent the best of the scoring supervisors. They typically have two to three years experience as scoring supervisors, and they have strong leadership qualities as well as a thorough understanding of STAR scoring.

Accuracy Monitoring

The accuracy of all scoring is regularly monitored using several procedures: First, scoring supervisors and scoring directors constantly monitor the degree to which readers are consistent in scores that they assign. This is done using data provided by the second readers employed to score ten percent of all student responses a second time. The consistency is measured in terms of the percentage of instances in which the first and second readers' scores are identical, adjacent, and nonadjacent; this is a commonly used measure of inter-rater reliability. If a scorer's rate of agreement begins to decline, the scorer is retrained by a scoring supervisor or scoring director and closely monitored thereafter. If the scorer's performance does not improve, the scorer is released.

Second, the consistency between readers' scores and those of scoring directors or supervisors is evaluated using student responses that are known as validity papers, because they have been previously scored by scoring directors or supervisors. One in every 40 papers read by each scorer is a validity paper. The consistency of the scorer's ratings with the scores on the validity papers is checked throughout the day to ensure that each scorer has applied the scoring rubrics accurately. The validity papers are introduced throughout the scoring process. If a scorer's validity falls below required levels, the scorer is retrained by a scoring supervisor or scoring director. If a scorer continues to show poor validity, the scorer is excused.

Third, scoring supervisors "back read" a certain percentage of the student responses that had been scored by the scorers. The scorer and supervisor scores are then compared to check the scorer's consistency and reliability and to ensure that the scorer is maintaining scoring standards. In addition, ePEN™ allows scoring directors to view the back-reading

completed by scoring supervisors to ensure that scoring supervisors are maintaining accuracy. Scoring directors also back read to scorers.

Fourth, to minimize score drift, scorers are required to score a Calibration Set before each scoring session. If a scorer is deficient on any of the accuracy indices, he or she is immediately retrained or released from the scoring process.

Score Verification Process

ETS psychometricians employ special procedures that adjust for item difficulty of one test form to another. As a result of this process, scoring tables are produced. Such tables map the current year's raw score to an appropriate scale score. Pearson utilizes these tables to generate scale scores for each student.

ETS verifies Pearson's scale scores by adhering to procedures such as the following:

- Independently generating the scale scores for students in a small number of school districts and comparing these scores with those generated by Pearson; the selection of school districts is based on the availability of data on all schools included in those districts, known as "complete districts"
- Reviewing longitudinal data for reasonableness; the results of the analyses are used to look at the tends and trends for the complete districts
- Reviewing longitudinal data for reasonableness using 99 percent of the entire testing population; the results are used to evaluate the trends for the state as well as few large school districts.

The results of the longitudinal analyses are provided to the CDE and jointly discussed. Any anomalies in the results are investigated further and jointly discussed with the CDE. Scores are released after explanations that satisfy both the CDE and ETS are obtained.

Overview of Score Aggregation Procedures

In order to provide meaningful results to the stakeholders, CMA scores for a given grade and content area are aggregated at the school, independently testing charter school, district, county, and state levels. The aggregated scores are generated both for individual scores as well as group scores. The following section presents the types of aggregation performed on CMA scores.

Individual Scores

The tables in this section provide state-level summary statistics describing student performance on each CMA. The students in grades three to five who took CMA for ELA and mathematics and those in grade five who took CMA for Science receive scale scores, and they are classified in terms of their performance level. For students in grades six through eight who took the CMA ELA, mathematics and science tests, only percent-correct scores by content area are reported. Students in grade four and seven who took the CMA writing test are also received a raw score for the writing test

Score Distributions and Summary Statistics

Table 7.1, on the next page, presents summary statistics that describe student performance on each CMA. Included in the table are the number of items in each test, the number of examinees taking each test, and the means and standard deviations of student scores expressed in terms of both raw scores and scale scores. The last two columns in the table list the raw score means and standard deviations as percentages of the total raw score points in each test.

The percentages of students in each performance category are presented in Table 7.2. The last column on the table presents the overall proportion of examinees that were classified at the proficient level or higher.

The numbers in the summary tables may not match exactly the results reported on the CDE Web site, as there may be slight differences in the samples used to compute the statistics. The P1 data file was used for the analyses in this chapter. This file contained the test results on more than 99 percent of the entire test taking population.

Table 7.1 Mean and Standard Deviation of Raw and Scale Scores for CMA

				Scale Score		Raw Score		Raw Scr. Pct Poss Range	
Subject	Grade	No. of Items	No. of Examinees	Mean	Std. Dev.	Mean	Std.Dev.	Mean	Std. Dev.
English–Language Arts	3	48	11,321	298	65	27.07	8.64	56.39	18.00
	4	48	17,121	316	72	25.56	7.84	53.26	16.33
	5	48	17,821	320	69	28.78	7.61	59.96	15.85
	6	54	15,900	–	–	28.30	7.37	52.40	13.64
	7	54	13,060	–	–	29.14	8.40	53.96	15.55
	8	54	12,434	–	–	28.57	8.06	52.91	14.92
Mathematics	3	48	11,227	315	70	29.03	8.87	60.47	18.48
	4	48	14,415	321	80	26.77	7.28	55.77	15.17
	5	48	15,656	325	77	28.37	8.06	59.11	16.80
	6	54	15,278	–	–	28.71	7.29	53.17	13.50
	7	54	13,145	–	–	25.07	6.45	46.43	11.94
Science	5	48	16,681	335	58	29.41	7.29	61.28	15.20
	8	54	11,384	–	–	28.12	7.47	52.08	13.83

Note: *Scale scores and proficiency levels were not available for grades six through eight during the spring 2009 operational administration.*

Table 7.2 Percentage of Examinees in Proficiency Levels for CMA

Subject	Grade	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Proficient/ Advanced
English–Language Arts	3	16%	37%	25%	14%	9%	23%
	4	17%	27%	24%	18%	12%	30%
	5	6%	33%	27%	21%	14%	35%
Mathematics	3	9%	34%	25%	24%	8%	32%
	4	10%	30%	24%	25%	11%	36%
	5	9%	30%	25%	25%	12%	37%
Science	5	6%	20%	31%	29%	14%	43%

Note: *Scale scores and proficiency levels were not available for grades six through eight during the spring 2009 operational administration.*

Table 7.B.1 through Table 7.B.3 in Appendix 7.B, starting on page 107, show the distributions of scale scores for each CMA for which scale scores were provided. The results are reported in terms of 15 score intervals, each of which contains 30 scale score points. Table 7.B.4 through Table 7.B.6 show the distribution of raw scores for the CMA tests for grades six through eight.

Group Scores

Statistics summarizing student performance by content area and grade for selected groups of students are provided in Table 7.C.1 through Table 7.C.7 for the CMA tests that reported scale scores and proficiency levels. In the tables, students are grouped by demographic characteristics, including gender, ethnicity, English proficiency, need for special education services, and economic status. The tables show the numbers of valid cases in each group as well as scale score means and standard deviations for each demographic group.

Table 7.3 defines the demographic groups included in the tables. Students' economic status was determined by considering the education level of their parents and whether or not they participated in the National School Lunch Program (NSLP).

Note that the statistics in these tables slightly differ from the statewide statistics reported on the CDE Web site because the P1 data file was used for the analyses in this chapter. In addition, students testing with invalid scores were excluded from the tables.

Table 7.3 Subgroup Definitions

Subgroup	Definition
Gender	Male Female
Ethnicity	American Indian or Alaska Native Asian – Chinese – Japanese – Korean – Vietnamese – Asian Indian – Laotian – Cambodian – Other Asian Pacific Islander – Native Hawaiian – Guamanian – Samoan – Tahitian – Other Pacific Islander Filipino Hispanic or Latino African American White (not Hispanic)
English Language Fluency	English-language fluency Initially fluent English proficient English learner Reclassified fluent English proficient
Economic Status	Not economically disadvantaged Economically disadvantaged
Primary Disability	Mental Retardation Hard of Hearing Deafness Speech/Language Impairment Visual Impairment Emotional disturbance Orthopedic Impairment Other Health Impairment Specific Learning Impairment Deaf Blindness Multiple Group Autism Traumatic Brain Injury

Reports to be Produced and Scores for each Report

The tests that make up the STAR Program provide results or score summaries that are reported for different purposes. The four major purposes include:

1. Communicating with parents and guardians;
2. Informing decisions needed to support student achievement;
3. Evaluating school programs; and
4. Providing data for state and federal accountability programs for schools

A detailed description of the uses and applications of STAR reports is presented in the next section.

Types of Score Reports

There are three categories of CMA reports. These categories and the specific reports in each category are given in the table below.

Table 7.4 Types of CMA Reports

1. Summary Reports	▪ STAR Student Master List Summary ▪ STAR Student Master List Summary, End-of-Course ▪ STAR Subgroup Summary (including the Ethnicity for Economic Status)
2. Individual Reports	▪ STAR Student Record Label ▪ STAR Student Master List ▪ STAR Student Report for CMA
3. Internet Reports	▪ CMA Scores (state, county, district, school) ▪ CMA Summary Scores (state, county, district, school)

These reports are sent to the independently testing charter schools, counties, or school districts; the school district forwards the appropriate reports to test sites or, in the case of the STAR Student Report, sends the reports to the child's parents or guardians and forwards a copy to the student's school or test site. Reports such as the STAR Student Report, Student Record Label, and Student Master List that include individual student results are not distributed beyond the student's school. Internet reports are described on the CDE Web site and are accessible to the public online at http://star.cde.ca.gov/.

Score Report Contents

The STAR Student Report provides scale scores, performance levels, and reporting cluster (subscore) results for each CMA taken by the students in grades three to five. Scale scores are reported on a scale ranging from 150 to 600. The performance levels are far below basic, below basic, basic, proficient, and advanced.

In addition, percent correct scores are provided at the cluster level. Also given for each cluster is the average percent correct cluster score obtained by students who achieved the lowest score that qualifies a student to be classified as proficient on the total test, as well as the average percent correct cluster score obtained by students who achieved the lowest score that qualifies a student to be classified as advanced on the total test. These two averages are given as a range of values in the report.

"For all students in grades six through eight, percent-correct results only are reported; students in grade seven also receive a numerical score (from 1–4) for the CMA for Writing. Performance levels and scale score are not reported this year for these grade levels.

Reports for students with disabilities and English learners who use accommodations include a notation that indicates that the student used accommodations. Scores for students who use accommodations are reported in the same way as they are for non-accommodated tests.

Further information about the STAR Student Report and the other reports is provided in Appendix 7.D. Beginning in 2009, an additional score report, Ethnicity for Economic Status, is produced for the CMA. This Subgroup Summary report aggregates and reports results by selected ethnic groups within economic status.

Score Report Applications

CMA results provide parents and guardians with information about their children's progress. The results are a tool for increasing communication and collaboration between parents or guardians and teachers. Along with teacher report cards and information from school and classroom tests, the STAR Student Report can be used by parents and guardians to talk with teachers about ways to improve their children's achievement of the California content standards.

Schools may use the CMA results to help make decisions about how to best support student achievement. CMA results, however, should never be used as the only source of information to make important decisions about a student's education.

CMA results help school districts and schools identify strengths and weaknesses in their instructional programs. Each year, school districts and school staffs examine CMA results at each grade level and content area tested. Their findings are used to help determine:

- The extent to which students are learning the academic standards,
- Instructional areas that can be improved,
- Teaching strategies that can be developed to address needs of students, and
- Decisions about how to use funds to ensure that students achieve the standards.

The results from the CMA are used for state and federal accountability programs to monitor each school's progress toward achieving established goals.

CMA results also are used to comply with federal Elementary and Secondary Education Act (ESEA) legislation that requires all schools to meet specific academic goals.

Criteria for Interpreting Test Scores

A school district may use CMA results to help make decisions about student placement, promotion, retention, or other considerations related to student achievement. However, it is important to remember that a single test can provide only limited information. Other relevant information should be considered as well. It is advisable for parents to evaluate their child's strengths and weaknesses in the relevant topics by reviewing classroom work and progress reports in addition to the child's CMA results (CDE, 2009). It is also important to note that a student's score in a content area contains measurement error and could vary somewhat if the student was retested.

Criteria for Interpreting Score Reports

The information presented on various reports must be interpreted with caution when making performance comparisons. When comparing scale score and performance level results for the CMA, the user is limited to comparisons within the same content area and grade. This is because the score scales are different for each content area and grade. Comparisons

between raw scores and cluster scores should be limited to comparisons within not only content area and grade but also test year. Comparing scores obtained in different grades or content areas should be avoided, because the results are not on the same scale. The user may compare scores for the same content area and grade, within a school, between schools, or between a school and its district, its county, or the state. The user can also make comparisons within the same grade and content area across years. For more details on the criteria for interpreting information provided on the score reports, see the *2009 STAR Post Test Guide* (CDE, 2009b).

Reference

California Department of Education. (2009a). *2009 STAR CST/CMA, CAPA, and STS printed reports*. http://www.startest.org/pdfs/STAR.reports.2009.pdf.

California Department of Education. (2009b). *2009 STAR post-test guide.* http://www.startest.org/pdfs/STAR.post-test_guide.2009.pdf.

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

Appendix 7.A—ELA for Writing (Grades Four and Seven) Rubrics

Grade Four Scoring Rubric

The scoring rubric shown in this section was used to assign scores to students' written responses on the grade four writing tests. This rubric includes two sets of criteria. The criteria under "The Writing" are adapted from the English–language arts content standards for Writing Strategies and Written Conventions of English. These criteria are used to evaluate written responses in all genres for their clarity of purpose, central idea, and organization and for their use of supporting evidence, sentence variety, and writing conventions. The criteria under "Narrative writing," "Summary writing," and "Response to literature writing," adapted from the grade four Writing Applications Content Standards for these genres, are used to evaluate student writing in the specific genres to which they apply.

Score: 4

The Writing:

- *Clearly* addresses the writing task
- Demonstrates a *clear* understanding of purpose
- Maintains a *consistent* point of view, focus, and organizational structure, including paragraphing when appropriate
- Includes a *clearly presented* central idea with *relevant* facts, details, and/or explanations
- Includes sentence *variety*
- Contains *some errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors do **not** interfere with the reader's understanding of the writing.)

Narrative writing:

- Provides a *thoroughly developed* sequence of *significant* events to relate ideas, observations, and/or memories
- Includes *vivid* descriptive language and sensory details that enable the reader to visualize the events or experiences

Summary writing:

- Summarizes text with *clear* identification of the main idea(s) and the most significant details, in student's own words

Response to literature writing:

- Demonstrates a *clear* understanding of the literary work
- Provides *effective* support for judgments through *specific* references to text and/or prior knowledge

Score: 3

The Writing:

- Addresses *most* of the writing task
- Demonstrates a *general* understanding of purpose
- Maintains a *mostly consistent* point of view, focus, and organizational structure, including paragraphing when appropriate
- Presents a central idea with *mostly* relevant facts, details, and/or explanations
- Includes some sentence *variety*

- Contains *errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors do **not** interfere with the reader's understanding of the writing.)

Narrative writing

- Provides an *adequately developed* sequence of significant events to relate ideas, observations, and/or memories
- Includes *some* descriptive language and sensory details that enable the reader to imagine the events or experiences

Summary writing

- Summarizes text with the main idea(s) and important details, generally in student's own words

Response to literature writing

- Demonstrates an understanding of the literary work
- Provides *some* support for judgments through references to text and/or prior knowledge

Score: 2

The Writing:

- Addresses *some* of the writing task
- Demonstrates *little* understanding of purpose
- Maintains an *inconsistent* point of view, focus, and/or organizational structure; may lack appropriate paragraphing
- S*uggests* a central idea with *limited* facts, details, and/or explanations
- Includes *little* sentence variety
- Contains *many errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors **may** interfere with the reader's understanding of the writing.)

Narrative writing:

- Provides a *minimally developed* sequence of events to relate ideas, observations, and/or memories
- Includes *limited* descriptive language and sensory details that enable the reader to imagine the events or experiences

Summary writing:

- Summarizes text with some of the main idea(s) and details, minimal use of student's own words

Response to literature writing:

- Demonstrates a *limited* understanding of the literary work
- Provides *weak* support for judgments

Score: 1

The Writing:

- Addresses *only one part*, if any, of the writing task
- Demonstrates *no* understanding of purpose
- *Lacks* a clear point of view, focus, and/or organizational structure; may contain inappropriate paragraphing
- *Lacks* a central idea but may contain *marginally related* facts, details, and/or explanations

- Includes *no* sentence variety
- Contains *serious errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors interfere with the reader's understanding of the writing.)

Narrative writing:

- *Lacks* a sequence of events to relate ideas, observations, and/or memories
- *Lacks* descriptive language and sensory details that enable the reader to imagine the events or experiences

Summary writing:

- Summarizes text with few, if any, the main idea(s) and/or details, little or no use of the student's own words

Response to literature writing:

- Demonstrates *little or no* understanding of the literary work
- *Fails* to provide support for judgments

Grade Seven Scoring Rubric

The scoring rubric that follows was used to assign scores to students' written responses on the grade seven writing tests. This rubric includes two sets of criteria. The criteria under "The Writing" are adapted from the English–language arts California content standards for Writing Strategies and Written Conventions of English. These criteria are used to evaluate written responses in all genres for their clarity of purpose, central idea, and organization; their coherence; and their use of supporting evidence, sentence variety, and writing conventions. The criteria under "Fictional or autobiographical narrative writing," "Response to literature writing," "Persuasive writing," and "Summary writing," adapted from the grade seven content standards for Writing Applications for these genres, are used to evaluate student writing in the specific genres to which they apply.

Score: 4

The Writing:

- *Clearly* addresses the writing task
- Demonstrates a *clear* understanding of purpose and audience
- Maintains a *consistent* point of view, focus, and organizational structure, including the *effective* use of transitions
- Includes a *clearly presented* central idea with *relevant* facts, details, and/or explanations
- Includes sentence *variety*
- Contains *some errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors do **not** interfere with the reader's understanding of the writing.)

Fictional or autobiographical narrative writing:

- Provides a *thoroughly developed* plot line, including major and minor characters and a *definite* setting
- Includes *appropriate* strategies (e.g., dialogue; suspense; narrative action)

Response to literature writing:

- Develops interpretations that demonstrate a *thoughtful,* comprehensive grasp of the text
- Organizes *accurate and coherent* interpretations around *clear* ideas, premises, or images from the literary work
- Provides *specific* textual examples and details to support the interpretations

Persuasive writing:

- *Authoritatively* defends a clear position with precise and relevant evidence and *convincingly* addresses the reader's concerns, biases, and expectations

Summary writing:

- Summarizes text with clear identification of the main idea(s) and most significant details, in student's own words, and clearly reflects underlying meaning

Score: 3

The Writing:

- Addresses most of the writing task
- Demonstrates a *general* understanding of purpose and audience

- Maintains a *mostly consistent* point of view, focus, and organizational structure, including use of isolated and/or single word transitions
- Presents a central idea with *mostly relevant* facts, details, and/or explanations
- Includes some sentence *variety*
- Contains *some errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors do **not** interfere with the reader's understanding of the writing.)

Fictional or autobiographical narrative writing:

- Provides an *adequately developed* plot line, including major and minor characters and a *definite* setting
- Includes *appropriate* strategies (e.g., dialogue; suspense; narrative action)

Response to literature writing:

- Develops interpretations that demonstrate a comprehensive grasp of the text
- Organizes accurate and *reasonably* coherent interpretations around *clear* ideas, premises, or images from the literary work
- Provides textual examples and details to support the interpretations

Persuasive writing

- *Generally* defends a position with relevant evidence and addresses the reader's concerns, biases, and expectations

Summary writing

- Summarizes text with the main idea(s) and *important* details, *mostly* in student's own words, and generally reflects underlying meaning

Score: 2

The Writing:

- Addresses *some* of the writing task
- Demonstrates *little* understanding of purpose and audience
- Maintains an *inconsistent* point of view, focus, and/or organizational structure, which may include *ineffective or awkward* transitions that do not unify important ideas
- *Suggests* a central idea with *limited* facts, details, and/or explanations
- Includes *little* sentence variety
- Contains *many errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors **may** interfere with the reader's understanding of the writing.)

Fictional or autobiographical narrative writing:

- Provides a *minimally developed* plot line, including characters and a setting
- *Attempts* to use strategies but with *minimal* effectiveness (e.g., dialogue; suspense; narrative action)

Response to literature writing:

- Develops interpretations that demonstrate a *limited* grasp of the text
- Includes interpretations that *lack* accuracy or coherence as related to ideas, premises, or images from the literary work
- Provides *few, if any,* textual examples and details to support the interpretations

Persuasive writing:

- Defends a position with *little, if any,* evidence and *may* address the reader's concerns, biases, and expectations

Summary writing:

- Summarizes text with *some* of the main idea(s) and details, which may be superficial, minimal use of the student's own words, and minimal reflection of underlying meaning

Score: 1

The Writing:

- Addresses *only one part*, if any, of the writing task
- Demonstrates *no* understanding of purpose and audience
- *Lacks* a point of view, focus, organizational structure, and transitions that unify important ideas
- *Lacks* a central idea but may contain *marginally related* facts, details, and/or explanations
- Includes *no* sentence variety
- Contains *serious errors* in the conventions of the English language (grammar, punctuation, capitalization, and spelling) (These errors interfere with the reader's understanding of the writing.)

Fictional or autobiographical narrative writing:

- *Lacks* a developed plot line
- *Fails* to use strategies (e.g., dialogue; suspense; narrative action)

Response to literature writing:

- Demonstrates *little* grasp of the text
- *Lacks* an interpretation or may be a simple retelling of the passage
- *Lacks* textual examples and details

Persuasive writing:

- *Fails* to defend a position with *any* evidence and *fails* to address the reader's concerns, biases, and expectations

Summary writing:

- Summarizes text with few, if any, of the main ideas and/or details, little or no use of the student's own words, little or no reflection of underlying meaning

Appendix 7.B—Scale Score Distributions

Table 7.B.1 Distribution of CMA Scale Scores for English–Language Arts

Scale Score	Grade 3	Grade 4	Grade 5
570 – 600	3	20	4
540 – 569	14	38	26
510 – 539	37	61	59
480 – 509	0	246	100
450 – 479	145	423	468
420 – 449	316	656	752
390 – 419	473	1,159	1,717
360 – 389	916	1,620	2,213
330 – 359	1,505	2,508	2,420
300 – 329	1,986	2,864	3,127
270 – 299	1,577	2,272	2,212
240 – 269	2,122	3,082	2,392
210 – 239	1,500	1,531	1,598
180 – 209	651	530	562
150 – 179	76	111	171

Table 7.B.2 Distribution of CMA Scale Scores for Mathematics

Scale Score	Grade 3	Grade 4	Grade 5
570 – 600	19	48	126
540 – 569	25	49	0
510 – 539	67	194	120
480 – 509	115	158	377
450 – 479	150	478	517
420 – 449	485	644	753
390 – 419	700	1,240	940
360 – 389	1,574	1,648	1,677
330 – 359	1,208	1,956	2,498
300 – 329	2,032	2,164	2,561
270 – 299	1,427	2,125	1,944
240 – 269	1,760	1,790	2,242
210 – 239	1,129	851	1,185
180 – 209	469	766	530
150 – 179	67	304	186

Table 7.B.3 Distribution of CMA Scale Scores for Science

Scale Score	Grade 5
570 – 600	3
540 – 569	13
510 – 539	45
480 – 509	74
450 – 479	364
420 – 449	710
390 – 419	1,813
360 – 389	2,356
330 – 359	3,349
300 – 329	3,590
270 – 299	2,101
240 – 269	1,492
210 – 239	645
180 – 209	109
150 – 179	17

Table 7.B.4 Distribution of CMA Raw Scores for English–Language Arts Grades Six, Seven, and Eight

	Grade 6				Grade 7				Grade 8			
Raw Score	No. Tested	Pct.	Cum. No. Tested	Pct. Below	No. Tested	%	Cum. No. Tested	Pct. Below	No. Tested	Pct.	Cum. No. Tested	Pct. Below
51–54	1	–	15,900	100%	54	–	13,099	100%	15	–	12,434	100%
48–50	27	–	15,899	100%	91	1%	13,045	100%	88	1%	12,419	100%
45–47	118	1%	15,872	100%	312	2%	12,954	99%	255	2%	12,331	99%
42–44	410	3%	15,754	99%	631	5%	12,642	97%	495	4%	12,076	97%
39–41	887	6%	15,344	97%	988	8%	12,011	92%	757	6%	11,581	93%
36–38	1,440	9%	14,457	91%	1,254	10%	11,023	84%	1,043	8%	10,824	87%
33–35	1,874	12%	13,017	82%	1,417	11%	9,769	75%	1,247	10%	9,781	79%
30–32	2,240	14%	11,143	70%	1,439	11%	8,352	64%	1,436	12%	8,534	69%
27–29	2,274	14%	8,903	56%	1,533	12%	6,913	53%	1,679	14%	7,098	57%
24–26	2,077	13%	6,629	42%	1,599	12%	5,380	41%	1,688	14%	5,419	44%
21–23	1,965	12%	4,552	29%	1,500	11%	3,781	29%	1,603	13%	3,731	30%
18–20	1,481	9%	2,587	16%	1,273	10%	2,281	17%	1,228	10%	2,128	17%
15–17	789	5%	1,106	7%	717	5%	1,008	8%	645	5%	900	7%
12–14	264	2%	317	2%	235	2%	291	2%	220	2%	255	2%
9–11	51	–	53	–	49	–	56	–	32	–	35	–
6–8	1	–	2	–	7	–	7	–	3	–	3	–
3–5	1	–	1	–	0	–	0	–	0	–	0	–
0–2	0	–	0	–	0	–	0	–	0	–	0	–

Percentages less than 1 are not shown.

Table 7.B.5 Distribution of CMA Raw Scores for Mathematics Grades Six and Seven

	Grade 6				Grade 7			
Raw Score	No. Tested	Pct.	Cum. No. Tested	Pct. Below	No. Tested	Percent	Cum. No. Tested	Pct. Below
51–54	9	–	15,278	100%	0	–	13,145	100%
48–50	46	–	15,269	100%	6	–	13,145	100%
45–47	165	1%	15,223	100%	38	–	13,139	100%
42–44	417	3%	15,058	99%	85	1%	13,101	100%
39–41	887	6%	14,641	96%	250	2%	13,016	99%
36–38	1,382	9%	13,754	90%	528	4%	12,766	97%
33–35	1,856	12%	12,372	81%	924	7%	12,238	93%
30–32	2,097	14%	10,516	69%	1,281	10%	11,314	86%
27–29	2,244	15%	8,419	55%	1,877	14%	10,033	76%
24–26	2,197	14%	6,175	40%	2,352	18%	8,156	62%
21–23	1,870	12%	3,978	26%	2,417	18%	5,804	44%
18–20	1,276	8%	2,108	14%	1,964	15%	3,387	26%
15–17	603	4%	832	5%	1,009	8%	1,423	11%
12–14	194	1%	229	1%	350	3%	414	3%
9–11	29	–	35	–	59	–	64	–
6–8	6	–	6	–	5	–	5	–
3–5	0	–	0	–	0	–	0	–
0–2	0	–	0	–	0	–	0	–

Percentages less than 1 are not shown.

Table 7.B.6 Distribution of CMA Raw Scores for Science Grade Eight

	Grade 8			
Raw Score	No. Tested	Pct.	Cum. No. Tested	Pct. Below
51–54	7	–	11,384	100%
48–50	50	–	11,377	100%
45–47	155	1%	11,327	100%
42–44	336	3%	11,172	99%
39–41	586	6%	10,836	97%
36–38	889	9%	10,250	91%
33–35	1,182	12%	9,361	82%
30–32	1,377	14%	8,179	70%
27–29	1,664	14%	6,802	56%
24–26	1,680	13%	5,138	42%
21–23	1,636	12%	3,458	29%
18–20	1,140	9%	1,822	16%
15–17	516	5%	682	7%
12–14	147	2%	166	2%
9–11	17	–	19	–
6–8	1	–	2	–
3–5	1	–	1	–
0–2	0	–	0	–

Percentages less than 1 are not shown.

Appendix 7.C—Demographic Summaries

Table 7.C.1 Demographic Summary for ELA, Grade Three

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. Of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Reading for Understanding	Language
All valid scores	11,321	298	65	16%	37%	25%	14%	9%	63%	55%	53%
Male	7,512	296	65	17%	37%	24%	13%	8%	63%	54%	51%
Female	3,729	303	65	13%	36%	26%	16%	10%	63%	57%	55%
Gender unknown	80	330	81	15%	21%	23%	23%	19%	70%	63%	59%
American Indian	135	310	77	16%	27%	30%	12%	16%	63%	60%	55%
Asian American	381	303	67	15%	33%	27%	14%	9%	65%	54%	55%
Pacific Islander	70	303	64	11%	43%	21%	14%	10%	63%	55%	56%
Filipino	156	305	61	11%	36%	28%	17%	8%	68%	54%	55%
Hispanic	6,718	292	62	17%	40%	24%	12%	7%	60%	53%	51%
African American	1,282	299	67	17%	36%	25%	13%	10%	63%	55%	52%
White	2,294	313	69	12%	30%	27%	19%	13%	68%	59%	56%
Ethnicity unknown	285	318	73	10%	33%	25%	17%	15%	68%	61%	56%
English only	6,153	305	68	14%	34%	25%	16%	11%	65%	57%	54%
Initially fluent English prof.	145	305	68	14%	37%	24%	16%	9%	64%	57%	54%
English learner	4,682	289	60	18%	41%	24%	12%	6%	60%	52%	50%
Reclassified fluent English prof.	37	328	89	5%	35%	22%	16%	22%	69%	62%	59%
English prof. unknown	304	308	64	10%	37%	25%	18%	10%	67%	59%	54%
Autism	652	306	66	14%	33%	28%	14%	11%	68%	55%	54%
Deaf-blindness	0	–	–	–	–	–	–	–	–	–	–
Deafness	77	248	41	35%	53%	6%	5%	0%	44%	39%	43%
Emotional disturbance	218	309	78	17%	31%	21%	17%	14%	67%	58%	53%
Hard of hearing	103	285	54	13%	49%	25%	10%	4%	59%	50%	51%
Mental retardation	305	256	51	36%	44%	14%	5%	2%	49%	43%	41%
Multiple disability	23	274	63	35%	43%	0%	13%	9%	55%	48%	46%
Orthopedic impairment	103	306	68	12%	36%	22%	18%	12%	62%	57%	56%
Other health impairment	780	308	64	12%	33%	27%	17%	10%	67%	58%	54%
Specific learning disability	6,039	298	65	16%	37%	25%	14%	9%	62%	55%	52%
Speech or language Impairment	2,302	299	63	14%	37%	26%	15%	7%	63%	55%	53%
Traumatic brain injury	47	293	58	19%	30%	30%	17%	4%	60%	54%	53%
Visual impairment	43	314	67	7%	35%	35%	9%	14%	64%	61%	57%
Disability unknown	629	302	65	13%	39%	24%	15%	10%	64%	57%	53%
Not economically disadvantaged	2,522	316	68	11%	29%	28%	18%	13%	68%	59%	57%
Economically disadvantaged	8,313	292	63	17%	39%	24%	13%	7%	61%	54%	51%
Economic status unknown	486	313	71	12%	33%	23%	18%	13%	67%	59%	56%
Primary Ethnicity—Not Economically Disadvantaged											
American Indian	24	331	77	8%	29%	21%	21%	21%	73%	64%	59%
Asian American	147	320	68	13%	24%	29%	18%	15%	71%	58%	60%
Pacific Islander	19	330	67	11%	26%	21%	21%	21%	70%	61%	65%
Filipino	83	311	59	11%	29%	34%	18%	8%	71%	55%	57%
Hispanic	860	307	66	13%	32%	28%	16%	10%	65%	58%	55%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. Of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Reading for Understanding	Language
African American	225	310	67	14%	31%	25%	16%	14%	67%	57%	56%
White	1,109	322	69	10%	28%	28%	20%	16%	70%	61%	59%
Ethnicity unknown	55	330	62	2%	31%	31%	20%	16%	73%	63%	61%
Primary Ethnicity—Economically Disadvantaged											
American Indian	104	299	72	18%	29%	31%	10%	13%	60%	57%	53%
Asian American	219	290	64	18%	41%	25%	11%	5%	60%	52%	52%
Pacific Islander	50	293	60	12%	48%	22%	12%	6%	61%	52%	54%
Filipino	62	299	65	13%	39%	24%	16%	8%	63%	55%	53%
Hispanic	5,697	289	61	18%	41%	24%	12%	6%	60%	53%	50%
African American	999	296	66	18%	37%	25%	12%	9%	62%	55%	52%
White	1,079	304	68	15%	32%	26%	17%	10%	64%	57%	53%
Ethnicity unknown	103	310	76	14%	37%	19%	14%	17%	66%	60%	53%
Primary Ethnicity—Unknown Economic Status											
American Indian	7	–	–	–	–	–	–	–	–	–	–
Asian American	15	334	62	7%	13%	47%	20%	13%	75%	65%	62%
Pacific Islander	1	–	–	–	–	–	–	–	–	–	–
Filipino	11	294	55	0%	73%	9%	9%	9%	72%	51%	45%
Hispanic	161	301	69	17%	36%	20%	16%	11%	62%	56%	54%
African American	58	307	75	14%	36%	22%	17%	10%	65%	58%	52%
White	106	323	65	8%	29%	24%	23%	16%	71%	63%	58%
Ethnicity unknown	127	319	74	10%	31%	26%	18%	14%	68%	61%	57%

Table 7.C.2 Demographic Summary for ELA, Grade Four

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Score	Std. Dev of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Understanding	Language
All valid scores	17,121	316	72	17%	29%	24%	18%	12%	63%	48%	52%
Male	11,444	313	72	18%	30%	23%	17%	12%	62%	48%	52%
Female	5,662	320	71	15%	28%	26%	19%	13%	64%	49%	53%
Gender unknown	15	301	68	20%	47%	7%	13%	13%	62%	48%	44%
American Indian	203	319	72	17%	25%	26%	22%	11%	64%	50%	52%
Asian American	545	320	73	16%	26%	25%	21%	12%	62%	49%	54%
Pacific Islander	114	319	74	19%	28%	18%	19%	16%	63%	48%	53%
Filipino	230	329	72	12%	27%	24%	23%	14%	65%	50%	56%
Hispanic	10,173	308	69	19%	32%	23%	16%	10%	61%	47%	50%
African American	1,888	308	71	20%	31%	22%	17%	10%	61%	47%	51%
White	3,800	337	76	11%	23%	24%	22%	20%	67%	53%	57%
Ethnicity unknown	168	339	73	13%	18%	23%	29%	17%	67%	54%	57%
English Only	9,560	325	75	15%	26%	23%	20%	15%	65%	50%	54%
Initially fluent English prof.	340	329	74	10%	31%	23%	17%	19%	67%	50%	55%
English learner	7,078	302	66	21%	33%	24%	15%	8%	60%	45%	49%
Reclassified fluent English prof.	100	348	76	6%	22%	25%	28%	19%	71%	54%	60%
English prof. unknown	43	300	67	26%	30%	16%	16%	12%	58%	49%	47%
Autism	792	319	76	16%	30%	24%	17%	14%	63%	50%	54%
Deaf-blindness	1	–	–	–	–	–	–	–	–	–	–
Deafness	77	268	60	38%	40%	12%	8%	3%	43%	43%	42%
Emotional disturbance	344	325	80	17%	25%	21%	20%	17%	66%	52%	53%
Hard of hearing	151	313	68	17%	30%	23%	17%	13%	62%	49%	52%
Mental retardation	284	259	54	45%	35%	13%	6%	1%	45%	39%	40%
Multiple disability	34	279	58	32%	26%	29%	9%	3%	53%	41%	44%
Orthopedic impairment	144	317	80	21%	27%	19%	17%	15%	61%	50%	52%
Other health impairment	1,252	333	76	13%	25%	23%	21%	18%	68%	52%	55%
Specific learning disability	9,899	314	72	18%	29%	23%	17%	12%	62%	48%	52%
Speech or language impairment	2,873	319	68	14%	30%	25%	19%	12%	64%	49%	53%
Traumatic brain injury	42	313	68	19%	21%	36%	7%	17%	64%	49%	50%
Visual impairment	49	305	73	16%	39%	20%	10%	14%	59%	45%	51%
Disability unknown	1,179	318	67	14%	27%	28%	20%	11%	64%	48%	53%
Not economically disadvantaged	4,560	338	76	11%	23%	24%	23%	20%	68%	53%	57%
Economically disadvantaged	12,498	307	68	19%	31%	23%	16%	10%	61%	47%	50%
Economic status unknown	63	309	75	25%	27%	13%	22%	13%	63%	47%	50%
Primary Ethnicity—Not economically disadvantaged											
American Indian	58	315	74	17%	33%	19%	21%	10%	60%	49%	52%
Asian American	224	333	70	10%	23%	28%	25%	14%	67%	51%	57%
Pacific Islander	46	335	77	11%	28%	22%	17%	22%	67%	50%	58%
Filipino	128	328	67	13%	27%	23%	26%	13%	66%	50%	56%
Hispanic	1,466	331	74	13%	24%	25%	21%	18%	67%	51%	55%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Score	Std. Dev of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Understanding	Language
African American	441	321	75	18%	24%	21%	22%	14%	64%	49%	54%
White	2,123	349	77	8%	21%	23%	24%	24%	70%	55%	59%
Ethnicity unknown	74	354	81	12%	11%	27%	23%	27%	69%	57%	61%
Primary Ethnicity—Economically disadvantaged											
American Indian	145	320	71	17%	21%	29%	22%	11%	66%	50%	52%
Asian American	321	312	73	20%	28%	23%	17%	11%	60%	47%	53%
Pacific Islander	68	308	71	25%	28%	15%	21%	12%	60%	46%	51%
Filipino	102	329	77	12%	27%	25%	20%	16%	65%	51%	55%
Hispanic	8,676	304	67	20%	33%	23%	15%	9%	60%	46%	49%
African American	1,440	305	69	20%	33%	22%	16%	9%	60%	46%	50%
White	1,665	323	73	15%	25%	26%	19%	14%	64%	50%	54%
Ethnicity unknown	81	326	62	14%	22%	22%	35%	7%	64%	52%	55%
Primary Ethnicity—Unknown Economic Status											
American Indian	0	–	–	–	–	–	–	–	–	–	–
Asian American	0	–	–	–	–	–	–	–	–	–	–
Pacific Islander	0	–	–	–	–	–	–	–	–	–	–
Filipino	0	–	–	–	–	–	–	–	–	–	–
Hispanic	31	285	66	35%	26%	16%	23%	0%	60%	41%	45%
African American	7	–	–	–	–	–	–	–	–	–	–
White	12	366	70	0%	25%	17%	17%	42%	70%	66%	60%
Ethnicity unknown	13	330	76	15%	31%	8%	23%	23%	69%	53%	53%

Table 7.C.3 Demographic Summary for ELA, Grade Five

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Reading for Understanding	Language
All valid scores	17,821	320	69	6%	33%	27%	21%	14%	68%	51%	64%
Male	11,862	317	70	6%	34%	26%	20%	13%	68%	51%	63%
Female	5,948	326	68	5%	31%	27%	23%	15%	68%	52%	66%
Gender unknown	11	275	66	18%	36%	27%	18%	0%	58%	44%	51%
American Indian	227	330	67	1%	33%	26%	21%	18%	73%	53%	66%
Asian American	527	316	65	5%	36%	28%	19%	13%	67%	49%	64%
Pacific Islander	104	327	72	7%	27%	24%	29%	13%	68%	52%	67%
Filipino	215	323	74	7%	26%	35%	16%	16%	69%	50%	66%
Hispanic	10,763	314	67	6%	36%	27%	19%	12%	66%	50%	63%
African American	2,069	314	68	7%	35%	26%	21%	11%	68%	50%	63%
White	3,769	339	74	4%	25%	25%	24%	22%	73%	56%	68%
Ethnicity unknown	147	319	73	5%	39%	21%	18%	16%	68%	53%	62%
English Only	9,776	327	71	5%	30%	26%	22%	17%	70%	53%	66%
Initially fluent English prof.	401	336	69	3%	25%	27%	26%	18%	72%	55%	68%
English learner	7,445	309	65	7%	38%	28%	18%	10%	64%	49%	62%
Reclassified fluent English prof.	156	341	70	3%	29%	19%	27%	22%	71%	57%	69%
English prof. unknown	43	297	71	16%	40%	21%	9%	14%	60%	49%	57%
Autism	699	311	69	7%	38%	26%	19%	11%	67%	47%	63%
Deaf-blindness	0	–	–	–	–	–	–	–	–	–	–
Deafness	81	275	62	16%	53%	16%	10%	5%	53%	40%	55%
Emotional disturbance	362	325	77	6%	32%	25%	19%	17%	72%	53%	63%
Hard of hearing	155	323	65	4%	31%	28%	24%	13%	66%	52%	66%
Mental retardation	383	255	54	23%	57%	14%	5%	2%	49%	37%	48%
Multiple disability	55	285	59	15%	40%	33%	9%	4%	63%	44%	55%
Orthopedic impairment	113	314	74	7%	35%	25%	19%	14%	64%	50%	63%
Other health impairment	1,313	332	71	5%	27%	27%	23%	18%	72%	54%	67%
Specific learning disability	11,138	321	69	5%	33%	26%	21%	14%	68%	52%	64%
Speech or language impairment	2,417	321	65	5%	31%	29%	22%	12%	68%	51%	65%
Traumatic brain injury	49	313	79	14%	27%	24%	18%	16%	66%	48%	63%
Visual impairment	35	311	74	6%	43%	20%	14%	17%	66%	54%	58%
Disability unknown	1,021	320	65	6%	31%	28%	22%	13%	70%	51%	64%
Not economically disadvantaged	4,560	338	72	4%	26%	26%	24%	21%	73%	55%	68%
Economically disadvantaged	13,191	314	67	6%	36%	27%	20%	11%	66%	50%	63%
Economic status unknown	70	285	62	14%	46%	23%	13%	4%	58%	43%	56%
Primary Ethnicity—Not economically disadvantaged											
American Indian	55	356	78	0%	29%	16%	22%	33%	79%	59%	71%
Asian American	221	324	64	4%	33%	26%	21%	16%	70%	51%	66%
Pacific Islander	33	335	78	9%	21%	21%	33%	15%	72%	53%	68%
Filipino	132	329	70	5%	24%	36%	18%	17%	70%	51%	68%
Hispanic	1,527	332	70	4%	29%	27%	23%	18%	72%	54%	67%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Vocabulary	Reading for Understanding	Language
African American	488	324	69	6%	29%	27%	24%	14%	71%	52%	65%
White	2,051	348	73	3%	22%	24%	26%	25%	75%	58%	70%
Ethnicity unknown	53	343	77	6%	23%	25%	23%	25%	78%	58%	65%
Primary Ethnicity—Economically disadvantaged											
American Indian	168	324	62	1%	33%	30%	21%	14%	71%	52%	65%
Asian American	302	310	65	5%	38%	29%	17%	10%	64%	48%	63%
Pacific Islander	71	323	69	6%	30%	25%	27%	13%	66%	51%	67%
Filipino	83	314	78	11%	28%	35%	13%	13%	66%	49%	63%
Hispanic	9,215	311	66	7%	37%	27%	19%	10%	65%	50%	63%
African American	1,571	311	67	7%	37%	25%	21%	10%	67%	49%	62%
White	1,701	329	73	6%	29%	25%	22%	18%	71%	54%	65%
Ethnicity unknown	80	311	68	4%	48%	20%	15%	14%	65%	51%	60%
Primary Ethnicity—Unknown Economic Status											
American Indian	4	–	–	–	–	–	–	–	–	–	–
Asian American	4	–	–	–	–	–	–	–	–	–	–
Pacific Islander	0	–	–	–	–	–	–	–	–	–	–
Filipino	0	–	–	–	–	–	–	–	–	–	–
Hispanic	21	285	71	14%	57%	5%	10%	14%	57%	43%	57%
African American	10	284	44	10%	40%	40%	10%	0%	66%	36%	60%
White	17	289	72	24%	18%	47%	12%	0%	61%	48%	54%
Ethnicity unknown	14	279	63	14%	50%	14%	21%	0%	54%	44%	54%

Table 7.C.4 Demographic Summary for Mathematics, Grade Three

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
All valid scores	11,227	315	70	9%	34%	25%	24%	8%	57%	62%	66%
Male	7,454	316	71	9%	34%	25%	24%	8%	57%	62%	67%
Female	3,706	313	68	9%	35%	26%	24%	6%	57%	61%	65%
Gender unknown	67	339	73	6%	25%	24%	30%	15%	63%	69%	72%
American Indian	132	319	74	14%	32%	19%	27%	9%	58%	62%	67%
Asian American	378	325	78	10%	29%	23%	26%	12%	61%	62%	67%
Pacific Islander	72	313	80	10%	43%	18%	22%	7%	56%	59%	66%
Filipino	157	316	74	11%	32%	24%	24%	10%	57%	61%	67%
Hispanic	6,670	313	69	9%	35%	26%	23%	7%	56%	61%	66%
African American	1,272	305	67	12%	38%	25%	20%	5%	54%	59%	65%
White	2,278	325	72	7%	31%	25%	27%	10%	59%	65%	69%
Ethnicity unknown	268	324	73	7%	32%	23%	25%	12%	60%	63%	69%
English Only	6,097	317	71	9%	34%	24%	25%	8%	57%	62%	67%
Initially fluent English prof.	144	324	72	6%	35%	22%	28%	9%	58%	65%	69%
English learner	4,662	312	68	10%	34%	27%	23%	7%	57%	61%	65%
Reclassified fluent English prof.	37	341	87	8%	19%	32%	22%	19%	64%	65%	72%
English prof. unknown	287	322	68	6%	33%	27%	26%	8%	59%	65%	68%
Autism	650	305	69	11%	41%	23%	20%	6%	54%	59%	63%
Deaf-blindness	0	–	–	–	–	–	–	–	–	–	–
Deafness	77	299	67	13%	40%	25%	16%	6%	55%	58%	56%
Emotional disturbance	215	307	75	13%	37%	21%	20%	8%	54%	60%	63%
Hard of hearing	101	320	62	6%	32%	30%	25%	8%	59%	64%	67%
Mental retardation	300	261	55	26%	53%	13%	5%	2%	42%	48%	50%
Multiple disability	23	285	59	17%	43%	22%	17%	0%	49%	56%	56%
Orthopedic impairment	96	301	73	14%	41%	18%	20%	8%	53%	58%	61%
Other health impairment	781	311	67	8%	37%	26%	22%	7%	56%	61%	66%
Specific learning disability	5,990	319	70	8%	32%	26%	25%	8%	58%	63%	68%
Speech or language impairment	2,295	317	71	10%	33%	23%	26%	8%	58%	62%	66%
Traumatic brain injury	46	317	74	11%	26%	35%	15%	13%	57%	62%	69%
Visual impairment	43	322	86	12%	33%	23%	16%	16%	57%	63%	67%
Disability unknown	610	321	71	7%	32%	28%	23%	9%	59%	62%	68%
Not economically disadvantaged	2,500	326	71	7%	29%	26%	28%	10%	60%	65%	69%
Economically disadvantaged	8,267	312	69	10%	36%	25%	23%	7%	56%	61%	65%
Economic status unknown	460	322	71	8%	32%	26%	25%	10%	59%	64%	68%
Primary Ethnicity—Not economically disadvantaged											
American Indian	24	351	79	8%	21%	8%	46%	17%	66%	74%	72%
Asian American	146	335	79	10%	23%	23%	29%	16%	65%	64%	68%
Pacific Islander	19	312	58	11%	32%	37%	16%	5%	57%	62%	66%
Filipino	84	322	68	7%	32%	24%	26%	11%	60%	62%	68%
Hispanic	854	322	69	7%	31%	28%	27%	8%	59%	64%	68%
African American	220	312	66	10%	31%	35%	19%	6%	57%	60%	67%
White	1,102	331	72	6%	29%	24%	30%	11%	61%	66%	70%
Ethnicity unknown	51	321	65	6%	27%	29%	27%	10%	58%	63%	70%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
Primary Ethnicity—Economically disadvantaged											
American Indian	103	309	71	16%	35%	21%	21%	7%	55%	59%	66%
Asian American	217	317	76	12%	33%	24%	23%	9%	58%	60%	65%
Pacific Islander	52	315	87	10%	46%	12%	25%	8%	56%	58%	66%
Filipino	62	308	80	16%	31%	23%	21%	10%	54%	60%	65%
Hispanic	5,662	311	69	10%	35%	26%	23%	7%	56%	61%	65%
African American	996	303	67	12%	39%	22%	21%	5%	53%	58%	64%
White	1,073	319	71	8%	34%	25%	25%	9%	57%	64%	68%
Ethnicity unknown	102	321	77	9%	34%	22%	22%	14%	59%	62%	68%
Primary Ethnicity—Unknown Economic Status											
American Indian	5	–	–	–	–	–	–	–	–	–	–
Asian American	15	351	74	0%	27%	20%	40%	13%	66%	70%	79%
Pacific Islander	1	–	–	–	–	–	–	–	–	–	–
Filipino	11	310	78	9%	36%	27%	18%	9%	52%	62%	70%
Hispanic	154	316	70	9%	34%	25%	23%	9%	57%	63%	66%
African American	56	307	71	14%	34%	25%	21%	5%	53%	62%	64%
White	103	327	67	5%	25%	34%	25%	11%	60%	65%	71%
Ethnicity unknown	115	328	74	6%	33%	22%	28%	11%	62%	63%	69%

Table 7.C.5 Demographic Summary for Mathematics, Grade Four

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
All valid scores	14,415	321	80	10%	30%	24%	25%	11%	61%	49%	53%
Male	9,365	322	82	11%	30%	24%	24%	12%	62%	49%	53%
Female	5,039	321	77	9%	31%	24%	26%	10%	61%	49%	53%
Gender unknown	11	291	71	9%	55%	27%	0%	9%	55%	42%	49%
American Indian	183	327	75	7%	27%	30%	25%	11%	63%	50%	54%
Asian American	432	340	92	9%	25%	20%	27%	19%	65%	53%	53%
Pacific Islander	99	317	83	12%	27%	30%	22%	8%	61%	47%	53%
Filipino	192	338	83	9%	23%	22%	33%	13%	66%	51%	56%
Hispanic	8,593	319	78	10%	32%	25%	24%	10%	61%	49%	52%
African American	1,721	303	78	16%	33%	23%	21%	7%	57%	47%	50%
White	3,065	335	83	8%	26%	22%	28%	15%	64%	51%	56%
Ethnicity unknown	130	322	80	8%	32%	29%	18%	12%	61%	49%	54%
English Only	8,044	323	81	11%	29%	23%	25%	12%	61%	50%	53%
Initially fluent English prof.	267	335	89	10%	24%	25%	26%	15%	64%	52%	55%
English learner	5,988	318	78	10%	32%	25%	24%	10%	61%	48%	52%
Reclassified fluent English prof.	78	340	75	10%	18%	21%	36%	15%	65%	53%	56%
English prof. unknown	38	296	60	8%	47%	24%	18%	3%	57%	42%	51%
Autism	708	311	90	16%	32%	20%	20%	12%	59%	47%	50%
Deaf-blindness	1	–	–	–	–	–	–	–	–	–	–
Deafness	73	313	81	11%	32%	29%	21%	8%	59%	49%	49%
Emotional disturbance	302	301	83	19%	34%	21%	18%	9%	57%	46%	49%
Hard of hearing	134	336	77	7%	26%	26%	27%	14%	64%	54%	52%
Mental retardation	283	251	57	37%	43%	13%	6%	1%	45%	38%	41%
Multiple disability	34	282	63	12%	56%	18%	12%	3%	52%	43%	47%
Orthopedic impairment	135	292	86	21%	37%	16%	16%	9%	52%	46%	51%
Other health impairment	1,063	324	80	10%	29%	23%	27%	11%	62%	50%	53%
Specific learning disability	8,221	323	78	9%	30%	25%	25%	11%	62%	50%	53%
Speech or language impairment	2,360	326	81	10%	29%	23%	26%	12%	63%	50%	53%
Traumatic brain injury	36	322	83	6%	36%	31%	17%	11%	61%	51%	49%
Visual impairment	45	292	77	11%	51%	16%	13%	9%	53%	43%	49%
Disability unknown	1,020	329	75	7%	28%	27%	26%	12%	63%	50%	55%
Not economically disadvantaged	3,681	334	85	9%	26%	23%	27%	15%	64%	51%	55%
Economically disadvantaged	10,684	317	78	11%	32%	24%	24%	9%	61%	48%	52%
Economic status unknown	50	297	70	12%	44%	20%	16%	8%	56%	43%	51%
Primary Ethnicity—Not economically disadvantaged											
American Indian	52	326	66	2%	31%	35%	25%	8%	64%	49%	54%
Asian American	177	352	96	9%	19%	21%	26%	24%	68%	55%	54%
Pacific Islander	38	316	90	13%	26%	29%	21%	11%	58%	51%	53%
Filipino	109	342	85	11%	19%	22%	34%	14%	66%	52%	56%
Hispanic	1,183	328	82	9%	27%	24%	27%	12%	63%	51%	53%
African American	389	314	81	12%	33%	23%	23%	9%	60%	48%	51%
White	1,675	342	85	8%	24%	21%	29%	18%	65%	52%	57%
Ethnicity unknown	58	333	95	9%	33%	22%	19%	17%	61%	51%	58%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
Primary Ethnicity—Economically disadvantaged											
American Indian	131	328	78	8%	26%	28%	25%	12%	63%	50%	54%
Asian American	255	331	89	10%	29%	18%	28%	15%	63%	52%	52%
Pacific Islander	61	317	79	11%	28%	31%	23%	7%	63%	44%	52%
Filipino	83	332	80	7%	28%	22%	31%	12%	64%	50%	56%
Hispanic	7,384	317	77	10%	32%	25%	24%	9%	61%	48%	52%
African American	1,327	301	77	17%	33%	23%	20%	7%	56%	47%	49%
White	1,381	326	80	9%	29%	24%	27%	11%	62%	50%	54%
Ethnicity unknown	62	314	64	8%	31%	35%	19%	6%	60%	49%	51%
Primary Ethnicity—Unknown Economic Status											
American Indian	0	–	–	–	–	–	–	–	–	–	–
Asian American	0	–	–	–	–	–	–	–	–	–	–
Pacific Islander	0	–	–	–	–	–	–	–	–	–	–
Filipino	0	–	–	–	–	–	–	–	–	–	–
Hispanic	26	282	59	15%	46%	23%	15%	0%	52%	44%	48%
African American	5	–	–	–	–	–	–	–	–	–	–
White	9	–	–	–	–	–	–	–	–	–	–
Ethnicity unknown	10	300	69	10%	40%	30%	10%	10%	63%	39%	47%

Table 7.C.6 Demographic Summary for Mathematics, Grade Five

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
All valid scores	15,656	325	77	9%	30%	25%	25%	12%	63%	58%	51%
Male	10,193	323	78	10%	31%	24%	23%	12%	63%	58%	51%
Female	5,452	329	76	8%	28%	25%	27%	12%	65%	59%	52%
Gender unknown	11	287	63	9%	64%	18%	0%	9%	58%	42%	49%
American Indian	198	325	72	8%	33%	25%	23%	11%	63%	59%	52%
Asian American	426	336	82	8%	27%	23%	27%	16%	67%	60%	53%
Pacific Islander	89	329	76	7%	29%	26%	26%	12%	65%	58%	52%
Filipino	178	335	84	8%	31%	17%	27%	16%	66%	59%	53%
Hispanic	9,329	325	76	9%	30%	26%	24%	11%	63%	58%	51%
African American	1,956	311	75	12%	35%	23%	21%	9%	60%	55%	49%
White	3,348	334	80	8%	27%	23%	27%	15%	65%	61%	54%
Ethnicity unknown	132	309	77	11%	41%	20%	18%	11%	59%	55%	50%
English Only	8,757	325	79	10%	30%	24%	24%	12%	63%	58%	52%
Initially fluent English prof.	336	346	76	6%	22%	22%	35%	16%	68%	64%	55%
English learner	6,389	324	75	8%	30%	26%	24%	11%	64%	58%	51%
Reclassified fluent English prof.	132	342	76	5%	23%	27%	29%	16%	66%	64%	54%
English prof. unknown	42	300	76	14%	45%	14%	19%	7%	58%	50%	50%
Autism	664	311	84	16%	35%	20%	16%	13%	59%	54%	50%
Deaf-blindness	0	–	–	–	–	–	–	–	–	–	–
Deafness	72	326	94	15%	25%	24%	19%	17%	65%	56%	49%
Emotional disturbance	364	309	74	12%	39%	21%	19%	10%	58%	56%	48%
Hard of hearing	134	348	84	4%	22%	26%	29%	19%	68%	63%	54%
Mental retardation	371	252	51	36%	47%	13%	4%	1%	45%	39%	40%
Multiple disability	53	292	64	17%	47%	19%	9%	8%	55%	48%	49%
Orthopedic impairment	119	292	79	24%	34%	17%	18%	8%	55%	50%	47%
Other health impairment	1,226	328	79	9%	29%	22%	27%	14%	63%	60%	52%
Specific learning disability	9,584	328	76	8%	29%	26%	26%	12%	64%	59%	52%
Speech or language impairment	2,066	328	79	9%	28%	25%	25%	13%	64%	59%	52%
Traumatic brain injury	46	310	79	20%	28%	13%	28%	11%	61%	53%	52%
Visual impairment	37	325	82	11%	30%	16%	32%	11%	64%	58%	51%
Disability unknown	920	332	74	7%	26%	28%	27%	12%	65%	60%	53%
Not economically disadvantaged	4,030	335	79	8%	25%	24%	28%	15%	65%	61%	54%
Economically disadvantaged	11,560	322	76	9%	31%	25%	24%	11%	63%	58%	51%
Economic status unknown	66	282	77	23%	44%	17%	11%	6%	54%	45%	44%
Primary Ethnicity—Not economically disadvantaged											
American Indian	45	329	71	4%	31%	31%	20%	13%	63%	61%	53%
Asian American	173	340	83	8%	23%	25%	29%	15%	68%	60%	54%
Pacific Islander	30	330	72	10%	17%	30%	30%	13%	67%	57%	53%
Filipino	116	344	85	7%	30%	15%	28%	20%	68%	61%	54%
Hispanic	1,348	330	76	8%	25%	26%	28%	12%	64%	60%	52%
African American	466	320	79	13%	29%	24%	24%	11%	62%	57%	51%
White	1,802	341	80	7%	24%	23%	28%	18%	66%	62%	55%
Ethnicity unknown	50	330	87	12%	26%	20%	24%	18%	62%	60%	55%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Number Sense	Algebra and Data Analysis	Measurement and Geometry
Primary Ethnicity—Economically disadvantaged											
American Indian	149	325	72	8%	34%	23%	24%	11%	63%	58%	51%
Asian American	249	334	81	7%	29%	22%	24%	18%	66%	59%	53%
Pacific Islander	59	328	79	5%	36%	24%	24%	12%	64%	58%	52%
Filipino	62	317	79	11%	34%	23%	24%	8%	62%	55%	51%
Hispanic	7,965	324	75	9%	30%	25%	24%	11%	63%	58%	51%
African American	1,480	308	74	12%	37%	22%	20%	8%	60%	54%	49%
White	1,528	326	79	9%	30%	23%	25%	13%	63%	59%	52%
Ethnicity unknown	68	301	69	9%	47%	21%	18%	6%	57%	53%	47%
Primary Ethnicity—Unknown Economic Status											
American Indian	4	–	–	–	–	–	–	–	–	–	–
Asian American	4	–	–	–	–	–	–	–	–	–	–
Pacific Islander	0	–	–	–	–	–	–	–	–	–	–
Filipino	0	–	–	–	–	–	–	–	–	–	–
Hispanic	16	292	91	19%	56%	6%	6%	13%	55%	48%	43%
African American	10	264	51	10%	50%	40%	0%	0%	49%	44%	40%
White	18	269	85	44%	28%	11%	11%	6%	51%	41%	42%
Ethnicity unknown	14	277	59	14%	64%	14%	0%	7%	55%	41%	44%

Table 7.C.7 Demographic Summary for Science, Grade Five

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Physical Sciences	Life Sciences	Earth Sciences
All valid scores	16,681	335	58	6%	20%	31%	29%	14%	63%	63%	58%
Male	11,052	337	60	7%	19%	30%	29%	15%	64%	64%	58%
Female	5,617	331	55	6%	21%	34%	28%	11%	62%	62%	57%
Gender unknown	12	295	57	25%	33%	17%	25%	0%	56%	56%	40%
American Indian	213	344	59	4%	17%	30%	32%	17%	65%	65%	60%
Asian American	484	325	55	9%	23%	33%	27%	9%	62%	59%	55%
Pacific Islander	103	341	54	4%	20%	26%	36%	14%	66%	64%	59%
Filipino	202	333	57	7%	18%	36%	28%	11%	63%	62%	57%
Hispanic	9,960	331	56	7%	21%	33%	28%	11%	62%	62%	56%
African American	1,979	322	57	11%	22%	34%	24%	9%	61%	60%	53%
White	3,598	356	61	4%	13%	26%	33%	25%	67%	69%	63%
Ethnicity unknown	142	338	68	6%	22%	26%	31%	15%	63%	64%	58%
English Only	9,243	342	60	6%	17%	29%	30%	17%	64%	65%	59%
Initially fluent English prof.	365	353	56	3%	14%	28%	33%	22%	66%	67%	64%
English learner	6,881	326	54	7%	23%	34%	27%	9%	61%	60%	55%
Reclassified fluent English prof.	148	352	52	3%	10%	30%	41%	16%	67%	67%	64%
English prof. unknown	44	305	55	14%	30%	36%	16%	5%	56%	56%	48%
Autism	679	327	63	10%	23%	29%	24%	13%	63%	58%	56%
Deaf-blindness	0	–	–	–	–	–	–	–	–	–	–
Deafness	81	299	60	22%	30%	26%	14%	9%	54%	51%	48%
Emotional disturbance	345	336	68	12%	16%	27%	28%	18%	63%	63%	57%
Hard of hearing	149	334	53	6%	15%	38%	31%	9%	65%	61%	57%
Mental retardation	376	278	49	28%	39%	23%	9%	1%	49%	47%	41%
Multiple disability	55	299	43	7%	45%	29%	18%	0%	58%	53%	45%
Orthopedic impairment	115	320	55	12%	21%	37%	23%	8%	59%	60%	54%
Other health impairment	1,239	345	59	5%	16%	30%	31%	18%	65%	66%	60%
Specific learning disability	10,346	338	58	5%	19%	31%	30%	15%	64%	64%	58%
Speech or language impairment	2,263	332	54	5%	21%	35%	28%	11%	63%	62%	57%
Traumatic brain injury	47	315	67	11%	28%	32%	17%	13%	58%	58%	51%
Visual impairment	33	327	61	6%	24%	30%	27%	12%	60%	59%	58%
Disability unknown	953	336	56	6%	18%	32%	30%	13%	64%	63%	57%
Not economically disadvantaged	4,300	350	60	5%	15%	28%	32%	21%	66%	67%	62%
Economically disadvantaged	12,314	330	57	7%	21%	32%	28%	11%	62%	62%	56%
Economic status unknown	67	307	56	10%	36%	30%	18%	6%	58%	56%	48%
Primary Ethnicity—Not economically disadvantaged											
American Indian	51	363	63	2%	12%	24%	39%	24%	66%	71%	67%
Asian American	192	330	57	8%	24%	28%	28%	12%	63%	60%	57%
Pacific Islander	35	339	57	3%	23%	26%	37%	11%	68%	62%	58%
Filipino	126	337	55	6%	14%	39%	29%	12%	65%	63%	58%

				Percent in Performance Level					Mean Percent Correct in Content Area		
	No. Tested	Mean Scale Scores	Std. Dev. of Scale Scores	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Physical Sciences	Life Sciences	Earth Sciences
Hispanic	1,434	342	58	6%	16%	31%	30%	17%	65%	65%	59%
African American	466	330	58	9%	20%	31%	28%	12%	63%	61%	56%
White	1,944	364	59	3%	11%	24%	34%	28%	69%	71%	65%
Ethnicity unknown	52	354	64	4%	8%	33%	35%	21%	66%	68%	63%
Primary Ethnicity—Economically disadvantaged											
American Indian	158	340	55	4%	18%	31%	31%	15%	64%	64%	59%
Asian American	288	322	53	10%	22%	36%	26%	7%	62%	59%	53%
Pacific Islander	68	342	53	4%	19%	26%	35%	15%	65%	65%	60%
Filipino	76	325	60	9%	24%	30%	26%	11%	60%	60%	56%
Hispanic	8,509	329	55	7%	22%	33%	28%	10%	62%	61%	56%
African American	1,503	320	56	11%	23%	34%	23%	8%	60%	59%	53%
White	1,636	348	60	5%	16%	28%	31%	21%	65%	67%	61%
Ethnicity unknown	76	336	69	5%	28%	24%	30%	13%	62%	63%	57%
Primary Ethnicity—Unknown Economic Status											
American Indian	4	–	–	–	–	–	–	–	–	–	–
Asian American	4	–	–	–	–	–	–	–	–	–	–
Pacific Islander	0	0	0	0%	0%	0%	0%	0%	0%	0%	0%
Filipino	0	0	0	0%	0%	0%	0%	0%	0%	0%	0%
Hispanic	17	316	45	0%	41%	35%	18%	6%	61%	58%	49%
African American	10	287	36	10%	50%	40%	0%	0%	54%	50%	41%
White	18	325	72	11%	22%	28%	22%	17%	56%	61%	59%
Ethnicity unknown	14	293	50	21%	43%	14%	21%	0%	55%	54%	41%

Appendix 7.D—Types of Score Reports

Table 7.D.1 Score Reports Reflecting CMA Results

2009 STAR CMA PRINTED REPORTS	
DESCRIPTION	**DISTRIBUTION**
The CMA Student Report	
This report provides parents/guardians and teachers with the student's results, presented in tables and graphs. It mainly includes: • Scale scores • Performance levels • Number and percent correct in each reporting cluster • Comparison of the student's scores on specific content areas (reporting clusters) to the scores of students statewide who scored at the lowest score for proficient and the lowest score for advanced on the total test	This report includes individual student results and is not distributed beyond parents/guardians and the student's school. Two color copies of this report are provided for each student: One is for the student's current teacher, and one is to be distributed to parents/ guardians by the district.
Student Record Label	
These reports are printed on adhesive labels to be affixed to the student's permanent school records. Each pupil shall have an individual record of accomplishment; that includes STAR testing results (see California *Education Code* section 60607(a)). Significant information includes: • Scale scores • Performance levels	This report includes individual student results and is not distributed beyond the student's school.
Student Master List	
This report is an alphabetical roster of individual student results. It mainly includes: • Percent correct for each reporting cluster within each subject area tested • A scale score and a performance level for each subject area tested • Writing score (grades four and seven)	This report provides administrators and teachers with a quick reference to all students' results within each grade Because this report includes individual student results it is not distributed beyond the student's school.

2009 STAR CMA PRINTED REPORTS	
DESCRIPTION	**DISTRIBUTION**
Student Master List Summary	
This report summarizes student results at the school, district, county, and state level for each grade. It does *not* include any individual student information. The following data are summarized by subject tested: • Number of students enrolled, number and percent of students tested, number and percent of valid scores, and number tested with scores • Mean percent correct, mean scale score, and scale score standard deviation for each subject area tested (grades three through five) • Number and percent of students scoring at each performance level (grades three through five) • The number of items for each reporting cluster and the mean percent correct (grades three through five) • For the CMA for grades four and seven only, the percent of students achieving each Writing Application score	This report is a resource for evaluators, researchers, teachers, parents/guardians, community members, and administrators. One copy is sent to the school and one to the district. This report is also produced for districts, counties, and the state. ***Note:*** The data on this report may be shared with parents/guardians, community members, and the media only if the data are for 11 or more students.
Subgroup Summary	
This set of reports disaggregates and reports results by the following subgroups: • All students • Disability status • Economic status • Gender • English proficiency • Primary ethnicity These reports contain no individual student-identifying information and are aggregated at the school, district, county, and state level. For each subgroup within a report, and for the total number of students, the following are included: • Total number tested in the subgroup • Percent tested in subgroup as a percent of all students tested • Number and percent of valid scores • Number tested who received scores • Mean scale score (grades three through five) • Standard deviation of scale score (grades three through five) • Number and percent of students scoring at each CMA performance level (grades three through five) • Percent correct for each content area (six through eight)	This report is a resource for evaluators, researchers, teachers, parents/guardians, community members, and administrators. One copy is sent to the school and one copy to the district. This report is also produced for districts, counties, and the state. ***Note:*** The data on this report may be shared with parents/guardians, community members, and the media only if the data are for 11 or more students.

2009 STAR CMA PRINTED REPORTS	
DESCRIPTION	**DISTRIBUTION**
Subgroup Summary—Ethnicity for Economic Status	
This report, a part of the Subgroup Summary, disaggregates and reports results by cross-referencing each ethnicity with economic status for CMA students in grades three through five only. The economic status for each student is "economically disadvantaged," "not economically disadvantaged," or "economic status unknown." A student is defined as "economically disadvantaged" if both parents have not received a high school diploma or the student is eligible to participate in the free or reduced-price lunch program also known as the National School Lunch Program (NSLP). For each subgroup within a report, and for the total number of students, the following data are included: • Total number tested in the subgroup • Percent tested in subgroup as a percent of all students tested • Number and percent of valid scores • Number tested who received scores • Mean scale score (grades three through five) • Standard deviation of scale score (grades three through five) • Number and percent of students scoring at each CMA performance level (grades three through five)	This report is a resource for evaluators, researchers, teachers, parents/guardians, community members, and administrators. One copy is sent to the school, and one copy to the school district. This report is also produced for school districts, counties, and the state. ***Note:*** The data on this report may be shared with parents/guardians, community members, and the media only if the data are for 11 or more students.

Chapter 8: Analyses

This chapter summarizes the item-level and test-level statistics obtained during the administration of the CMA during the spring of 2009.

The statistics presented in this chapter are divided into five sections in the following order:

1. Classical Item Analyses
2. Reliability Analyses
3. Analyses in Support of Validity Evidence
4. Item Response Theory (IRT) Analyses
5. Differential Item Functioning (DIF) Analyses

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

1. Appendix 8.A presents the classical item analyses including *p*-value and point-biserial for each item in each operational test. The appendix also presents information about the distribution of scores on the Writing tasks administered in grades four and seven for the overall population and for the various subgroups. In addition, the average and median *p*-value and point-biserial for the operational tests are presented in Table 8.2 on page 128.
2. Appendix 8.B presents results of the reliability analyses of total test scores and subscores for the population as a whole and for selected subgroups. Also presented are results of the analyses of the consistency and accuracy of the performance classifications. Finally, inter-rater reliability results for the Writing tasks administered at grades four and seven are shown.
3. Appendix 8.C presents tables showing the correlation between scores obtained in the different content areas measured by the CMA. They are provided as an evidence of the validity of the interpretation and uses of CMA scores. The results for the overall test population are presented in Table 8.5. The results for various subgroups are summarized in Appendix 8.C.
4. Appendix 8.D presents the results of IRT analyses including the distribution of items based on their fit to the Rasch model. This appendix also includes summaries of Rasch item difficulty statistics (*b*-values) for the operational and field-test items. In addition, this appendix presents the raw-to-scale score conversion tables obtained for the operational CMA tests for grades three through five.
5. Appendix 8.E presents the results of the DIF analyses applied to all operational and field-test items for which sufficient student samples were available. In this appendix, items flagged for significant DIF are listed. Also given are the distributions of items across DIF categories.

Samples Used for the Analyses

The analyses were conducted at different times in the testing process and involved varying proportions of the full CMA data. The classical item-analysis presented in Appendix 8.A, the reliability statistics included in Appendix 8.B, the content area correlations presented in Appendix 8.C, the IRT results presented in Appendix 8.D, and the item-level DIF results presented in Appendix 8.E were calculated using the P1 data file. This file contained more than 99 percent of the entire test-taking population. For the raw-to-scale score conversion

table in Appendix 8.D, the sample of all valid cases available by early June were used. Summary statistics describing the P1 and calibration samples are presented in Table 8.1.

Table 8.1 Summary Statistics for P1 and Calibration Samples

		P1			Calibration Sample			
Subject	**Grade**	**N**	**Mean**	**SD**	**N**	**% of P1**	**Mean**	**SD**
English–Language Arts	3	11,321	27.07	8.64	3,445	30%	27.9	8.92
	4	17,121	25.56	7.84	4372	26%	26.43	8.17
	5	17,821	28.78	7.61	4,815	27%	28.52	7.76
Mathematics	3	11,227	29.03	8.87	2,860	25%	29.12	8.92
	4	14,415	26.77	7.28	3,893	27%	26.47	7.21
	5	15,656	28.37	8.06	4,205	27%	27.87	8.00
Science	5	16,681	29.41	7.29	4,377	26%	29.50	7.35

Classical Analyses

Multiple-Choice Items

Classical item statistics that included overall and item-by-item proportion correct indices and the point-biserial correlation indices were computed for the operational items. The point-biserial correlation is a special case of the Pearson product-moment correlation used to measure the strength of the relationship between two variables, one dichotomous and one continuously measured—in this case, the item score (right/wrong) and the total test score. The formula for the Pearson product-moment correlation is:

$$R_{it} = \frac{Cov(i,t)}{\sigma_{xi}\sigma_t} \tag{8.1}$$

where,

Cov(i,t) is the covariance between an item i and total score t

σxi is the standard deviation for an item i

σt is the standard deviation for t

The classical statistics for the overall test are presented in Table 8.2. The item-by-item values for the indices are presented in Table 8.A.1 through Table 8.A.3, which start on page 150.

Table 8.2 Average and Median Proportion Correct and Point-Biserial

				Mean		Median	
Subject	**CMA**	**No. of Items**	**No. of Examinees**	***p*-value**	**Pt-Rbis**	***p*-value**	**Pt-Rbis**
English–Language Arts	3	48	11,321	0.56	0.47	0.58	0.49
	4	48	17,121	0.53	0.43	0.54	0.46
	5	48	17,821	0.60	0.44	0.58	0.45
	6	54	16,892	0.52	0.36	0.52	0.39
	7	54	13,908	0.54	0.40	0.54	0.40
	8	54	13,162	0.53	0.40	0.54	0.40

				Mean		Median	
Subject	CMA	No. of Items	No. of Examinees	*p*-value	Pt-Rbis	*p*-value	Pt-Rbis
Mathematics	3	48	11,227	0.61	0.5	0.64	0.54
	4	48	14,415	0.56	0.42	0.54	0.43
	5	48	15,656	0.59	0.46	0.59	0.49
	6	54	15,278	0.53	0.35	0.52	0.37
	7	54	13,145	0.46	0.31	0.46	0.31
Science	5	48	16,681	0.61	0.42	0.61	0.45
	8	54	11,384	0.52	0.37	0.51	0.38

Writing Tasks

As described earlier, students in grades four and seven were administered two different Writing prompts at different times in the STAR testing cycle. Students were given one of the two writing prompts depending upon when they were tested.

The distributions of Writing scores for the overall population and for the various subgroups are presented in Table 8.A.4 and Table 8.A.5, which appear on pages 153 and 154. The subgroups include gender, ethnicity, economic status, primary disability, and language proficiency.

The mean scores obtained on the Writing tests for the overall population and for various subgroups are presented in Table 8.A.6 and Table 8.A.7. To quantify the differences between mean scores of subgroups, effect sizes were calculated; Cohen's *d* (1992) was used as the measure of effect size. Cohen's *d* is defined as the difference between two means divided by the pooled standard deviation adjusted for sample size. Cohen (1992) suggested that 0.2 is indicative of a small effect, 0.5 a medium, and 0.8 a large effect size. The effect sizes for various subgroup differences are presented in Table 8.A.8.

Reliability Analyses

Reliability focuses on the extent to which differences in test scores reflect true differences in the knowledge, ability, or skill being tested, rather than fluctuations due to chance or random factors. 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 score variance) and partly due to random unsystematic 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 score variance. Several different ways of estimating this proportion exist. The estimates of reliability reported here 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 students' state of health or testing environment. Reliability coefficients may 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 were retested. The formula for the internal consistency reliability as measured by Cronbach's Alpha (Cronbach, 1951) is reported on the next page:

$$\alpha = \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_t^2}\right] \tag{8.2}$$

where,

n is the number of items,

σ_i^2 is the variance of scores on the i-th item, and

σ_t^2 is the variance of the total score (either the total raw score or scale score).

The standard error of measurement (SEM) provides a measure of score instability in the score metric. The SEM was computed as follows:

$$\sigma_e = \sigma_t\sqrt{1-\alpha} \tag{8.3}$$

where,

α is the reliability estimated using (8.2) above, and

σ_t is the standard deviation of the total raw scores.

The SEM is particularly useful in determining the confidence interval (CI) that captures an examinee's true score. Assuming that measurement error is normally distributed, it can be said that upon infinite replications of the testing occasion, approximately 95 percent of the CIs of $\pm$1.96 SEM around the observed score would contain an examinee's true score (Crocker & Algina, 1986). For example, if an examinee's observed score on a given test equals 15 points, and the SEM equals 1.92, one can be 95 percent confident that the examinee's true score lies between 11 and 19 points (15 $\pm$ 3.76 rounded to the nearest integer).

Table 8.3 gives the reliability for each of the operational CMA tests, along with the number of items and examinees upon which those analyses were performed.

Table 8.3 Reliabilities and SEMs for the CMA

					Scale Score[1]			Raw Score		
Subject	**CMA**	**No. of Items**	**No. of Examinees**	**Reliability**	**Mean**	**Std. Dev.**	**SEM**	**Mean**	**Std. Dev.**	**SEM**
English–Language Arts	**3**	48	11,321	0.87	298	65	23.65	27.07	8.64	3.14
	4	48	17,121	0.83	316	72	29.53	25.56	7.84	3.20
	5	48	17,821	0.84	320	69	28.01	28.78	7.61	3.08
	6	54	15,900	0.78	–	–	–	28.30	7.37	3.42
	7	54	13,060	0.83	–	–	–	29.14	8.40	3.43
	8	54	12,434	0.82	–	–	–	28.57	8.06	3.40
Mathematics	**3**	48	11,227	0.89	315	70	23.75	29.03	8.87	3.00
	4	48	14,415	0.81	321	80	34.56	26.77	7.28	3.14
	5	48	15,656	0.85	325	77	29.47	28.37	8.06	3.08
	6	54	15,278	0.77	–	–	–	28.71	7.29	3.47
	7	54	13,145	0.70	–	–	–	25.07	6.45	3.51
Science	**5**	48	16,681	0.82	335	58	24.59	29.41	7.29	3.08
	8	54	11,384	0.79	–	–	–	28.12	7.47	3.46

[1] Scale scores were not available for grades six through eight during the spring 2009 administration.

Intercorrelations, Reliabilities, and SEMs for Reporting Clusters

For each CMA, number-correct scores are computed for the three reporting clusters in each operational test for which subscores are reported in grades three to five tests. Intercorrelations and reliability estimates for these clusters are presented in Table 8.C.1 through Table 8.C.12 starting on page 182. Consistent with previous years, the reliabilities across reporting clusters varied significantly according to the number of items in each cluster.

Subgroup Reliabilities and SEMs

The reliabilities of the ELA and mathematics CMA tests in grades three through five and the science CMA in grade five were examined for various subgroups of the examinee population. The subgroups included in these analyses were defined by their gender, ethnicity, economic status, primary disabilities, and language proficiency. As of 2009, reliability analyses are also presented for the subgroups categorized by whether or not examinees belonging to a certain ethnic subgroup were economically disadvantaged for students in grades three through eight.

For each subgroup analysis, reliability and SEM information is reported for the total test scores and also for the cluster scores. The reliability and SEM information for the cluster scores are reported for only grades three through five tests. Table 8.B.1 through Table 8.B.3 present the overall reliabilities. Table 8.B.4 through Table 8.B.7 present the reliabilities for the subgroups based on gender, economic status, English-language fluency, and primary ethnicity. The next set of tables, Table 8.B.8 through Table 8.B.10, show the same analyses for the subgroups based on primary ethnicity and economic status, and gender and economic status. Table 8.B.11 and Table 8.B.12 present the reliabilities for primary disabilities.

The overall reliabilities for the various subgroups are compiled in Table 8.B.13 through Table 8.B.19. Table 8.B.20 through Table 8.B.29 present the cluster-level reliabilities for the various subgroups in grades three through five. Note that the reliabilities are reported only for samples that are comprised of 11 or more examinees.

Conditional Standard Errors of Measurement

As part of the IRT-based equating procedures, scale score conversion tables and conditional standard errors of measurement (CSEMs) are produced. CSEMs for CMA scale scores are based on item response theory and are calculated by the IRTEQUATE module in GENASYS.

The CSEM is estimated as a function of measured ability. It is typically smaller in scale score units toward the center of the scale in the test metric where more items are located and larger at the extremes where there are fewer items. An examinee's CSEM under the IRT framework is equal to the inverse of the square root of the test information function:

$$\mathrm{CSEM}(\hat{\theta}) = \frac{1}{\sqrt{\mathrm{I}(\theta)}} a, \tag{8.4}$$

where

$\mathrm{CSEM}(\hat{\theta})$ is the standard error of measurement and

$\mathrm{I}(\theta)$ is the test information function.

The statistic is multiplied by a, where a is the original scaling factor needed to transform theta to the scale score metric. The value of a varies by grade and subject.

Standard errors of measurement vary across the scale. When a test has cut scores, it is important to provide CSEMs at the cut scores.

Table 8.4 presents the scale score CSEMs at the lowest score required for a student to be classified in the below basic, basic, proficient, and advanced performance levels for each CMA.

The CSEMs tended to be higher at the advanced cut points for all tests. The pattern of lower values of CSEMs at the basic and proficient levels are expected since (1) more items tend to be of middle difficulty; and (2) items at the extremes still provide information toward the middle of the scale. This results in more precise scores in the middle of the scale and less precise scores in the extremes of the scale.

Table 8.4 Scale Score CSEM at Cut Points

		Below Basic		Basic		Proficient		Advanced	
Subject Area	**CMA**	**Min SS**	**CSEM**	**Min SS**	**CSEM**	**Min SS**	**CSEM**	**Min SS**	**CSEM**
English–Language Arts	3	234	22	300	22	350	24	397	28
	4*	250	28	300	27	350	28	407	31
	5	220	27	300	26	350	28	402	31
* Including writing prompt									
Mathematics	3	230	22	300	21	350	24	425	32
	4	229	34	300	32	350	33	431	37
	5	234	27	300	26	350	28	422	33
Science	5	250	24	300	23	350	24	403	27

Decision Classification Analyses

The methodology used for estimating the reliability of classification decisions is described in Livingston and Lewis (1995) and is implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.14). RELCLASS-COMP estimates decision consistency using an estimated multivariate distribution of reported classifications on the current form of the exam and classifications on a hypothetical alternate form using the reliability of the test and strong true score theory. RELCLASS-COMP also 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 a cut score is estimated by collapsing the multivariate distribution at the passing score boundary into an *n* by *n* table (where *n* is the number of performance levels) and summing the entries in the diagonal. Figure 8.1 and Figure 8.2 display the tables used in the analyses.

Figure 8.1 Decision Accuracy for Achieving a Performance Level

		Decision made on a form actually taken	
		Does not achieve a performance level	Achieves a performance level
True status on all-forms average	Does not achieve a performance level	Correct classification	Mis-classification
	Achieves a performance level	Mis-classification	Correct classification

Figure 8.2 Decision Consistency for Achieving a Performance Level

		Decision made on the second form taken	
		Does not achieve a performance level	Achieves a performance level
Decision made on the first form taken	Does not achieve a performance level	Correct classification	Mis-classification
	Achieves a performance level	Mis-classification	Correct classification

The results of these analyses for the CMA that include performance classifications are presented in Table 8.B.30 through Table 8.B.37 in Appendix 8.B. These tables are available for the ELA and mathematics multiple-choice tests in grades three through five and the science multiple-choice test in grade five. For grade four ELA, there is an additional table that includes both the multiple-choice and Writing Application tests.

Each table includes the contingency tables for the various performance-level classifications. The proportion of accurately classified students is determined by summing across the diagonals of the upper tables, and the proportion of consistently classified students is determined by summing the diagonals of the lower tables.

The classifications are collapsed to below-proficient versus proficient and above, which are the critical categories for AYP analyses, and these results are also presented in the tables.

Writing Score Reliability

The reliability of the total scores for the students who took the Writing prompt for ELA in grades four and seven were computed using the following composite reliability formula (Feldt & Brennan, 1989):

$$\alpha_c = 1 - \frac{\sum_{j=1}^{k} \left(w_j \sigma_{e_j} \right)^2}{\sigma_c^2} \tag{8.5}$$

where,

k is the number of part scores in the composite,

w_j is the weight associated with the j-th part score,

σ_{e_j} is the SEM of the j-th part score, and

σ_c^2 is the variance of the composite score.

The reliability of the Writing tasks can be found indirectly by examining the correlation between the multiple-choice (MC) and Writing components in relation to the MC reliability. The lower bound reliability for a constructed response (CR) item in a test with multiple-choice items and only one CR item can be found using the squared correlation between the MC and Writing (CR) portions of the test and dividing by the reliability of the MC portion of the test ($(Corr_{Writing\text{-}MC})^2/Rel_{MC}$) (Sax, 1989).

The SEM for Writing (i.e., the CR portion of the test) can then be found using the following equation:

$$\sigma_{e_{CR}} = \sigma_{CR} * \sqrt{1 - reliability_{CR}} \tag{8.6}$$

where,

σ_{CR} is the standard deviation of the Writing scores

Prompt and Rater Agreement Summary

As described earlier, in order to monitor the accuracy of ratings, two raters scored approximately 10 percent of the examinees' Writing responses. The two sets of ratings were used to carry out inter-rater agreement and generalizability analyses to assess the reliability of the writing scores.

Inter-rater Reliability Analyses

The inter-rater reliability or consistency is defined as the degree to which the measure assessed by two or more raters is in agreement when the same instrument is administered to the same person. It is an indicator of homogeneity and is most frequently measured using intraclass correlation which looks at the exact agreement between raters over and above that expected by chance. The index is defined as,

$$ICC = r_I = (ms_{between} - ms_{within})/(ms_{between} + [k - 1]ms_{within}) \tag{8.7}$$

where,

$ms_{between}$ is the mean-square estimate of between-subjects variance, and

ms_{within} is the mean-square estimate of within-subjects variance.

For categorical ratings, Cohen's Kappa statistic has the properties of an intraclass correlation coefficient and can be used for inter-rater reliability. Cohen's Kappa was therefore used as a primary indicator of the inter-rater reliability of the Writing scoring. In addition, the percentages of ratings on which the raters were in exact agreement or differed by just one point were also computed.

The reliability analyses were performed on approximately ten percent of the overall testing population that were scored by two raters. From those double-scored Writing responses, samples were selected that had valid ratings (1, 2, 3, or 4) for both raters. The results of these analyses are presented in Table 8.B.38 and Table 8.B.39 in Appendix 8.B.

Generalizability Analyses

Generalizability analyses were performed on the Writing scores to quantify the proportion of variance explained by various possible sources of variation including raters, writing prompt, and persons (desired variance). A generalizability study (g-study) was performed to estimate variance components for selected sources of variation also known as "facets." A d-study was performed to estimate the generalizability coefficient (Brennan, 2001; Crocker & Algina, 1986).

The well-known computer programs, GENOVA, as well as its extension, urGENOVA, were used to carry out these analyses (Brennan, 2001; Crick & Brennan, 1983). Since two raters scored each student's response but each student did not receive the same prompt, a nested unbalanced design was studied (Lee & Kantor, 2007; Wang et al., 2007) as described below:

Design = (Person : Task) x Rater

The model assumes that the raters are selected from an infinite pool of raters and all the raters are randomly equivalent. The model also assumes that the Writing prompts are randomly selected from a universe of prompts and that students' writing responses are randomly assigned to the raters. The results of the study are presented in the in the subsection "Generalizability Analyses for Writing" tables on page 139.

Validity Evidence

Validity refers to the degree to which each interpretation or use of a test score is supported by evidence that is gathered (APA, AERA, & NCME, 1999; ETS, 2002). It is a central concern underlying the development, administration, and scoring of a test and the uses and interpretations of test scores.

Validation is the process of accumulating evidence to support each proposed score interpretation or use. It does not involve a single study or gathering one particular kind of evidence. Validation involves multiple investigations and various kinds of evidence (APA, AERA, & NCME, 1999; Cronbach, 1971; ETS, 2002; Kane, 2006). The process begins with test design and continues through the entire assessment process including item development and field testing, analyses of item and test data, test scaling, scoring, and score reporting.

This section presents the evidence gathered to support the intended uses and interpretations of scores for the CMA testing program. The description is organized in the manner prescribed by APA, AERA, and NCME's *The Standards for Educational and Psychological Testing* (1999). These standards require a clear definition of the purpose of the test, which includes a description of the qualities, called constructs, that are to be assessed by a test, the population to be assessed, as well as how the scores are to be interpreted and used.

In addition, the *Standards* identify five kinds of evidence that can provide support for score interpretations and uses, which are as follows:

1. Evidence based on test content;
2. Evidence based on relations to other variables;
3. Evidence based on response processes;
4. Evidence based on internal structure; and
5. Evidence based on the consequences of testing.

These kinds of evidence are also defined as important elements of validity information in documents developed by the U.S. Department of Education for the peer review of testing programs administered by states in response to the Elementary and Secondary Education Act (USDOE, 2009; see Appendix A).

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

Purpose of the CMA

As mentioned in Chapter 1, the CMA tests contribute to calculating school and district API. Additionally, the CMA for ELA and mathematics in grades three through five are used in determining AYP that applies toward meeting the requirement of the federal Elementary and Secondary Education Act (ESEA), which is to have all students score proficient or above by 2014.

The Constructs to Be Measured

The CMA tests, administered in English, are designed to show how well students perform relative to the California content standards. These content standards were approved by the SBE; they describe what students should know and be able to do at each grade level.

Test blueprints and specifications written to define the procedures used to measure the content standards provide an operational definition of the construct to which each set of standards refers—that is, they define, for each subject area to be assessed, the tasks to be presented, the administration instructions to be given, and the rules used to score examinee responses. They control as many aspects of the measurement procedure as possible so that the testing conditions will remain the same over test administrations (Cronbach, 1971; Cronbach, Gleser, Nanda, & Rajaratnam, 1972) to minimize construct irrelevant score variance (Messick, 1989). The content blueprints for the CMA can be found on the CDE Web page at http://www.cde.ca.gov/ta/tg/sr/cmablueprints.asp. ETS has developed all CMA test items to conform to the SBE-approved content standards and test blueprints.

The Scores Generated and the Interpretations and Uses of these Scores

Total scores expressed as scale scores, student performance levels, and subscores for each reporting cluster are generated for the ELA and mathematics tests in grades three through five and the science test in grade five. For the other tests in grades six through eight, students' total scores were expressed in the raw score metric and no performance levels were provided. On the basis of a student's total score, an inference is drawn about how much knowledge and skill in the subject area the student has. The total score also is used to classify students in terms of their level of knowledge and skill in the subject area. These levels are called performance levels and are as follows: advanced, proficient, basic, below basic, and far below basic.

Reporting cluster scores, also called subscores, are used to draw inferences about a student's achievement in each of several specific knowledge or skill areas covered by each test. Reporting cluster results compare an individual student's percent-correct score to the average percent correct for the state as a whole. The range of scores for students who scored proficient is also provided for each cluster using a percent correct metric. The reference points for this range are: (1) the average percent correct for students who received the lowest score qualifying for the proficient category; and (2) the average percent correct for students who received the lowest score qualifying for the advanced category. A detailed description of the uses and applications of CMA scores is presented in Chapter 7.

The tests that make up the CMA program provide results or score summaries that are used for different purposes. The four major purposes are:

1. Communicating with parents and guardians;
2. Informing decisions needed to support student achievement;
3. Evaluating school programs; and
4. Providing data for state and federal accountability programs for schools.

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

Intended Test Population(s)

California public school students who meet certain eligibility criteria are the intended test population for the CMA. Students in grades three through eight are tested in ELA; students in grades three through seven are tested in mathematics. In addition, students in grades five and eight take a science test.

Validity Evidence Collected

Evidence Based on Content

According to the APA, AERA, and NCME's *The Standards for Educational and Psychological Testing* (1999), analyses that demonstrate a strong relationship between a test's content and the construct that the test was designed to measure can provide important evidence of validity. In current K–12 testing, the construct of interest usually is operationally defined by state content standards and the test blueprints that specify the content, format, and scoring of items that are admissible measures of the knowledge and skills described in the content standards. Evidence that the items meet these specifications and represent the domain of knowledge and skills referenced by the standards supports the inference that students' scores on these items can appropriately be regarded as measures of the intended construct.

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

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

Description of the state standards—As was noted in Chapter 1, the SBE adopted rigorous content standards in 1997 and 1998 four major content areas: ELA, history–social science, mathematics, and science. These standards were designed to guide instruction and learning for all students in the state and to bring California students to world-class levels of achievement.

Specifications and blueprints—ETS maintains item development specifications for each CMA. The item specifications describe the characteristics of the items that should be written to measure each content standard. A thorough description of the specifications can be found in Chapter 3, staring on page 60. Once the items are developed, ETS selects all CMA test items to conform to the SBE-approved California content standards and test blueprints. Test blueprints for the components of the CMA were proposed by ETS and reviewed and approved by the Assessment Review Panel (ARPs). The ARPs are

advisory panels to the CDE and ETS on areas related to item development for the CMA. They were also reviewed and approved by the CDE and presented to the SBE for adoption. There have been no recent changes in the blueprints for the CMA. The content blueprints for the CMA can be found on the CDE "STAR CMA Blueprints" Web page at http://www.cde.ca.gov/ta/tg/sr/cmablueprints.asp.

Item development process—A detailed description of the content and psychometric criteria applicable to the construction of 2009 CMA is presented in Chapter 4, starting on page 71.

Item review process—Chapter 3 explains in detail the extensive item review process applied to items written for use in CMA content-area tests. In brief, items written for the CMA undergo multiple review cycles and involve multiple groups of reviewers. One of the reviews is carried out by an external reviewer, that is, the ARP. The ARP is responsible for reviewing all newly developed items for alignment to the California content standards.

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

Guidelines for test construction were established with the goal of maintaining parallel forms to the greatest extent possible from year to year. Details can be found in Chapter 4, starting on page 71.

Additionally, an external review panel, the Statewide Pupil Assessment Review (SPAR), is responsible for reviewing and approving the achievement tests to be used statewide for the testing of students in California public schools, grades two through eleven. More information about the SPAR is given in Chapter 3, starting on page 67.

Evidence Based on Relations to Other Variables

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

Differential Item Functioning Analyses

Analyses of DIF can provide evidence of the degree to which a score interpretation or use is valid for individuals who differ in particular demographic characteristics. For the CMA, DIF analyses were performed on all operational items and all field-test items for which sufficient student samples were available.

The results of the DIF analyses are presented in Appendix 8.E. Very few items exhibited significant DIF, suggesting that, in general, scores based on the CMA items would have the same meaning for individuals who differed in their demographic characteristics.

Intercorrelations Between Content Areas

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

Results in the tables appear to be consistent with expectations. In general, students' ELA scores correlated moderately with their mathematics scores. They correlated somewhat more highly with their scores on the CMA for science tests.

Table 8.C.1 throughTable 8.C.12 in Appendix 8.C provide the content area correlations by gender, ethnicity, English proficiency level, economic status, and primary disability. Similar patterns of correlations between students' ELA, mathematics, and science scores were found within the subgroups.

Table 8.5 2009 CMA Content Area Correlations (All Valid Scores)

Grade		ELA	Mathematics	Science
3	**ELA**	**11,321**	0.62	
	Mathematics	11,146	**11,227**	
4	**ELA**	**17,121**	0.54	
	Mathematics	13,772	**14,415**	
5	**ELA**	**17,821**	0.53	0.65
	Mathematics	14,677	**15,656**	0.53
	Science	16,254	14,364	**16,681**
6	**ELA**	**15,900**	0.52	
	Mathematics	13,862	**15,278**	
7	**ELA**	**13,060**	0.51	
	Mathematics	11,670	**13,145**	
8	**ELA**	**12,434**		0.59
	Science	10,940		**11,384**

Intercorrelations between the reporting clusters are presented in Table 8.B.1 through Table 8.B.3 for the CMA. In general, moderate correlations between cluster scores should be expected since, by design, the clusters measure various aspects of the same construct. The findings given in the tables show that moderate intercorrelations were obtained. As also would be expected, the intercorrelations were higher among clusters that assessed more similar skills than they were among clusters that assessed somewhat dissimilar skills. For example, in mathematics, the clusters related to number sense correlated more highly with Algebra and Data Analysis than they did with Measurement and Geometry.

Note that the correlations are reported only for samples that are comprised of more than 10 examinees.

Generalizability Analyses for Writing

Generalizability analyses were performed on students' Writing scores for grades four and seven to assess the proportion of variance explained by raters, Writing prompt, and persons. The details on the methodology are described in the subsection "Generalizability Analyses" in the "Writing Score Reliability" section on page 133. Details about the results can be found in Table 8.B.41 and Table 8.B.42.

A decision study (d-study) was conducted to look at the generalizability-coefficients (g-coefficient) for the Writing scores; the g-coefficient for grade four was 0.84 and the g-coefficient for grade seven was 0.83. The largest variance component was attributed to the "person" variation, which is the desired variation to occur among the examinee or "person" scores. Variation attributable to the construct-irrelevant rater variable was small, and the variation attributable to the prompt was negligible for both grades.

Evidence Based on Response Processes

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

Evidence of Rater Reliability, Inter-rater Agreement

Rater consistency for the writing prompt is critical to the CMA Writing scores and their interpretations. These findings provide evidence of the degree to which raters agree in their observations about the qualities evident in students' essay responses. In order to evaluate the reliability of the student scores on the writing prompts administered in grades four and seven, two raters scored approximately ten percent of the examinee responses. The data collected were used to evaluate inter-rater reliability and inter-rater agreement.

Inter-rater reliability

Cohen's Kappa statistics findings provide evidence of the degree to which a student's score may vary from rater to rater. Without explicit criteria to guide the rating process, two independent raters may not assign the same score to a given response.

The results showed moderate levels of agreement between raters that scores examinees' written responses to the prompts administered in both grades four and seven.[1]

Inter-rater agreement

As noted previously, approximately ten percent of the test population's responses to the writing prompts were scored by two raters. The percentage of students for whom the raters were in exact agreement was 68 percent in grade four and 71 percent in grade seven. Approximately 99 percent of the tests for both grade four and grade seven exhibited exact or adjacent agreement between the two sets of ratings.

Evidence Based on Internal Structure

As suggested by the *Standards,* evidence of validity can also be obtained from studies of the properties of the item scores and the relationship between these scores and scores on components of the test. To the extent that the score properties and relationships found are consistent with the definition of the construct measured by test, support is gained for interpreting these score as measures of the construct.

For the CMA, it is assumed that a single construct underlies the total scores obtained on each test. Evidence to support this assumption can be gathered from the results of item analyses, evaluations of internal consistency, and studies of model-data fit, dimensionality, and reliability.

With respect to the subscores that are reported, these scores are intended to reflect examinees' knowledge and/or skill in an area that is part of the construct underlying the total test. Analyses of the intercorrelations among the subscores themselves and between the

[1] Research has shown the value of Kappa statistics between 0.41 and 0.60 as exhibiting moderate levels of agreement between the two ratings. (Reference: http://www.niwa.co.nz/education-and-training/training-courses)

subscores and total test score can be used for this purpose. Information about the internal consistency of the items on which each subscore is based also is useful to provide.

Classical Statistics

Biserial correlations calculated for the items in a test show the degree to which the items discriminate between students with low and high scores on a test. To the degree that the correlations are high, evidence that the items assess the same construct is provided. The biserials for the items in the CMA tests are presented in Table 8.A.1 through Table 8.A.7. As shown in Table 8.2, this index was greater than or equal to 0.40 for 10 CMA tests and was between 0.31 and 0.39 for the remaining four CMA tests.

Also germane to the validity of a score interpretation are the ranges of items for the items on which a test score will be based. The finding that items have difficulties that span the range of examinee ability provides evidence that examinees at all levels of ability are adequately measured by the items. Information on item *p*-values are given in Table 8.1; the distributions of item *b*-values are given in Table 8.D.14 through Table 8.D.19. The data indicate that 13 of the 14 CMA tests had average *p*-values that ranged from 0.50 to 0.61, and they had a range of item *b*-values, as shown in Table 8.D.14 through Table 8.D.19.

Reliability

Reliability is a prerequisite for validity. The finding of reliability in student scores supports the validity of the inference that the scores reflect a stable construct. This section will describe briefly findings concerning the total test level, as well as reliability results for the reporting clusters.

Overall reliability—The reliability analyses on each of the operational CMA tests are presented in Table 8.3. The results indicate that the reliabilities of the CMA for ELA, mathematics, and science tests were moderately high, ranging from 0.70 to 0.89.

Subgroup reliabilities—The reliabilities of the operational CMA tests were also examined for various subgroups of the examinee population that differed in their demographic characteristics. The characteristics considered were gender, ethnicity, economic status, primary disabilities, language proficiency, and ethnicity-by-economic status. The results of these analyses can be found in Table 8.B.4 though Table 8.B.19.

Reporting cluster reliabilities—For each CMA, number-correct scores are computed for the reporting clusters The reliabilities of these scores are presented in Table 8.B.20 through Table 8.B.29 for the 14 CMA tests. The reliabilities of reporting clusters are invariably lower than those for the total tests because they are based on very few items. Cluster reliabilities are also affected by the number of items in each cluster, with cluster scores based on fewer items having somewhat lower reliabilities than cluster scores based on more items. Because the reliabilities of scores at the cluster level are lower, schools supplement the score results with other information when interpreting the results.

Reliability of Performance Classifications—The methodology used for estimating the reliability of classification decisions is described in the section “Decision Classification Analyses” on page 132. The results of these analyses are presented in Table 8.B.30 through Table 8.B.37 in Appendix 8.B, which start on page 177. When the decisions are collapsed to below proficient versus proficient and above, which are the critical categories for AYP analyses, the proportion of students that were classified accurately ranged from 0.87 to 0.92 across the CMA tests for which performance level were available. Similarly, the proportion of students that were classified consistently ranged from 0.82 to 0.89 for students classified into below proficient versus proficient and advanced.

Dimensionality

Dimensionality analyses were conducted by a CDE psychometrics team (Internal memorandum, September 11, 2009). The study investigated the factor structures of the CMA at grades three and five as part of ESEA for peer review.

Two factors corresponding to the ELA and mathematics domain were found for the CMA in these grades, as would be expected, since these tests were designed to measure different constructs.

Evidence Based on Consequences of Testing

As observed in the *Standards*, tests are usually administered "with the expectation that some benefit will be realized from the intended use of the scores" (p. 18). When this is the case, evidence that the expected benefits accrue will provide support for intended use of the scores. The CDE and ETS are in the process of determining what kinds of information can be gathered to assess the consequences of administration of the CMA.

IRT Analyses

ETS uses a computer system called the Generalized Analysis System (GENASYS) for the IRT item calibration. As part of this system, a proprietary version of the PARSCALE computer program (Muraki & Bock, 1995) is used and parameterized to result in one-parameter calibrations. Research at ETS has suggested that PARSCALE calibrations done in this manner produce results that are virtually identical to results based on WINSTEPS (Way, Kubiak, Henderson, & Julian, 2002). The details on all equating procedures are given in Chapter 2 starting on page 8

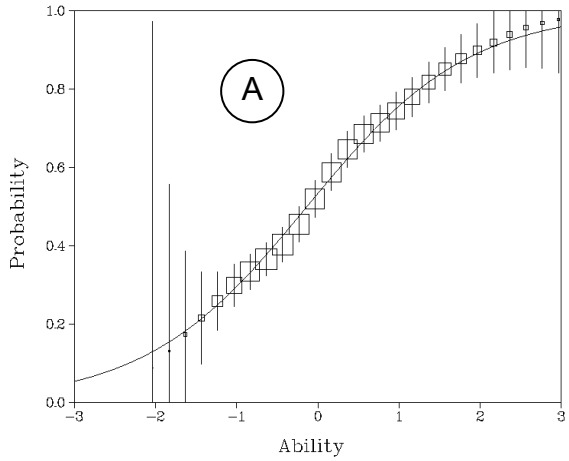
The estimation process is constrained by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588) and by setting the lower asymptote for all multiple-choice items to zero. The resulting estimation is equivalent to the Rasch model for multiple-choice items and the Rasch partial credit model for polytomously scored items, which was used to obtain calibrations for the essays in the CMA for ELA tests administered in grades four and seven . This approach is in keeping with previous CMA calibration procedures accomplished using the WINSTEPS program (Linacre, 2000).

The PARSCALE calibrations are run in two stages, following procedures used with other ETS testing programs. In the first stage, estimation imposed normal constraints on the updated prior ability 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 (EM) cycle with no constraints. For both stages, the metric of the scale is controlled by the constant discrimination parameter.

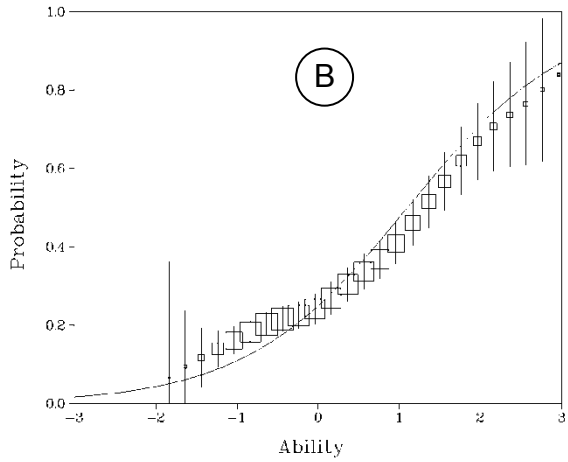
IRT Model-Data Fit Analyses

Because the Rasch model is used in equating the CMA, an important part of IRT analyses is the assessment of model-data fit. ETS psychometricians classify operational and field-test items for the CMA into discrete categories based on an evaluation of how well each item was fit by the Rasch model. The flagging procedure has categories of A, B, C, D, and F that are assigned based on an evaluation of graphical model-data fit information. Descriptors for each category are provided on the next page. As an illustration, the IRT item characteristic curves and empirical data (item-ability regressions) for five items field-tested in 2005 are shown in Figure 8.3. These five items represent the various rating categories. The item number in the calibration and ETS identification number for each item ("accession number") are listed next to each item, as well as the corresponding rating categories.

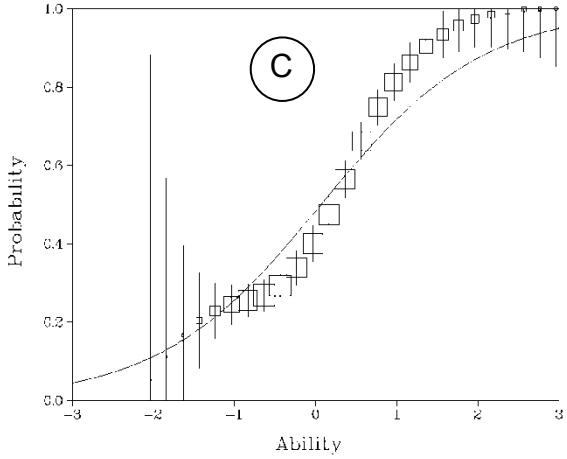
Figure 8.3 Items from the 2005 History–Social Science Grade Ten Field-Test Calibration



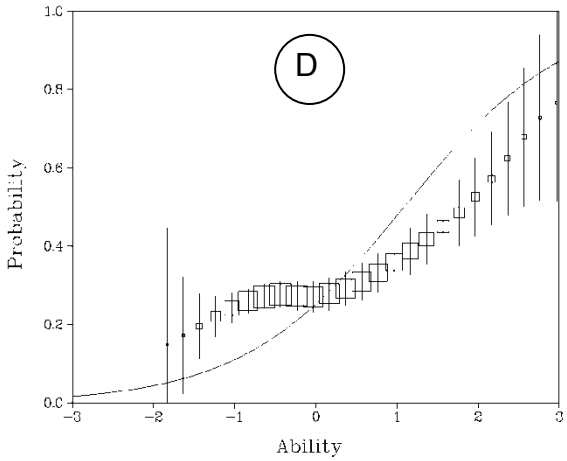
Version 30, Seq 29 (#236) CSV23487 4-Choice P+ = 0.563
a = 0.588 F, b = –0.135, c = 0.000 F, CHI = 5.41, N = 5,912



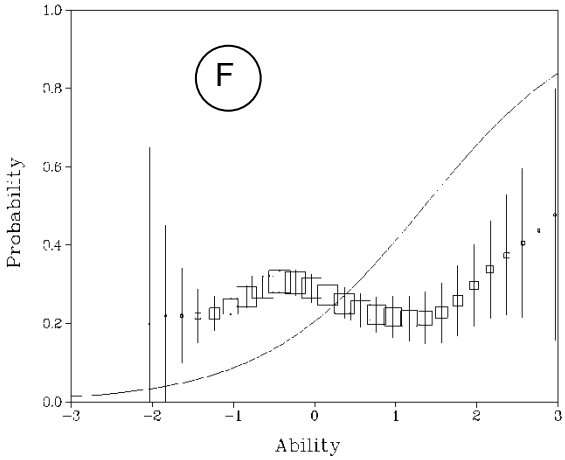
Version 1, Seq 28 (#61) CSV22589 4 Choice P+ = 0.307
a = 0.588 F, b = 1.104, c = 0.000 F, CHI = 66.70, N = 6,348



Version 18, Seq 30 (#165) CSV20282 4-Choice P+ = 0.523
a = 0.588 F, b = 0.066, c = 0.000 F, CHI = 208.99, N = 6,183



Version 9, Seq 32 (#113) CSV20317 4-Choice P+ = 0.314
a = 0.588 F, b = 1.089, c = 0.000 F, CHI = 361.31, N = 6,047



Version 21, Seq 31 (#184) CSV20311 4-Choice P+ = 0.263
a = 0.588 F, b = 1.356, c = 0.000 F, CHI = 1027.57, N = 6,277

Flag A (Item 236, CSV23487)

- 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 061, CSV22589)

- 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 165, CSV20282)

- 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 113, CSV20317)

- Theoretical curve outside of error range at some regions across ability range
- Large Chi-square value relative to the other items in the calibration with similar sample sizes

Flag F (Item 184, CSV20311)

- Theoretical curve outside of error range at most regions across ability range
- Probability of answering item correctly may be higher at lower ability than higher ability (U-shaped empirical curve)
- Very large Chi-square value relative to the other items with similar sample sizes and classical item statistics tend also to be very poor.

In general, items with flagging categories of A, B, or C are all considered acceptable. Ratings of D are considered questionable, and the ratings of F indicate a poor model fit.

Model Fit Assessment Results

The model fit assessment is performed twice in the administration cycle. The assessment is first performed before scoring tables are produced and released. The assessment is performed again as part of the final item analyses when much larger samples are available. The flags produced as a result of this assessment are placed in the item bank. The test developers are asked to avoid the items flagged as D if possible and to carefully review them if they must be used. Test developers are instructed to avoid using items rated F for operational test assembly without a review by a psychometrician.

The distributions of the operational items across the IRT model data fit classifications are presented in Table 8.D.1 through Table 8.D.3. The distributions for the field test items are presented in Table 8.D.4 through Table 8.D.6.

Summaries of Scaled IRT b-values

Once the IRT *b*-values are placed on the item bank scale, analyses are performed to assess the overall test difficulty and the distribution of items in a particular range of item difficulty.

Table 8.D.7 through Table 8.D.13 present univariate statistics (mean, standard deviation, minimum, and maximum) for the scaled IRT *b*-values. The results for the overall test are presented separately for the operational items and the field test items.

Scaling Results

Complete raw-to-scale score conversion tables for the 2009 CMA are presented in Table 8.D.20 through Table 8.D.28 in Appendix 8.D, starting on page 195. The raw scores and corresponding unrounded converted scale scores are listed in those tables. For the 2009 CMA, scale scores were truncated at both ends of the scale so that the minimum reported scale score was 150 and the maximum reported scale score was 600. The scale scores defining the cut scores for all performance categories are presented in Table, which is on page 87 in Chapter 6.

DIF Analyses

Analyses of differential item functioning (DIF) assess differences in the item performance of groups of students that differ in their demographic characteristics.

DIF analyses were performed on all operational items and all field-test items for which sufficient student samples were available. The sample size requirements for the field-test DIF analyses were 100 in the focal group and 400 in the combined focal and reference groups. These sample sizes were based on standard operating procedures with respect to DIF analyses at ETS. The DIF analyses utilized the Mantel-Haenszel (MH) DIF statistic (Mantel & Haenszel, 1959; Holland & Thayer, 1985). This statistic is based on the estimate of constant odds ratio and is described as:

The α_{MH} is the constant odds ratio taken from Dorans and Holland (1993, equation 7) and computed as:

$$\alpha_{MH} = \frac{\left(\sum_m R_{rm} \frac{W_{fm}}{N_{tm}}\right)}{\left(\sum_m R_{fm} \frac{W_{rm}}{N_{tm}}\right)} \tag{8.8}$$

$$MH\ D\text{-}DIF = -2.35 \ln[\alpha_{MH}] \tag{8.9}$$

where,

R = number right,

W = number wrong,

N = total in:

fm = focal group at ability m,

rm = reference group at ability m, and

tm = total group at ability l m.

Items analyzed for DIF at ETS are classified into one of three categories: A, B, or C. Category A contains items with negligible DIF. Category B contains items with slight to moderate DIF. Category C contains items with moderate to large values of DIF.

These categories have been used by ETS testing programs for more than 14 years. The definitions of the categories based on evaluations of the item-level MH D-DIF statistics is as follows:

DIF Category	Definition
A (negligible)	MH D-DIF not significantly different from zero, or has an absolute value less than one.
B (moderate)	MH D-DIF is significantly different from zero, and is either (1) less than 1.5; or (2) not significantly different from one.
C (large)	MH D-DIF is significantly different from one, and has an absolute value greater than 1.5.

The factors considered in the DIF analyses included gender, ethnicity, level of English proficiency[2], and primary disability. The results of the DIF analyses are presented in Appendix 8.E.

Table 8.E.1 lists the operational items exhibiting significant DIF. Table 8.E.2 represents the analogous list for the field-test items. Test developers are instructed to avoid selecting field-test items flagged as having shown DIF that disadvantage a focal group (C-DIF) for future operational test forms unless their inclusion is deemed essential to meeting test-content specifications.

Table 8.E.3 to Table 8.E.15 show the distributions of operational items across the DIF category classifications for the CMA. In these tables, classifications of B- or C- indicate DIF against a focal group; classifications of B+ and C+ indicate DIF in favor of a focal group. The last two columns of each table show the total number of items flagged for DIF in one or more comparisons. Table 8.E.16 to Table 8.E.28 provide the analogous results for the field-test items. Table 8.6 lists specific subgroups that were used for DIF analyses for the CMA.

Table 8.6 Subgroup Classification for DIF Analyses

DIF Type	Reference Group	Focal Group
Gender	Male	Female
Race/Ethnicity	White	African American American Indian Asian Combined Asian Group (Asian/Pacific Islander/Filipino) Filipino Hispanic/Latin American Pacific Islander

[2] Analyses of English learners on the CMA for ELA are presented for readers' interest; however, differential performance on an item that is due to the language difficulties of nonnative speakers does not indicate that an item is unfair or biased.

DIF Type	Reference Group	Focal Group
Disability	Specific Learning Disability	Autism Deaf-Blindness Deafness Emotional Disturbance Hard of Hearing Mental Retardation Multiple Disabilities Orthopedic Impairment Other Health Impairment Speech or Language Impairment Traumatic Brain Injury Visual Impairment

References

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

Brennan, R. L. (2001). *Generalizability theory*. New York: Springer Verlag.

Brennan, R. L. (2001). *Manual for GENOVA*. Iowa City, IA: Iowa Testing Programs, University of Iowa.

Cohen, J. A. (1992). Power primer. *Psychological Bulletin, Vol. 112,* pp. 55–159.

Crick, J. E. and Brennan, R. L. (1983). *Manual for GENOVA: A generalized analysis of variance system* (American College Testing Technical Bulletin No. 43). Iowa City, IA: ACT, Inc.

Crocker, L. and Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, Vol. 16, pp. 292–334.

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

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

Dorans, N. J. and Holland, P. W. (1993). *DIF detection and description: Mantel-Haenszel and standardization*. In P. W. Holland and H. Wainer (Eds.), *Differential item functioning*. Hillsdale, NJ: Erlbaum, pp. 35–66.

Educational Testing Service. (2002). *ETS standards for quality and fairness*. Princeton, NJ: Educational Testing Service.

Feldt, L. S. and Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement*. New York: Macmillan.

Hambleton, R. K. and Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer-Nijhoff.

Holland, P. W. and Thayer, D. T. (1985). *An alternative definition of the ETS delta scale of item difficulty*. RR-85–43.

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

Lee, Y., and Kantor, R. (2005). *Dependability of new ESL writing test scores: evaluating prototype tasks and alternative rating schemes.* TOEFL Monograph Series-31.

Livingston, S. A., and Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal on Educational Measurement*, Vol. 32, pp. 179–97.

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

Appendix 8.A—Classical Analyses

Table 8.A.1 Item-by-item *p*-value and Point-Biserial for ELA

Grade	3		4		5		6		7		8	
Items	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis
1	0.63	0.53	0.68	0.56	0.87	0.70	0.84	0.59	0.54	0.25	0.89	0.66
2	0.67	0.43	0.55	0.44	0.83	0.49	0.63	0.46	0.54	0.40	0.63	0.53
3	0.63	0.48	0.66	0.41	0.48	0.30	0.6	0.42	0.64	0.49	0.38	0.16
4	0.79	0.66	0.63	0.50	0.73	0.56	0.73	0.56	0.76	0.48	0.70	0.31
5	0.48	0.49	0.53	0.49	0.68	0.49	0.34	0.06	0.59	0.40	0.78	0.57
6	0.66	0.47	0.76	0.66	0.49	0.39	0.70	0.53	0.39	0.26	0.43	0.40
7	0.71	0.68	0.60	0.53	0.59	0.55	0.68	0.53	0.65	0.49	0.56	0.29
8	0.42	0.28	0.50	0.46	0.58	0.42	0.63	0.49	0.74	0.58	0.33	0.37
9	0.41	0.23	0.54	0.45	0.54	0.42	0.60	0.57	0.67	0.56	0.59	0.49
10	0.59	0.51	0.34	0.21	0.34	0.15	0.67	0.45	0.58	0.55	0.71	0.54
11	0.61	0.58	0.28	0.25	0.48	0.43	0.61	0.45	0.49	0.37	0.67	0.46
12	0.70	0.56	0.50	0.36	0.54	0.46	0.60	0.48	0.69	0.64	0.54	0.46
13	0.63	0.54	0.65	0.46	0.44	0.38	0.38	0.16	0.59	0.34	0.32	0.22
14	0.66	0.58	0.72	0.54	0.59	0.42	0.41	0.31	0.36	0.38	0.50	0.38
15	0.71	0.55	0.58	0.55	0.57	0.34	0.38	0.16	0.48	0.35	0.45	0.39
16	0.46	0.45	0.62	0.41	0.63	0.44	0.37	0.21	0.50	0.37	0.44	0.34
17	0.64	0.49	0.54	0.51	0.77	0.58	0.21	0.01	0.55	0.53	0.56	0.43
18	0.60	0.45	0.38	0.22	0.29	0.08	0.35	0.33	0.36	0.20	0.65	0.53
19	0.63	0.61	0.48	0.46	0.38	0.27	0.52	0.35	0.45	0.39	0.53	0.42
20	0.56	0.52	0.37	0.34	0.6	0.47	0.51	0.40	0.41	0.39	0.54	0.43
21	0.21	0.26	0.48	0.46	0.52	0.39	0.56	0.28	0.66	0.51	0.39	0.39
22	0.59	0.44	0.68	0.61	0.70	0.56	0.47	0.39	0.55	0.35	0.40	0.26
23	0.48	0.24	0.56	0.49	0.47	0.49	0.41	0.23	0.60	0.47	0.36	0.27
24	0.51	0.43	0.48	0.38	0.44	0.44	0.19	0.10	0.57	0.52	0.46	0.52
25	0.65	0.66	0.35	0.19	0.52	0.34	0.53	0.27	0.65	0.57	0.46	0.34
26	0.45	0.40	0.61	0.54	0.70	0.54	0.41	0.26	0.42	0.43	0.56	0.48
27	0.55	0.43	0.47	0.41	0.76	0.62	0.54	0.44	0.55	0.50	0.63	0.38
28	0.57	0.50	0.63	0.56	0.53	0.45	0.51	0.48	0.46	0.24	0.59	0.51
29	0.65	0.17	0.70	0.60	0.78	0.48	0.65	0.47	0.39	0.28	0.63	0.53
30	0.45	0.42	0.49	0.46	0.54	0.39	0.31	0.18	0.46	0.37	0.69	0.58
31	0.68	0.46	0.67	0.61	0.69	0.54	0.55	0.46	0.52	0.22	0.30	0.19
32	0.48	0.53	0.72	0.53	0.89	0.50	0.48	0.42	0.44	0.25	0.47	0.42
33	0.54	0.5	0.56	0.47	0.66	0.52	0.60	0.43	0.44	0.3	0.5	0.41
34	0.49	0.34	0.5	0.31	0.77	0.55	0.43	0.35	0.43	0.25	0.55	0.44
35	0.41	0.32	0.45	0.23	0.89	0.67	0.64	0.55	0.46	0.27	0.56	0.32
36	0.42	0.46	0.65	0.45	0.55	0.32	0.61	0.42	0.47	0.37	0.51	0.41
37	0.39	0.28	0.33	0.15	0.82	0.53	0.55	0.32	0.62	0.49	0.56	0.34
38	0.73	0.52	0.52	0.43	0.48	0.40	0.47	0.25	0.52	0.30	0.37	0.30
39	0.44	0.31	0.41	0.32	0.74	0.56	0.41	0.18	0.67	0.49	0.42	0.30
40	0.51	0.47	0.46	0.37	0.62	0.47	0.36	0.24	0.51	0.33	0.45	0.37
41	0.54	0.56	0.47	0.4	0.39	0.29	0.80	0.57	0.47	0.41	0.44	0.34
42	0.63	0.56	0.43	0.28	0.58	0.49	0.48	0.31	0.48	0.32	0.76	0.55
43	0.47	0.44	0.47	0.46	0.43	0.31	0.49	0.31	0.64	0.47	0.57	0.44
44	0.52	0.38	0.39	0.3	0.49	0.39	0.49	0.21	0.62	0.51	0.48	0.3
45	0.63	0.59	0.42	0.35	0.79	0.62	0.65	0.34	0.56	0.22	0.39	0.35
46	0.56	0.49	0.57	0.49	0.43	0.24	0.54	0.4	0.55	0.32	0.39	0.33
47	0.75	0.69	0.62	0.52	0.58	0.40	0.52	0.38	0.72	0.59	0.57	0.54
48	0.58	0.56	0.55	0.33	0.59	0.45	0.59	0.44	0.73	0.55	0.71	0.39
49							0.52	0.43	0.53	0.40	0.38	0.25
50							0.49	0.34	0.43	0.31	0.54	0.52
51							0.62	0.45	0.56	0.46	0.32	0.08
52							0.52	0.34	0.42	0.35	0.67	0.50
53							0.40	0.12	0.55	0.48	0.58	0.39
54							0.73	0.58	0.49	0.43	0.72	0.47

Table 8.A.2 Item-by-item *p*-value and Point-Biserial for Mathematics

Grade	3		4		5		6		7	
Items	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis	*p*-value	Pt-Bis
1	0.81	0.64	0.92	0.64	0.95	0.42	0.73	0.38	0.52	0.16
2	0.57	0.52	0.80	0.49	0.78	0.52	0.81	0.23	0.38	0.31
3	0.73	0.53	0.80	0.59	0.75	0.63	0.59	0.27	0.24	0.31
4	0.74	0.64	0.78	0.62	0.70	0.53	0.43	0.21	0.52	0.22
5	0.57	0.55	0.79	0.47	0.64	0.53	0.55	0.44	0.34	0.17
6	0.65	0.30	0.60	0.57	0.93	0.66	0.47	0.22	0.46	0.25
7	0.65	0.60	0.58	0.44	0.59	0.52	0.58	0.43	0.65	0.26
8	0.44	0.48	0.54	0.39	0.54	0.51	0.65	0.37	0.47	0.26
9	0.66	0.57	0.48	0.53	0.41	0.32	0.66	0.27	0.64	0.30
10	0.51	0.51	0.41	0.35	0.59	0.56	0.50	0.44	0.49	0.34
11	0.47	0.45	0.38	0.29	0.42	0.25	0.54	0.26	0.34	0.16
12	0.35	0.29	0.39	0.39	0.59	0.31	0.44	0.35	0.40	0.17
13	0.52	0.40	0.57	0.44	0.45	0.54	0.57	0.37	0.36	0.20
14	0.40	0.42	0.37	0.17	0.63	0.57	0.44	0.43	0.56	0.25
15	0.57	0.57	0.46	0.46	0.47	0.42	0.47	0.30	0.61	0.39
16	0.14	0.02	0.51	0.42	0.55	0.30	0.59	0.38	0.41	0.26
17	0.60	0.44	0.53	0.50	0.59	0.43	0.62	0.45	0.62	0.47
18	0.53	0.50	0.60	0.51	0.53	0.52	0.48	0.30	0.57	0.32
19	0.64	0.56	0.70	0.49	0.73	0.41	0.49	0.40	0.41	0.43
20	0.55	0.55	0.66	0.55	0.63	0.55	0.46	0.25	0.54	0.45
21	0.63	0.56	0.80	0.60	0.82	0.55	0.38	0.24	0.43	0.34
22	0.61	0.44	0.81	0.51	0.70	0.57	0.71	0.54	0.41	0.24
23	0.61	0.67	0.63	0.38	0.69	0.64	0.55	0.37	0.42	0.42
24	0.75	0.56	0.45	0.40	0.77	0.44	0.53	0.45	0.59	0.38
25	0.75	0.58	0.56	0.57	0.48	0.35	0.60	0.57	0.68	0.48
26	0.78	0.62	0.57	0.49	0.63	0.50	0.45	0.36	0.41	0.18
27	0.66	0.46	0.48	0.34	0.29	0.28	0.47	0.41	0.38	0.42
28	0.41	0.27	0.36	0.24	0.50	0.46	0.72	0.60	0.47	0.41
29	0.24	0.26	0.43	0.39	0.60	0.49	0.53	0.18	0.32	0.11
30	0.72	0.62	0.45	0.31	0.68	0.57	0.44	0.31	0.45	0.28
31	0.64	0.50	0.61	0.46	0.52	0.46	0.55	0.46	0.41	0.28
32	0.71	0.57	0.32	0.29	0.73	0.65	0.28	0.07	0.55	0.41
33	0.79	0.45	0.64	0.55	0.45	0.35	0.65	0.64	0.61	0.38
34	0.76	0.56	0.86	0.45	0.66	0.41	0.64	0.50	0.46	0.36
35	0.67	0.57	0.63	0.39	0.46	0.36	0.52	0.34	0.49	0.42
36	0.70	0.49	0.59	0.49	0.55	0.48	0.61	0.47	0.42	0.30
37	0.64	0.49	0.42	0.33	0.36	0.32	0.70	0.58	0.45	0.43
38	0.38	0.36	0.44	0.33	0.44	0.40	0.36	0.20	0.34	0.28
39	0.53	0.34	0.34	0.11	0.46	0.27	0.52	0.43	0.51	0.30
40	0.66	0.57	0.38	0.27	0.50	0.33	0.40	0.16	0.30	0.34
41	0.71	0.36	0.44	0.39	0.66	0.49	0.48	0.22	0.52	0.37
42	0.73	0.56	0.52	0.36	0.67	0.51	0.42	0.28	0.47	0.37
43	0.75	0.57	0.65	0.48	0.38	0.24	0.47	0.41	0.20	–0.08
44	0.71	0.54	0.50	0.30	0.71	0.63	0.55	0.35	0.34	0.08
45	0.71	0.59	0.39	0.17	0.48	0.40	0.48	0.25	0.45	0.31
46	0.42	0.48	0.51	0.37	0.55	0.50	0.43	0.25	0.43	0.42
47	0.59	0.62	0.50	0.38	0.68	0.58	0.47	0.29	0.45	0.31
48	0.73	0.61	0.61	0.52	0.45	0.47	0.44	0.25	0.32	0.18
49							0.63	0.46	0.60	0.43
50							0.75	0.48	0.57	0.43
51							0.49	0.27	0.53	0.36
52							0.52	0.37	0.54	0.37
53							0.36	0.12	0.44	0.31
54							0.53	0.37	0.57	0.31

Table 8.A.3 Item-by-item *p*-value and Point-Biserial for Science

Grade	5		8	
Items	***p*-value**	**Pt-Bis**	***p*-value**	**Pt-Bis**
1	0.85	0.34	0.59	0.18
2	0.68	0.41	0.53	0.28
3	0.45	0.38	0.63	0.50
4	0.82	0.61	0.42	0.27
5	0.48	0.29	0.65	0.40
6	0.38	0.18	0.47	0.26
7	0.75	0.55	0.59	0.40
8	0.64	0.53	0.58	0.45
9	0.45	0.19	0.86	0.66
10	0.80	0.40	0.70	0.47
11	0.89	0.49	0.38	0.36
12	0.42	0.12	0.48	0.24
13	0.45	0.31	0.58	0.48
14	0.59	0.48	0.42	0.33
15	0.63	0.34	0.60	0.37
16	0.82	0.46	0.51	0.27
17	0.71	0.47	0.44	0.42
18	0.68	0.42	0.58	0.40
19	0.51	0.35	0.47	0.47
20	0.44	0.35	0.48	0.41
21	0.75	0.55	0.64	0.51
22	0.85	0.58	0.46	0.39
23	0.65	0.49	0.48	0.38
24	0.84	0.60	0.53	0.34
25	0.64	0.42	0.41	0.36
26	0.43	0.17	0.45	0.39
27	0.58	0.52	0.50	0.42
28	0.64	0.49	0.77	0.53
29	0.46	0.37	0.55	0.46
30	0.59	0.47	0.41	0.32
31	0.56	0.41	0.64	0.55
32	0.78	0.60	0.61	0.53
33	0.65	0.52	0.40	0.34
34	0.55	0.49	0.36	0.07
35	0.50	0.41	0.58	0.41
36	0.54	0.34	0.41	0.22
37	0.51	0.38	0.69	0.47
38	0.54	0.44	0.56	0.36
39	0.63	0.50	0.54	0.42
40	0.70	0.50	0.77	0.53
41	0.35	0.23	0.56	0.26
42	0.41	0.27	0.44	0.32
43	0.53	0.47	0.41	0.35
44	0.77	0.59	0.37	0.17
45	0.57	0.50	0.31	0.08
46	0.77	0.59	0.49	0.33
47	0.44	0.15	0.44	0.24
48	0.73	0.61	0.42	0.25
49			0.35	0.21
50			0.62	0.54
51			0.36	0.15
52			0.47	0.51
53			0.59	0.43
54			0.58	0.36

Table 8.A.4 Distribution of Essay Scores for ELA Grade Four—Overall and by Subgroup (all %)

Score	Total	Females	Males	Eng-best	Initially Fluent	Lrnr	Redesig. Fluent	Not Econ Disadv.	Econ Disadv.
0	1.54	0.31	1.23	0.99	0.03	0.51	0.01	0.46	1.07
1	12.12	3.35	8.77	6.35	0.19	5.54	0.04	2.53	9.57
2	45.26	13.55	31.70	23.97	0.87	20.17	0.24	10.47	34.82
3	33.24	12.60	20.65	19.57	0.75	12.65	0.27	10.39	22.83
4	7.85	3.45	4.40	5.17	0.24	2.38	0.07	3.39	4.46
Total	100	33.26	66.74	56.04	2.08	41.26	0.62	27.24	72.76

Score	Am. Ind.	Asian	Pac. Isl	Filipino	Hisp.	Af. Am.	White	Autism	Deafness
0	0.19	0.06	0.02	0.05	0.76	0.23	0.40	0.18	0.02
1	0.52	0.40	0.08	0.14	7.75	0.02	2.16	1.16	0.13
2	0.40	1.31	0.24	0.56	28.61	1.44	9.18	1.96	0.13
3	0.11	1.03	0.28	0.42	18.97	4.88	8.69	1.08	0.05
4	0.19	0.32	0.07	0.15	4.01	3.38	2.61	0.29	0.01
Total	1.23	3.13	0.69	1.33	60.09	10.50	23.04	4.67	0.33

Score	Emotional Dist.	Hard of Hearing	Mental Retard.	Mult. Disab.	Orthoped. Impair.	Other Health Impair.	Specific Learning Disab.	Speech or Lang. Disab.	Traumatic Brain Injury	Visual Impair.
0	0.11	0.05	0.06	0.00	0.04	0.06	0.73	0.25	0.02	0.02
1	0.29	0.11	0.71	0.03	0.14	0.78	6.64	2.17	0.02	0.05
2	0.90	0.39	0.58	0.14	0.36	3.58	28.79	8.15	0.12	0.13
3	0.49	0.32	0.19	0.04	0.21	2.69	21.73	6.17	0.07	0.08
4	0.10	0.05	0.02	0.00	0.08	0.76	5.17	1.40	0.01	0.01
Total	1.89	0.92	1.56	0.21	0.83	7.88	63.06	18.15	0.23	0.28

Table 8.A.5 Distribution of Essay Scores for ELA Grade Seven—Overall and by Subgroup (all %)

Score	Total	Females	Males	Eng-best	Initially Fluent	Lrnr	Redesig. Fluent	Not Econ Disadv.	Econ Disadv.
0	1.36	0.27	1.10	0.88	0.05	0.42	0.02	0.49	0.87
1	8.17	1.82	6.34	4.34	0.09	3.67	0.06	1.51	6.64
2	43.92	13.89	30.03	22.62	0.75	19.77	0.75	10.00	33.90
3	39.89	15.00	24.89	23.41	0.98	14.59	0.94	12.14	27.78
4	6.65	2.86	3.80	4.72	0.21	1.52	0.20	2.69	3.98
Total	100	33.84	66.16	55.97	2.08	39.98	1.97	26.83	73.17
Score	**Am. Ind.**	**Asian**	**Pac. Isl**	**Filipino**	**Hisp.**	**Af. Am.**	**White**	**Autism**	**Deafness**
0	0.00	0.05	0.00	0.03	0.71	0.24	0.32	0.19	0.03
1	0.11	0.23	0.01	0.10	5.22	1.12	1.34	0.61	0.12
2	0.35	1.38	0.18	0.42	28.11	5.37	8.21	1.41	0.23
3	0.51	1.33	0.11	0.42	22.46	4.91	10.12	0.72	0.06
4	0.06	0.17	0.03	0.10	3.25	0.82	2.19	0.17	0.00
Total	1.04	3.15	0.32	1.07	59.76	12.46	22.20	3.10	0.44

Score	Emotional Dist.	Hard of Hearing	Mental Retard.	Mult. Disab.	Orthoped. Impair.	Other Health Impair.	Specific Learning Disab.	Speech or Lang. Disab.	Traumatic Brain Injury	Visual Impair.
0	0.09	0.04	0.08	0.00	0.01	0.14	0.69	0.10	0.00	0.00
1	0.36	0.11	0.86	0.05	0.08	0.51	4.66	0.70	0.05	0.02
2	1.11	0.32	1.32	0.09	0.33	3.04	31.29	4.48	0.19	0.12
3	0.88	0.27	0.38	0.07	0.17	3.12	30.15	3.85	0.08	0.12
4	0.19	0.07	0.04	0.00	0.02	0.60	4.91	0.64	0.00	0.07
Total	2.64	0.81	2.67	0.21	0.61	7.41	71.70	9.77	0.31	0.33

Table 8.A.6 Mean Scores for ELA Grade Four Essay—Overall and by Subgroup

Overall		Females		Males		English–best		Initially Fluent		Learner		Redes. Fluent		Not Econ. Disadv.		Econ. Disadv.	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
13,089	2.34	4,352	2.47	8,734	2.27	7,329	2.39	272	2.47	5,395	2.26	81	2.57	3,558	2.50	9,502	2.28

Am. Indian		Asian		Pac. Islander		Filipino		Hispanic		Af. American		White		Autism		Deafness	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
160	2.32	406	2.37	89	2.44	172	2.38	7,797	2.29	1,362	2.25	2,990	2.48	586	2.03	42	1.67

Emotional Disturbance		Hard of Hearing		Mental Retardation		Multiple Disabilities		Orthopedic Impair.		Other Health Impair.		Specific Learning Disability		Speech or Lang. Impair.		Traumatic Brain Injury	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
237	2.08	115	2.23	196	1.61	26	2.04	104	2.17	989	2.42	7,916	2.38	2,278	2.35	29	2.17

Visual Imp.	
N	Mean
35	2.06

Table 8.A.7 Mean Scores for ELA Grade Seven Essay—Overall and by Subgroup

Overall		Females		Males		English–best		Initially Fluent		Learner		Redes. Fluent		Not Econ. Disadv.		Econ. Disadv.	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
10,851	2.42	3,672	2.54	7,179	2.36	6,066	2.48	225	2.59	4,333	2.33	214	2.63	2,905	2.56	7,923	2.37

Am. Indian		Asian		Pac. Islander		Filipino		Hispanic		Af. American		White		Autism		Deafness	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
112	2.51	340	2.42	35	2.49	115	2.43	6,446	2.37	1,344	2.40	2,394	2.56	327	2.02	46	1.72

Emotional Disturbance		Hard of Hearing		Mental Retardation		Multiple Disabilities		Orthopedic Impair.		Other Health Impair.		Specific Learning Disability		Speech or Lang. Impair.		Traumatic Brain Injury	
N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
278	2.27	85	2.26	282	1.79	22	2.09	64	2.19	782	2.47	7,563	2.47	1,031	2.43	33	2.09

Visual Imp.	
N	Mean
35	2.71

Table 8.A.8 Effect Sizes for ELA Grades Four and Seven Essay—by Subgroup

	Gender	Econ. Status	English Fluency					
	(M-F)	(No-Yes)	Eng-Ini. Des.	Eng-Lrnr	Eng-Redes.	Ini. Des.-Lrnr	Ini. Des.-Redes.	Lrnr-Redes.
4	–0.22	0.3	–0.09	0.17	–0.21	0.27	–0.12	–0.4
7	–0.22	0.28	–0.17	0.22	–0.18	0.42	–0.01	–0.43

	Primary Ethnicity										
	W-Am. Ind	W-Asian	W-Pac. Isl.	W-Filipino	W-Hisp.	W-Af. Am.	Am. Ind.-Asian	Am. Ind.-Pac. Isl.	Am. Ind.-Filipino	Am. Ind.-Hisp.	Am. Ind.-Af. Am.
4	0.21	0.12	0.03	0.07	0.25	0.27	–0.08	–0.17	–0.14	0.03	0.06
7	0.12	0.19	0.15	0.13	0.27	0.21	0.07	0.03	0.01	0.15	0.09
	Asian-Pac. Isl.	Asian-Filipino	Asian-Hisp.	Asian-Af. Am.	Pac. Isl.-Filipino	Pac. Isl. -Hisp.	Pac. Isl. -Af. Am.	Filipino-Hispanic	Filipino-Af. Am.	Hisp. -Af. Am.	
4	–0.09	–0.06	0.12	0.15	0.04	0.22	0.25	0.18	0.21	0.03	
7	–0.04	–0.05	0.08	0.02	–0.02	0.11	0.06	0.13	0.07	–0.06	

	Disability	
	SLD-Disab Not Rptd	SLI-Disab Not Rptd
4	–0.01	–0.04
7	0.09	0.04

Appendix 8.B—Reliability Analyses

Table 8.B.1 Subscore Reliabilities and Correlations for ELA

Subscore Area	No. of items	Correlation			*Reliab.*	*SEM*
Grade 3		1	2	3		
Vocabulary	14	1.00	0.67	0.65	0.73	1.63
Reading for Understanding	17		1.00	0.63	0.71	1.86
Language	17			1.00	0.66	1.91
Grade 4						
Vocabulary	11	1.00	0.54	0.60	0.66	1.46
Reading for Understanding	16		1.00	0.56	0.61	1.87
Language	21			1.00	0.68	2.13
Grade 5						
Vocabulary	8	1.00	0.49	0.54	0.56	1.17
Reading for Understanding	18		1.00	0.58	0.64	1.98
Language	22			1.00	0.74	2.01

Table 8.B.2 Subscore Reliabilities and Correlations for Mathematics

Subscore Area	No. of items	Correlation			*Reliab.*	*SEM*
Grade 3		1	2	3		
Number Sense	24	1.00	0.67	0.63	0.81	2.15
Algebra and Data Analysis	13		1.00	0.63	0.72	1.52
Measurement and Geometry	11			1.00	0.64	1.41
Grade 4						
Number Sense	23	1.00	0.57	0.45	0.74	2.08
Algebra and Data Analysis	15		1.00	0.43	0.57	1.82
Measurement and Geometry	10			1.00	0.44	1.45
Grade 5						
Number Sense	21	1.00	0.62	0.48	0.75	1.97
Algebra and Data Analysis	17		1.00	0.52	0.74	1.82
Measurement and Geometry	10			1.00	0.49	1.47

Table 8.B.3 Subscore Reliabilities and Correlations for Science

Subscore Area	No. of items	Correlation			*Reliab.*	*SEM*
Grade 5		1	2	3		
Physical Sciences	16	1.00	0.53	0.51	0.55	1.74
Life Sciences	16		1.00	0.60	0.67	1.75
Earth Sciences	16			1.00	0.66	1.81

Table 8.B.4 Reliabilities and SEMs for the CMA by Gender

		Male			Female			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	7,512	0.87	3.15	3,729	0.87	3.11	80	0.90	2.95
	4*	11,444	0.84	3.21	5,662	0.83	3.19	15	0.83	3.19
	5	11,862	0.84	3.09	5,948	0.83	3.05	11	0.84	3.18
	6	10,587	0.78	3.43	5,305	0.78	3.40	8	–	–
	7*	8,648	0.83	3.44	4,402	0.83	3.41	10	–	–
	8	8,218	0.82	3.41	4,212	0.82	3.37	4	–	–
Mathematics	3	7,454	0.89	3.00	3,706	0.88	3.01	67	0.90	2.85
	4*	9,365	0.82	3.14	5,039	0.80	3.14	11	0.76	3.25
	5	10,193	0.85	3.10	5,452	0.85	3.05	11	0.77	3.26
	6	9,980	0.78	3.47	5,289	0.76	3.47	9	–	–
	7*	8,548	0.72	3.51	4,587	0.68	3.49	10	–	–
Science	5	11,052	0.83	3.06	5,617	0.80	3.11	12	0.83	3.20
	8	7,447	0.81	3.44	3,933	0.73	3.48	4	–	–

Table 8.B.5 Reliabilities and SEMs for the CMA by Economic Status

		Econ Disadv.			Not Econ Disadv.			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	8,313	0.86	3.17	2,522	0.88	3.05	486	0.88	3.06
	4*	12,498	0.82	3.23	4,560	0.85	3.13	63	0.86	3.16
	5	13,191	0.83	3.10	4,560	0.85	3.00	70	0.81	3.21
	6	11,802	0.77	3.44	4,035	0.80	3.37	63	0.79	3.46
	7*	9,400	0.82	3.45	3,592	0.85	3.38	68	0.84	3.46
	8	8,748	0.81	3.41	3,635	0.83	3.35	51	0.75	3.48
Mathematics	3	8,267	0.88	3.02	2,500	0.89	2.95	460	0.89	2.97
	4*	10,684	0.80	3.15	3,681	0.83	3.10	50	0.77	3.20
	5	11,560	0.85	3.09	4,030	0.86	3.05	66	0.86	3.15
	6	11,191	0.77	3.48	4,022	0.78	3.45	65	0.71	3.53
	7*	9,341	0.69	3.51	3,739	0.73	3.50	65	0.69	3.50
Science	5	12,314	0.81	3.11	4,300	0.83	2.99	67	0.81	3.19
	8	7,986	0.77	3.47	3,352	0.81	3.42	46	0.68	3.53

Table 8.B.6 Reliabilities and SEMs for the CMA by Language Fluency

		English Only			I-FEP[1]			English Learner			R-FEP[2]			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	6,153	0.88	3.10	145	0.87	3.11	4,682	0.85	3.19	37	0.91	2.96	304	0.86	3.10
	4*	9,560	0.84	3.17	340	0.84	3.16	7,078	0.80	3.24	100	0.84	3.09	43	0.83	3.21
	5	9,776	0.84	3.04	401	0.83	3.01	7,445	0.82	3.12	156	0.85	2.97	43	0.84	3.17
	6	8,621	0.79	3.40	291	0.81	3.41	6,741	0.76	3.45	218	0.82	3.35	29	0.74	3.49
	7*	7,380	0.84	3.41	268	0.85	3.40	5,111	0.80	3.47	260	0.85	3.33	41	0.82	3.48
	8	7,139	0.84	3.38	339	0.81	3.38	4,692	0.79	3.43	234	0.84	3.32	30	0.85	3.36
Mathematics	3	6,097	0.89	3.00	144	0.88	2.97	4,662	0.88	3.02	37	0.92	2.80	287	0.88	2.98
	4*	8,044	0.82	3.14	267	0.85	3.09	5,988	0.80	3.15	78	0.80	3.08	38	0.69	3.24
	5	8,757	0.86	3.09	336	0.85	2.98	6,389	0.85	3.08	132	0.85	3.01	42	0.85	3.18
	6	8,548	0.77	3.47	288	0.80	3.44	6,196	0.77	3.48	215	0.83	3.38	31	0.72	3.53
	7*	7,615	0.71	3.51	271	0.68	3.52	4,946	0.68	3.51	273	0.75	3.47	40	0.68	3.50
Science	5	9,243	0.83	3.04	365	0.81	2.99	6,881	0.80	3.13	148	0.78	3.03	44	0.80	3.19
	8	6,599	0.80	3.45	283	0.79	3.44	4,252	0.75	3.47	222	0.79	3.39	28	0.72	3.52

[1] Initially-Fluent English Proficient

[2] Reclassified-Fluent English Proficient

Table 8.B.7 Reliabilities and SEMs for the CMA by Primary Ethnicity

		African American			American Indian			Asian			Filipino			Hispanic			Pacific Islander			White			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	1,282	0.87	3.12	135	0.90	3.06	381	0.88	3.10	156	0.86	3.09	6,718	0.86	3.17	70	0.86	3.13	2,294	0.88	3.06	285	0.89	3.03
	4*	1,888	0.83	3.22	203	0.84	3.18	545	0.83	3.18	230	0.83	3.15	10,173	0.82	3.23	114	0.85	3.19	3,800	0.85	3.14	168	0.85	3.14
	5	2,069	0.83	3.10	227	0.83	3.05	527	0.82	3.08	215	0.86	3.01	10,763	0.83	3.10	104	0.84	3.04	3,769	0.85	2.99	147	0.85	3.09
	6	1,818	0.78	3.44	195	0.81	3.42	502	0.75	3.41	166	0.75	3.38	9,573	0.77	3.44	95	0.79	3.41	3,417	0.81	3.37	134	0.80	3.39
	7*	1,589	0.83	3.44	157	0.82	3.43	395	0.83	3.41	135	0.83	3.42	7,631	0.82	3.45	46	0.79	3.50	3,003	0.86	3.36	104	0.87	3.38
	8	1,466	0.81	3.42	175	0.84	3.38	360	0.79	3.38	125	0.79	3.42	7,185	0.81	3.41	56	0.78	3.47	2,970	0.85	3.34	97	0.82	3.42
Mathematics	3	1,272	0.88	3.06	132	0.90	2.96	378	0.90	2.92	157	0.90	2.95	6,670	0.88	3.02	72	0.89	3.05	2,278	0.89	2.96	268	0.90	2.94
	4*	1,721	0.80	3.19	183	0.79	3.14	432	0.85	3.08	192	0.83	3.08	8,593	0.80	3.15	99	0.82	3.16	3,065	0.83	3.10	130	0.80	3.15
	5	1,956	0.85	3.14	198	0.83	3.11	426	0.87	3.03	178	0.87	3.04	9,329	0.85	3.08	89	0.84	3.10	3,348	0.86	3.05	132	0.86	3.14
	6	1,824	0.74	3.50	192	0.75	3.48	442	0.80	3.42	162	0.79	3.45	8,964	0.77	3.48	84	0.77	3.47	3,476	0.79	3.45	134	0.80	3.45
	7*	1,626	0.66	3.51	154	0.71	3.51	383	0.72	3.49	148	0.75	3.49	7,537	0.69	3.51	50	0.72	3.51	3,141	0.74	3.50	106	0.75	3.49
Science	5	1,979	0.82	3.14	213	0.82	3.04	484	0.80	3.13	202	0.81	3.07	9,960	0.81	3.11	103	0.80	3.07	3,598	0.83	2.95	142	0.85	3.06
	8	1,366	0.75	3.50	161	0.79	3.44	337	0.78	3.45	109	0.79	3.45	6,511	0.76	3.47	52	0.78	3.49	2,763	0.82	3.40	85	0.80	3.43

Table 8.B.8 Reliabilities and SEMs for the CMA by Primary Ethnicity-for-Not-Economically-Disadvantaged

		African American			American Indian			Asian			Filipino			Hispanic			Pacific Islander			White			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	225	0.88	3.07	24	0.90	2.97	147	0.89	3.00	83	0.85	3.08	860	0.87	3.10	19	0.87	3.02	1,109	0.88	3.02	55	0.85	3.01
	4*	441	0.85	3.18	58	0.83	3.19	224	0.82	3.17	128	0.81	3.14	1,466	0.84	3.15	46	0.86	3.13	2,123	0.85	3.10	74	0.87	3.07
	5	488	0.84	3.06	55	0.87	2.89	221	0.82	3.05	132	0.84	2.99	1,527	0.84	3.03	33	0.86	3.00	2,051	0.85	2.95	53	0.86	2.95
	6	438	0.79	3.42	51	0.84	3.36	188	0.76	3.36	99	0.73	3.38	1,320	0.79	3.39	34	0.82	3.41	1,851	0.81	3.34	54	0.76	3.35
	7*	396	0.83	3.43	52	0.81	3.44	138	0.84	3.37	76	0.83	3.39	1,271	0.84	3.41	13	0.73	3.51	1,607	0.86	3.33	39	0.88	3.32
	8	398	0.83	3.39	56	0.75	3.39	125	0.82	3.33	69	0.76	3.42	1,285	0.82	3.37	24	0.80	3.44	1,638	0.85	3.32	40	0.83	3.43
Mathematics	3	220	0.87	3.03	24	0.91	2.78	146	0.91	2.86	84	0.89	2.95	854	0.88	2.97	19	0.84	3.09	1,102	0.89	2.93	51	0.87	2.95
	4*	389	0.81	3.17	52	0.74	3.15	177	0.87	3.02	109	0.83	3.06	1,183	0.82	3.11	38	0.84	3.15	1,675	0.84	3.08	58	0.85	3.11
	5	466	0.86	3.11	45	0.82	3.13	173	0.87	3.02	116	0.87	2.98	1,348	0.85	3.07	30	0.84	3.09	1,802	0.86	3.03	50	0.89	3.07
	6	456	0.74	3.50	54	0.71	3.42	167	0.80	3.40	96	0.81	3.41	1,276	0.78	3.46	29	0.78	3.47	1,887	0.78	3.43	57	0.74	3.44
	7*	413	0.70	3.52	49	0.70	3.51	132	0.78	3.46	85	0.76	3.50	1,280	0.69	3.52	16	0.81	3.51	1,719	0.75	3.49	45	0.75	3.49
Science	5	466	0.83	3.11	51	0.84	2.92	192	0.82	3.11	126	0.80	3.07	1,434	0.82	3.04	35	0.83	3.05	1,944	0.83	2.90	52	0.85	2.96
	8	371	0.75	3.49	57	0.83	3.40	108	0.76	3.45	62	0.80	3.45	1,159	0.78	3.44	22	0.78	3.47	1,538	0.83	3.38	35	0.83	3.35

Table 8.B.9 Reliabilities and SEMs for the CMA by Primary Ethnicity-for-Economically-Disadvantaged

		African American			American Indian			Asian			Filipino			Hispanic			Pacific Islander			White			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	999	0.87	3.13	104	0.89	3.10	219	0.86	3.17	62	0.87	3.13	5,697	0.85	3.19	50	0.84	3.18	1,079	0.88	3.10	103	0.90	3.06
	4*	1,440	0.82	3.24	145	0.84	3.17	321	0.84	3.19	102	0.84	3.17	8,676	0.81	3.24	68	0.84	3.22	1,665	0.84	3.19	81	0.80	3.20
	5	1,571	0.83	3.11	168	0.80	3.09	302	0.82	3.10	83	0.87	3.03	9,215	0.82	3.11	71	0.84	3.06	1,701	0.85	3.04	80	0.83	3.13
	6	1,370	0.77	3.44	143	0.79	3.44	313	0.73	3.44	65	0.78	3.39	8,229	0.76	3.44	60	0.77	3.41	1,552	0.80	3.40	70	0.83	3.41
	7*	1,184	0.82	3.44	104	0.83	3.43	257	0.83	3.43	59	0.83	3.47	6,333	0.81	3.46	33	0.79	3.50	1,381	0.85	3.40	49	0.85	3.42
	8	1,065	0.80	3.43	119	0.87	3.36	234	0.77	3.41	56	0.81	3.42	5,877	0.80	3.42	32	0.74	3.50	1,321	0.84	3.36	44	0.82	3.37
Mathematics	3	996	0.88	3.07	103	0.89	3.02	217	0.90	2.98	62	0.91	2.96	5,662	0.88	3.02	52	0.90	3.03	1,073	0.89	2.98	102	0.91	2.95
	4*	1,327	0.80	3.20	131	0.80	3.14	255	0.84	3.11	83	0.82	3.11	7,384	0.80	3.15	61	0.80	3.16	1,381	0.81	3.13	62	0.73	3.18
	5	1,480	0.84	3.15	149	0.84	3.10	249	0.87	3.03	62	0.85	3.13	7,965	0.85	3.08	59	0.84	3.10	1,528	0.86	3.08	68	0.84	3.14
	6	1,357	0.74	3.51	137	0.75	3.50	274	0.80	3.43	64	0.77	3.49	7,665	0.77	3.48	54	0.77	3.47	1,573	0.78	3.46	67	0.83	3.44
	7*	1,204	0.65	3.51	104	0.72	3.51	251	0.68	3.51	63	0.74	3.48	6,231	0.69	3.51	34	0.54	3.52	1,408	0.71	3.51	46	0.75	3.48
Science	5	1,503	0.81	3.15	158	0.81	3.07	288	0.79	3.14	76	0.83	3.08	8,509	0.81	3.12	68	0.79	3.08	1,636	0.83	3.01	76	0.85	3.07
	8	992	0.74	3.50	104	0.75	3.46	228	0.78	3.46	47	0.78	3.46	5,332	0.76	3.47	30	0.79	3.48	1,214	0.80	3.43	39	0.75	3.45

Table 8.B.10 Reliabilities and SEMs for the CMA by Gender by Economic Status

		Economically Disadvantaged									Not Economically Disadvantaged								
		Male			Female			Unknown			Male			Female			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts	3	5,520	0.86	3.18	2,781	0.86	3.14	12	0.91	3.00	1,706	0.88	3.07	812	0.88	3.01	4	–	–
****MC Only***	4*	8,275	0.82	3.23	4,218	0.81	3.22	5	–	–	3,129	0.85	3.14	1,430	0.85	3.11	1	–	–
	5	8,719	0.83	3.12	4,470	0.82	3.08	2	–	–	3,103	0.85	3.01	1,457	0.84	2.95	-	–	–
	6	7,813	0.77	3.45	3,988	0.77	3.41	1	–	–	2,735	0.80	3.38	1,298	0.80	3.34	2	–	–
	7*	6,159	0.82	3.46	3,237	0.82	3.43	4	–	–	2,443	0.85	3.39	1,149	0.85	3.35	-	–	–
	8	5,717	0.81	3.42	3,030	0.81	3.39	1	–	–	2,470	0.83	3.37	1,165	0.84	3.30	-	–	–
Mathematics	3	5,488	0.89	3.02	2,769	0.88	3.03	10	–	–	1,689	0.89	2.94	807	0.87	2.96	4	–	–
	4*	6,891	0.81	3.15	3,789	0.79	3.15	4	–	–	2,443	0.84	3.10	1,237	0.82	3.10	1	–	–
	5	7,516	0.85	3.10	4,042	0.85	3.06	2	–	–	2,638	0.86	3.07	1,392	0.86	3.02	-	–	–
	6	7,280	0.77	3.48	3,909	0.76	3.48	2	–	–	2,661	0.79	3.44	1,359	0.77	3.46	2	–	–
	7*	6,054	0.70	3.51	3,283	0.66	3.50	4	–	–	2,453	0.74	3.50	1,286	0.71	3.49	-	–	–
Science	5	8,122	0.82	3.09	4,189	0.79	3.14	3	–	–	2,892	0.84	2.98	1,408	0.81	3.02	-	–	–
	8	5,166	0.79	3.45	2,818	0.71	3.49	2	–	–	2,252	0.82	3.40	1,100	0.77	3.45	-	–	–

Table 8.B.11 Reliabilities and SEMs for the CMA by Primary Disability

Subject	Grade	Autism			Deaf-Blindness			Deafness			Emotional Dist.			Hard of Hearing			Mental Retard.			Mult. Disab.		
		N	Rel	SEM	N	Rel	SEM	N	Rel	SEM	N	Rel	SEM	N	Rel	SEM	N	Rel	SEM	N	Rel	SEM
English–Language Arts *MC Only	3	652	0.87	3.09	–	–	–	77	0.69	3.29	218	0.90	3.03	103	0.83	3.20	305	0.80	3.24	23	0.87	3.21
	4*	792	0.84	3.18	1	–	–	77	0.76	3.27	344	0.86	3.15	151	0.82	3.19	284	0.71	3.30	34	0.76	3.27
	5	699	0.83	3.11	–	–	–	81	0.80	3.20	362	0.86	3.05	155	0.82	3.05	383	0.73	3.27	55	0.79	3.19
	6	499	0.78	3.43	–	–	–	61	0.65	3.48	389	0.83	3.37	136	0.76	3.40	387	0.64	3.51	37	0.73	3.48
	7*	404	0.85	3.41	–	–	–	77	0.76	3.47	392	0.89	3.33	105	0.83	3.44	347	0.65	3.53	29	0.87	3.41
	8	328	0.84	3.39	–	–	–	54	0.58	3.44	352	0.88	3.32	112	0.79	3.44	340	0.67	3.48	45	0.79	3.43
Mathematics	3	650	0.88	3.05	–	–	–	77	0.88	3.08	215	0.90	3.02	101	0.86	2.99	300	0.81	3.22	23	0.86	3.15
	4	708	0.85	3.14	1	–	–	73	0.82	3.12	302	0.82	3.18	134	0.82	3.09	283	0.64	3.29	34	0.70	3.28
	5	664	0.87	3.12	–	–	–	72	0.90	3.03	364	0.85	3.14	134	0.86	2.98	371	0.69	3.27	53	0.79	3.24
	6	540	0.79	3.46	–	–	–	59	0.75	3.49	404	0.80	3.45	128	0.74	3.44	391	0.55	3.54	39	0.78	3.51
	7	415	0.74	3.50	–	–	–	79	0.65	3.51	410	0.74	3.50	103	0.75	3.48	357	0.51	3.48	27	0.77	3.50
Science	5	679	0.85	3.08	–	–	–	81	0.84	3.16	345	0.87	3.02	149	0.79	3.08	376	0.75	3.27	55	0.70	3.21
	8	313	0.82	3.43	–	–	–	81	0.65	3.45	332	0.81	3.45	97	0.79	3.43	319	0.58	3.51	44	0.79	3.47

Table 8.B.12 Reliabilities and SEMs for the CMA by Primary Disability (continued)

		Orthoped. Impair.			Other Health Impair.			Specific Lrn Disab.			Speech or Lang Impair.			Traumatic Brain Injury			Visual Impair.			Unknown		
Subject	**Grade**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**	**N**	**Rel**	**SEM**
English–Language Arts *MC Only	3	103	0.88	3.11	780	0.87	3.10	6,039	0.87	3.14	2,302	0.86	3.14	47	0.85	3.20	43	0.85	3.12	629	0.87	3.12
	4*	144	0.86	3.19	1,252	0.85	3.15	9,899	0.83	3.21	2,873	0.82	3.20	42	0.82	3.21	49	0.85	3.22	1,179	0.81	3.21
	5	113	0.85	3.10	1,313	0.84	3.02	11,138	0.84	3.07	2,417	0.82	3.06	49	0.88	3.04	35	0.86	3.12	1,021	0.82	3.08
	6	127	0.79	3.43	1,113	0.80	3.38	10,577	0.78	3.42	1,661	0.76	3.41	52	0.73	3.47	44	0.70	3.44	817	0.79	3.41
	7*	91	0.87	3.39	929	0.84	3.37	8,872	0.83	3.44	1,179	0.81	3.43	42	0.82	3.44	38	0.83	3.43	555	0.85	3.42
	8	93	0.80	3.41	862	0.84	3.36	8,738	0.82	3.39	976	0.79	3.40	46	0.83	3.43	22	0.81	3.40	466	0.82	3.41
Mathematics	3	96	0.90	3.04	781	0.88	3.03	5,990	0.88	2.99	2,295	0.89	2.98	46	0.90	2.99	43	0.91	2.97	610	0.88	2.98
	4	135	0.83	3.20	1,063	0.82	3.14	8,221	0.81	3.14	2,360	0.82	3.12	36	0.79	3.17	45	0.84	3.17	1,020	0.79	3.12
	5	119	0.87	3.15	1,226	0.86	3.08	9,584	0.85	3.07	2,066	0.86	3.06	46	0.88	3.10	37	0.86	3.10	920	0.84	3.07
	6	136	0.81	3.46	1,191	0.77	3.47	9,905	0.77	3.47	1,569	0.76	3.47	52	0.79	3.48	42	0.84	3.47	822	0.77	3.48
	7	103	0.77	3.46	1,034	0.72	3.51	8,785	0.70	3.51	1,182	0.68	3.50	44	0.67	3.53	38	0.72	3.48	568	0.68	3.51
Science	5	115	0.81	3.17	1,239	0.82	3.04	10,346	0.82	3.07	2,263	0.80	3.10	47	0.87	3.13	33	0.84	3.12	953	0.81	3.09
	8	93	0.79	3.46	779	0.80	3.45	7,927	0.79	3.45	892	0.76	3.46	47	0.78	3.50	20	0.84	3.39	440	0.77	3.46

Table 8.B.13 Overall Subgroup Reliabilities

Content Area	Grade	*Gender*		*Econ. Dis.*		*Language Fluency*			
		Male	Female	No	Yes	Eng. Only	Ini. Fluent	Learner	Red. Fluent
English–Language Arts *MC Only	3	0.87	0.87	0.88	0.86	0.88	0.87	0.85	0.91
	4*	0.84	0.83	0.85	0.82	0.84	0.84	0.80	0.84
	5	0.84	0.83	0.85	0.83	0.84	0.83	0.82	0.85
	6	0.78	0.78	0.80	0.77	0.79	0.81	0.76	0.82
	7*	0.83	0.83	0.85	0.82	0.84	0.85	0.80	0.85
	8	0.82	0.82	0.83	0.81	0.84	0.81	0.79	0.84
Mathematics	3	0.89	0.88	0.89	0.88	0.89	0.88	0.88	0.92
	4	0.82	0.80	0.83	0.80	0.82	0.85	0.80	0.80
	5	0.85	0.85	0.86	0.85	0.86	0.85	0.85	0.85
	6	0.78	0.76	0.78	0.77	0.77	0.80	0.77	0.83
	7	0.72	0.68	0.73	0.69	0.71	0.68	0.68	0.75
Science	5	0.83	0.80	0.83	0.81	0.83	0.81	0.80	0.78
	8	0.81	0.73	0.81	0.77	0.80	0.79	0.75	0.79

Table 8.B.14 Overall Subgroup Reliabilities—Primary Ethnicity

Content Area	Grade	*Primary Ethnicity*						
		African American	American Indian	Asian	Filipino	Hispanic	Pacific Islander	White
English–Language Arts *MC Only	3	0.87	0.90	0.88	0.86	0.86	0.86	0.88
	4*	0.83	0.84	0.83	0.83	0.82	0.85	0.85
	5	0.83	0.83	0.82	0.86	0.83	0.84	0.85
	6	0.78	0.81	0.75	0.75	0.77	0.79	0.81
	7*	0.83	0.82	0.83	0.83	0.82	0.79	0.86
	8	0.81	0.84	0.79	0.79	0.81	0.78	0.85
Mathematics	3	0.88	0.90	0.90	0.90	0.88	0.89	0.89
	4	0.80	0.79	0.85	0.83	0.80	0.82	0.83
	5	0.85	0.83	0.87	0.87	0.85	0.84	0.86
	6	0.74	0.75	0.80	0.79	0.77	0.77	0.79
	7	0.66	0.71	0.72	0.75	0.69	0.72	0.74
Science	5	0.82	0.82	0.80	0.81	0.81	0.80	0.83
	8	0.75	0.79	0.78	0.79	0.76	0.78	0.82

Table 8.B.15 Overall Subgroup Reliabilities by Primary Ethnicity—Not Economically Disadvantaged

Content Area	Grade	*Ethnicity*						
		African American	American Indian	Asian	Filipino	Hispanic	Pacific Islander	White
English–Language Arts *MC Only	3	0.88	0.90	0.89	0.85	0.87	0.87	0.88
	4*	0.85	0.83	0.82	0.81	0.84	0.86	0.85
	5	0.84	0.87	0.82	0.84	0.84	0.86	0.85
	6	0.79	0.84	0.76	0.73	0.79	0.82	0.81
	7*	0.83	0.81	0.84	0.83	0.84	0.73	0.86
	8	0.83	0.75	0.82	0.76	0.82	0.80	0.85
Mathematics	3	0.87	0.91	0.91	0.89	0.88	0.84	0.89
	4	0.81	0.74	0.87	0.83	0.82	0.84	0.84
	5	0.86	0.82	0.87	0.87	0.85	0.84	0.86
	6	0.74	0.71	0.80	0.81	0.78	0.78	0.78
	7	0.70	0.70	0.78	0.76	0.69	0.81	0.75
Science	5	0.83	0.84	0.82	0.80	0.82	0.83	0.83
	8	0.75	0.83	0.76	0.80	0.78	0.78	0.83

Table 8.B.16 Overall Subgroup Reliabilities by Primary Ethnicity—Economically Disadvantaged

		Ethnicity						
Content Area		**African American**	**American Indian**	**Asian**	**Filipino**	**Hispanic**	**Pacific Islander**	**White**
English–Language Arts *MC Only	3	0.87	0.89	0.86	0.87	0.85	0.84	0.88
	4*	0.82	0.84	0.84	0.84	0.81	0.84	0.84
	5	0.83	0.80	0.82	0.87	0.82	0.84	0.85
	6	0.77	0.79	0.73	0.78	0.76	0.77	0.80
	7*	0.82	0.83	0.83	0.83	0.81	0.79	0.85
	8	0.80	0.87	0.77	0.81	0.80	0.74	0.84
Mathematics	3	0.88	0.89	0.90	0.91	0.88	0.90	0.89
	4	0.80	0.80	0.84	0.82	0.80	0.80	0.81
	5	0.84	0.84	0.87	0.85	0.85	0.84	0.86
	6	0.74	0.75	0.80	0.77	0.77	0.77	0.78
	7	0.65	0.72	0.68	0.74	0.69	0.54	0.71
Science	5	0.81	0.81	0.79	0.83	0.81	0.79	0.83
	8	0.74	0.75	0.78	0.78	0.76	0.79	0.80

Table 8.B.17 Overall Subgroup Reliabilities by Gender/Economic Status

		Economically Disadvantaged		**Not Economically Disadvantaged**	
Content Area	**Grade**	**Male**	**Female**	**Male**	**Female**
English–Language Arts *MC Only	3	0.86	0.86	0.88	0.88
	4*	0.82	0.81	0.85	0.85
	5	0.83	0.82	0.85	0.84
	6	0.77	0.77	0.80	0.80
	7*	0.82	0.82	0.85	0.85
	8	0.81	0.81	0.83	0.84
Mathematics	3	0.89	0.88	0.89	0.87
	4*	0.81	0.79	0.84	0.82
	5	0.85	0.85	0.86	0.86
	6	0.77	0.76	0.79	0.77
	7*	0.70	0.66	0.74	0.71
Science	5	0.82	0.79	0.84	0.81
	8	0.79	0.71	0.82	0.77

Table 8.B.18 Overall Subgroup Reliabilities by Primary Disability

Content Area	**GRADE**	**Autism**	**Deaf-Blindness**	**Deafness**	**Emotional Dist.**	**Hard of Hearing**	**Mental Retard.**	**Mult. Disab.**
English–Language Arts *MC Only	3	0.87	–	0.69	0.90	0.83	0.80	0.87
	4*	0.84	–	0.76	0.86	0.82	0.71	0.76
	5	0.83	–	0.80	0.86	0.82	0.73	0.79
	6	0.78	–	0.65	0.83	0.76	0.64	0.73
	7*	0.85	–	0.76	0.89	0.83	0.65	0.87
	8	0.84	–	0.58	0.88	0.79	0.67	0.79
Mathematics	3	0.88	–	0.88	0.90	0.86	0.81	0.86
	4	0.85	–	0.82	0.82	0.82	0.64	0.70
	5	0.87	–	0.90	0.85	0.86	0.69	0.79
	6	0.79	–	0.75	0.80	0.74	0.55	0.78
	7	0.74	–	0.65	0.74	0.75	0.51	0.77
Science	5	0.85	–	0.84	0.87	0.79	0.75	0.70
	8	0.82	–	0.65	0.81	0.79	0.58	0.79

Table 8.B.19 Overall Subgroup Reliabilities by Primary Disability (continued)

Content Area	Grade	Orthoped. Impair.	Other Health Impair.	Specific Lrn Disab.	Speech or Lang Impair.	Traumatic Brain Injury	Visual Impair.
English–Language Arts *MC Only	3	0.88	0.87	0.87	0.86	0.85	0.85
	4*	0.86	0.85	0.83	0.82	0.82	0.85
	5	0.85	0.84	0.84	0.82	0.88	0.86
	6	0.79	0.80	0.78	0.76	0.73	0.70
	7*	0.87	0.84	0.83	0.81	0.82	0.83
	8	0.80	0.84	0.82	0.79	0.83	0.81
Mathematics	3	0.90	0.88	0.88	0.89	0.90	0.91
	4*	0.83	0.82	0.81	0.82	0.79	0.84
	5	0.87	0.86	0.85	0.86	0.88	0.86
	6	0.81	0.77	0.77	0.76	0.79	0.84
	7*	0.77	0.72	0.70	0.68	0.67	0.72
Science	5	0.81	0.82	0.82	0.80	0.87	0.84
	8	0.79	0.80	0.79	0.76	0.78	0.84

Table 8.B.20 Subscore Reliabilities and SEM by Gender

		Male		Female	
Subscore Area	**# of Items**	**Reliab.**	**SEM**	**Reliab.**	**SEM**
ELA Grade 3					
1. Vocabulary	14	0.73	1.63	0.72	1.62
2. Reading for Understanding	17	0.71	1.87	0.72	1.84
3. Language	17	0.66	1.92	0.67	1.90
ELA Grade 4					
1. Vocabulary	11	0.67	1.46	0.65	1.46
2. Reading for Understanding	16	0.61	1.87	0.59	1.86
3. Language	21	0.68	2.14	0.68	2.12
ELA Grade 5					
1. Vocabulary	8	0.57	1.17	0.54	1.17
2. Reading for Understanding	18	0.64	1.98	0.63	1.98
3. Language	22	0.74	2.03	0.73	1.98
Mathematics Grade 3					
1. Number Sense	24	0.81	2.14	0.79	2.15
2. Algebra & Data Analysis	13	0.72	1.52	0.71	1.52
3. Measurement & Geometry	11	0.64	1.41	0.63	1.42
Mathematics Grade 4					
1. Number Sense	23	0.75	2.08	0.72	2.08
2. Algebra & Data Analysis	15	0.57	1.82	0.56	1.82
3. Measurement & Geometry	10	0.44	1.45	0.43	1.45
Mathematics Grade 5					
1. Number Sense	21	0.75	1.98	0.75	1.94
2. Algebra & Data Analysis	17	0.74	1.83	0.74	1.81
3. Measurement & Geometry	10	0.50	1.47	0.47	1.46
Science Grade 5					
1. Physical Sciences	16	0.56	1.73	0.51	1.77
2. Life Sciences	16	0.68	1.74	0.64	1.77
3. Earth Sciences	16	0.68	1.80	0.62	1.83

Table 8.B.21 Subscore Reliabilities and SEM by Gender—Not Economically Disadvantaged

Subscore Area	# of Items	Male		Female	
		Reliab.	SEM	Reliab.	SEM
ELA Grade 3					
1. Vocabulary	14	0.75	1.55	0.72	1.54
2. Reading for Understanding	17	0.72	1.83	0.73	1.79
3. Language	17	0.69	1.89	0.70	1.85
ELA Grade 4					
1. Vocabulary	11	0.66	1.41	0.66	1.39
2. Reading for Understanding	16	0.65	1.84	0.66	1.82
3. Language	21	0.71	2.10	0.70	2.07
ELA Grade 5					
1. Vocabulary	8	0.54	1.12	0.53	1.10
2. Reading for Understanding	18	0.68	1.95	0.66	1.94
3. Language	22	0.75	1.98	0.74	1.90
Mathematics Grade 3					
1. Number Sense	24	0.81	2.10	0.79	2.12
2. Algebra & Data Analysis	13	0.72	1.50	0.69	1.50
3. Measurement & Geometry	11	0.65	1.37	0.64	1.39
Mathematics Grade 4					
1. Number Sense	23	0.78	2.04	0.75	2.04
2. Algebra & Data Analysis	15	0.60	1.81	0.60	1.80
3. Measurement & Geometry	10	0.44	1.44	0.48	1.43
Mathematics Grade 5					
1. Number Sense	21	0.76	1.96	0.76	1.92
2. Algebra & Data Analysis	17	0.75	1.80	0.74	1.78
3. Measurement & Geometry	10	0.53	1.46	0.48	1.46
Science Grade 5					
1. Physical Sciences	16	0.56	1.68	0.51	1.71
2. Life Sciences	16	0.69	1.69	0.67	1.71
3. Earth Sciences	16	0.69	1.76	0.64	1.78

Table 8.B.22 Subscore Reliabilities and SEM by Gender—Economically Disadvantaged

Subscore Area	# of Items	Male		Female	
		Reliab.	SEM	Reliab.	SEM
ELA Grade 3					
1. Vocabulary	14	0.72	1.65	0.71	1.65
2. Reading for Understanding	17	0.70	1.88	0.71	1.85
3. Language	17	0.63	1.93	0.65	1.91
ELA Grade 4					
1. Vocabulary	11	0.66	1.48	0.64	1.48
2. Reading for Understanding	16	0.58	1.88	0.55	1.87
3. Language	21	0.66	2.15	0.65	2.13
ELA Grade 5					
1. Vocabulary	8	0.56	1.19	0.53	1.19
2. Reading for Understanding	18	0.62	1.99	0.61	1.99
3. Language	22	0.74	2.05	0.72	2.00
Mathematics Grade 3					
1. Number Sense	24	0.81	2.16	0.79	2.16
2. Algebra & Data Analysis	13	0.72	1.52	0.71	1.53
3. Measurement & Geometry	11	0.63	1.42	0.63	1.43
Mathematics Grade 4					
1. Number Sense	23	0.74	2.09	0.71	2.10
2. Algebra & Data Analysis	15	0.56	1.83	0.55	1.83
3. Measurement & Geometry	10	0.44	1.46	0.41	1.45
Mathematics Grade 5					
1. Number Sense	21	0.74	1.99	0.74	1.95
2. Algebra & Data Analysis	17	0.73	1.83	0.73	1.81
3. Measurement & Geometry	10	0.49	1.48	0.47	1.47
Science Grade 5					
1. Physical Sciences	16	0.56	1.74	0.50	1.78
2. Life Sciences	16	0.66	1.76	0.63	1.79
3. Earth Sciences	16	0.67	1.82	0.61	1.84

Table 8.B.23 Subscore Reliabilities and SEM by Language Fluency

Subscore Area	# of Items	English Learner		English Only		I-FEP		R-FEP	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3									
1. Vocabulary	14	0.70	1.66	0.74	1.60	0.74	1.60	0.79	1.52
2. Reading for Understanding	17	0.67	1.90	0.73	1.83	0.76	1.82	0.78	1.77
3. Language	17	0.62	1.93	0.69	1.90	0.66	1.91	0.78	1.81
ELA Grade 4									
1. Vocabulary	11	0.65	1.49	0.66	1.44	0.67	1.41	0.47	1.44
2. Reading for Understanding	16	0.53	1.88	0.64	1.85	0.65	1.84	0.70	1.79
3. Language	21	0.64	2.15	0.70	2.12	0.69	2.11	0.73	2.05
ELA Grade 5									
1. Vocabulary	8	0.53	1.21	0.56	1.14	0.55	1.14	0.53	1.14
2. Reading for Understanding	18	0.59	2.00	0.66	1.97	0.62	1.97	0.68	1.93
3. Language	22	0.72	2.05	0.75	1.99	0.75	1.94	0.74	1.93
Mathematics Grade 3									
1. Number Sense	24	0.80	2.15	0.81	2.14	0.81	2.13	0.85	2.01
2. Algebra & Data Analysis	13	0.71	1.53	0.72	1.51	0.72	1.49	0.79	1.45
3. Measurement & Geometry	11	0.63	1.42	0.64	1.41	0.67	1.37	0.71	1.30
Mathematics Grade 4									
1. Number Sense	23	0.73	2.08	0.75	2.08	0.79	2.03	0.68	2.05
2. Algebra & Data Analysis	15	0.56	1.82	0.57	1.82	0.63	1.81	0.55	1.83
3. Measurement & Geometry	10	0.43	1.45	0.44	1.45	0.52	1.43	0.44	1.39
Mathematics Grade 5									
1. Number Sense	21	0.74	1.96	0.76	1.97	0.73	1.90	0.78	1.89
2. Algebra & Data Analysis	17	0.73	1.82	0.74	1.82	0.74	1.76	0.73	1.78
3. Measurement & Geometry	10	0.47	1.48	0.51	1.47	0.53	1.44	0.38	1.47
Science Grade 5									
1. Physical Sciences	16	0.52	1.77	0.56	1.72	0.52	1.71	0.45	1.76
2. Life Sciences	16	0.62	1.79	0.68	1.72	0.66	1.70	0.56	1.72
3. Earth Sciences	16	0.63	1.84	0.68	1.79	0.64	1.75	0.60	1.76

Table 8.B.24 Subscore Reliabilities and SEM by Ethnicity

Subscore Area	# of Items	African American		American Indian		Asian		Filipino		Hispanic		Pacific Islander		White	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3															
1. Vocabulary	14	0.73	1.62	0.77	1.60	0.76	1.57	0.70	1.57	0.71	1.66	0.75	1.59	0.75	1.56
2. Reading for Understanding	17	0.72	1.85	0.77	1.79	0.71	1.87	0.70	1.85	0.69	1.88	0.68	1.86	0.74	1.81
3. Language	17	0.68	1.90	0.73	1.88	0.69	1.88	0.67	1.89	0.63	1.93	0.65	1.92	0.69	1.89
ELA Grade 4															
1. Vocabulary	11	0.65	1.48	0.59	1.48	0.68	1.45	0.64	1.43	0.66	1.47	0.65	1.47	0.66	1.41
2. Reading for Understanding	16	0.60	1.87	0.65	1.83	0.61	1.86	0.61	1.86	0.57	1.87	0.65	1.87	0.65	1.83
3. Language	21	0.68	2.14	0.70	2.11	0.67	2.10	0.69	2.07	0.66	2.15	0.71	2.10	0.71	2.10
ELA Grade 5															
1. Vocabulary	8	0.56	1.17	0.49	1.14	0.54	1.18	0.54	1.15	0.54	1.19	0.52	1.17	0.57	1.11
2. Reading for Understanding	18	0.62	1.99	0.63	1.98	0.59	1.99	0.70	1.95	0.62	1.99	0.64	1.97	0.68	1.94
3. Language	22	0.74	2.03	0.72	2.00	0.73	2.01	0.78	1.94	0.73	2.03	0.77	1.96	0.76	1.96
Mathematics Grade 3															
1. Number Sense	24	0.79	2.18	0.82	2.12	0.85	2.05	0.84	2.11	0.80	2.16	0.78	2.19	0.81	2.12
2. Algebra & Data Analysis	13	0.69	1.56	0.75	1.50	0.73	1.51	0.74	1.49	0.71	1.52	0.76	1.54	0.72	1.49
3. Measurement & Geometry	11	0.64	1.44	0.64	1.40	0.70	1.38	0.68	1.39	0.63	1.42	0.63	1.43	0.65	1.38
Mathematics Grade 4															
1. Number Sense	23	0.73	2.13	0.74	2.06	0.80	2.01	0.79	1.99	0.73	2.09	0.74	2.10	0.76	2.04
2. Algebra & Data Analysis	15	0.55	1.84	0.40	1.86	0.64	1.80	0.56	1.82	0.56	1.83	0.52	1.84	0.60	1.81
3. Measurement & Geometry	10	0.39	1.48	0.48	1.44	0.50	1.44	0.47	1.43	0.43	1.45	0.56	1.44	0.44	1.44
Mathematics Grade 5															
1. Number Sense	21	0.74	2.02	0.72	1.99	0.79	1.90	0.78	1.92	0.74	1.96	0.73	1.97	0.76	1.95
2. Algebra & Data Analysis	17	0.71	1.87	0.69	1.84	0.74	1.81	0.74	1.81	0.73	1.82	0.72	1.82	0.76	1.79
3. Measurement & Geometry	10	0.49	1.48	0.49	1.48	0.47	1.47	0.50	1.46	0.48	1.48	0.55	1.47	0.53	1.46
Science Grade 5															
1. Physical Sciences	16	0.55	1.78	0.51	1.73	0.51	1.76	0.56	1.73	0.54	1.75	0.51	1.72	0.55	1.67
2. Life Sciences	16	0.67	1.79	0.70	1.71	0.62	1.81	0.62	1.79	0.65	1.77	0.66	1.74	0.67	1.67
3. Earth Sciences	16	0.64	1.85	0.67	1.78	0.65	1.82	0.65	1.78	0.64	1.83	0.63	1.82	0.69	1.74

Table 8.B.25 Subscore Reliabilities and SEM by Ethnicity—Not Economically Disadvantaged

Subscore Area	# of Items	African American		American Indian		Asian		Filipino		Hispanic		Pacific Islander		White	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3															
1. Vocabulary	14	0.72	1.58	0.67	1.49	0.76	1.49	0.70	1.52	0.73	1.59	0.78	1.48	0.75	1.52
2. Reading for Understanding	17	0.73	1.83	0.79	1.75	0.72	1.83	0.72	1.85	0.72	1.83	0.73	1.83	0.73	1.80
3. Language	17	0.72	1.87	0.74	1.88	0.75	1.81	0.67	1.89	0.67	1.90	0.68	1.84	0.69	1.87
ELA Grade 4															
1. Vocabulary	11	0.68	1.44	0.62	1.51	0.60	1.44	0.64	1.42	0.67	1.41	0.61	1.44	0.65	1.38
2. Reading for Understanding	16	0.66	1.85	0.62	1.82	0.59	1.87	0.59	1.86	0.64	1.84	0.71	1.85	0.66	1.82
3. Language	21	0.70	2.11	0.73	2.10	0.66	2.10	0.66	2.06	0.69	2.11	0.69	2.07	0.71	2.07
ELA Grade 5															
1. Vocabulary	8	0.54	1.15	0.55	1.03	0.47	1.15	0.46	1.15	0.54	1.14	0.61	1.10	0.54	1.08
2. Reading for Understanding	18	0.65	1.97	0.71	1.88	0.61	1.98	0.69	1.95	0.65	1.97	0.72	1.93	0.69	1.92
3. Language	22	0.74	2.00	0.80	1.91	0.74	1.98	0.75	1.92	0.74	1.98	0.76	1.97	0.74	1.93
Mathematics Grade 3															
1. Number Sense	24	0.78	2.15	0.85	2.01	0.85	2.01	0.82	2.09	0.79	2.13	0.62	2.22	0.81	2.10
2. Algebra & Data Analysis	13	0.66	1.57	0.73	1.36	0.72	1.50	0.72	1.49	0.69	1.51	0.74	1.58	0.72	1.47
3. Measurement & Geometry	11	0.63	1.42	0.74	1.33	0.76	1.32	0.65	1.39	0.63	1.39	0.57	1.44	0.65	1.37
Mathematics Grade 4															
1. Number Sense	23	0.75	2.11	0.76	2.03	0.81	1.95	0.81	1.97	0.75	2.06	0.81	2.08	0.77	2.02
2. Algebra & Data Analysis	15	0.55	1.83	0.37	1.86	0.70	1.76	0.55	1.83	0.59	1.81	0.49	1.85	0.61	1.80
3. Measurement & Geometry	10	0.37	1.48	0.30	1.48	0.46	1.45	0.47	1.43	0.45	1.44	0.51	1.45	0.46	1.43
Mathematics Grade 5															
1. Number Sense	21	0.77	1.98	0.68	2.02	0.79	1.88	0.79	1.88	0.74	1.96	0.75	1.95	0.76	1.93
2. Algebra & Data Analysis	17	0.71	1.86	0.68	1.85	0.73	1.82	0.74	1.77	0.74	1.80	0.73	1.82	0.76	1.77
3. Measurement & Geometry	10	0.52	1.47	0.45	1.49	0.45	1.48	0.52	1.44	0.49	1.47	0.28	1.52	0.53	1.45
Science Grade 5															
1. Physical Sciences	16	0.55	1.76	0.63	1.66	0.52	1.76	0.44	1.74	0.54	1.72	0.49	1.71	0.53	1.64
2. Life Sciences	16	0.71	1.75	0.69	1.66	0.69	1.78	0.63	1.78	0.67	1.73	0.64	1.76	0.66	1.65
3. Earth Sciences	16	0.61	1.84	0.65	1.71	0.66	1.80	0.65	1.77	0.66	1.79	0.68	1.80	0.68	1.71

Table 8.B.26 Subscore Reliabilities and SEM by Ethnicity—Economically Disadvantaged

Subscore Area	# of Items	African American		American Indian		Asian		Filipino		Hispanic		Pacific Islander		White	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3															
1. Vocabulary	14	0.73	1.63	0.78	1.62	0.74	1.64	0.68	1.64	0.70	1.67	0.72	1.62	0.74	1.61
2. Reading for Understanding	17	0.72	1.86	0.76	1.82	0.70	1.89	0.71	1.86	0.68	1.89	0.65	1.89	0.75	1.82
3. Language	17	0.67	1.90	0.70	1.90	0.62	1.93	0.66	1.90	0.62	1.93	0.58	1.97	0.68	1.90
ELA Grade 4															
1. Vocabulary	11	0.64	1.49	0.58	1.47	0.71	1.45	0.65	1.44	0.65	1.48	0.66	1.49	0.66	1.45
2. Reading for Understanding	16	0.58	1.88	0.66	1.83	0.62	1.86	0.64	1.86	0.55	1.88	0.61	1.87	0.62	1.85
3. Language	21	0.67	2.15	0.69	2.11	0.68	2.11	0.72	2.08	0.64	2.15	0.71	2.12	0.68	2.13
ELA Grade 5															
1. Vocabulary	8	0.57	1.18	0.47	1.16	0.57	1.19	0.63	1.15	0.54	1.20	0.46	1.21	0.58	1.13
2. Reading for Understanding	18	0.61	2.00	0.58	2.00	0.57	1.99	0.73	1.94	0.61	1.99	0.60	1.99	0.66	1.96
3. Language	22	0.74	2.04	0.68	2.03	0.72	2.03	0.80	1.98	0.73	2.04	0.77	1.96	0.76	1.99
Mathematics Grade 3															
1. Number Sense	24	0.80	2.18	0.80	2.16	0.84	2.08	0.85	2.12	0.80	2.16	0.82	2.17	0.81	2.14
2. Algebra & Data Analysis	13	0.69	1.56	0.74	1.54	0.74	1.52	0.79	1.48	0.72	1.52	0.77	1.53	0.71	1.51
3. Measurement & Geometry	11	0.63	1.45	0.62	1.42	0.64	1.43	0.68	1.41	0.62	1.43	0.65	1.42	0.66	1.39
Mathematics Grade 4															
1. Number Sense	23	0.73	2.14	0.73	2.07	0.78	2.05	0.77	2.03	0.73	2.09	0.66	2.12	0.74	2.06
2. Algebra & Data Analysis	15	0.55	1.84	0.42	1.86	0.57	1.83	0.58	1.82	0.55	1.83	0.52	1.83	0.58	1.82
3. Measurement & Geometry	10	0.40	1.48	0.53	1.42	0.52	1.44	0.49	1.44	0.43	1.45	0.59	1.43	0.41	1.45
Mathematics Grade 5															
1. Number Sense	21	0.73	2.03	0.73	1.99	0.79	1.91	0.74	2.01	0.74	1.96	0.72	1.98	0.76	1.98
2. Algebra & Data Analysis	17	0.71	1.87	0.69	1.83	0.75	1.81	0.73	1.86	0.73	1.82	0.72	1.82	0.75	1.81
3. Measurement & Geometry	10	0.48	1.48	0.48	1.47	0.48	1.48	0.46	1.48	0.47	1.48	0.64	1.42	0.51	1.47
Science Grade 5															
1. Physical Sciences	16	0.55	1.78	0.48	1.75	0.51	1.76	0.67	1.70	0.53	1.76	0.51	1.72	0.54	1.70
2. Life Sciences	16	0.66	1.80	0.68	1.72	0.56	1.83	0.61	1.80	0.64	1.78	0.67	1.74	0.68	1.70
3. Earth Sciences	16	0.64	1.85	0.65	1.80	0.64	1.83	0.67	1.79	0.63	1.83	0.60	1.83	0.70	1.77

Table 8.B.27 Subscore Reliabilities and SEM by Economic Status

Subscore Area	# of Items	Economically Disadvantaged		Not Economically Disadvantaged		Unknown Economic Status	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3							
1. Vocabulary	14	0.71	1.65	0.74	1.55	0.76	1.56
2. Reading for Understanding	17	0.70	1.87	0.73	1.82	0.73	1.81
3. Language	17	0.64	1.93	0.69	1.88	0.70	1.88
ELA Grade 4							
1. Vocabulary	11	0.65	1.48	0.66	1.40	0.63	1.48
2. Reading for Understanding	16	0.57	1.88	0.66	1.83	0.68	1.80
3. Language	21	0.66	2.15	0.70	2.09	0.74	2.10
ELA Grade 5							
1. Vocabulary	8	0.55	1.19	0.54	1.11	0.55	1.28
2. Reading for Understanding	18	0.62	1.99	0.68	1.95	0.62	1.99
3. Language	22	0.73	2.03	0.75	1.96	0.71	2.12
Mathematics Grade 3							
1. Number Sense	24	0.80	2.16	0.81	2.11	0.80	2.13
2. Algebra & Data Analysis	13	0.72	1.53	0.71	1.50	0.73	1.49
3. Measurement & Geometry	11	0.63	1.43	0.65	1.38	0.64	1.39
Mathematics Grade 4							
1. Number Sense	23	0.73	2.09	0.77	2.04	0.73	2.13
2. Algebra & Data Analysis	15	0.56	1.83	0.60	1.81	0.46	1.85
3. Measurement & Geometry	10	0.43	1.46	0.45	1.44	-0.10	1.51
Mathematics Grade 5							
1. Number Sense	21	0.74	1.97	0.76	1.95	0.77	2.02
2. Algebra & Data Analysis	17	0.73	1.83	0.75	1.80	0.73	1.86
3. Measurement & Geometry	10	0.48	1.48	0.51	1.46	0.58	1.47
Science Grade 5							
1. Physical Sciences	16	0.54	1.76	0.54	1.69	0.46	1.84
2. Life Sciences	16	0.65	1.77	0.68	1.70	0.68	1.81
3. Earth Sciences	16	0.65	1.83	0.68	1.76	0.63	1.86

Table 8.B.28 Subscore Reliabilities and SEM by Disability

Subscore Area	# of Items	Autism		Deafness		Emotional Disturbance		Hard of Hearing		Mental Retardation		Multiple Disability	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3													
1. Vocabulary	14	0.76	1.54	0.50	1.78	0.76	1.57	0.70	1.66	0.65	1.73	0.56	1.76
2. Reading for Understanding	17	0.72	1.86	0.44	1.93	0.80	1.78	0.60	1.92	0.56	1.94	0.77	1.87
3. Language	17	0.67	1.89	0.39	1.97	0.74	1.87	0.60	1.92	0.52	1.93	0.55	1.96
ELA Grade 4													
1. Vocabulary	11	0.63	1.47	0.60	1.54	0.70	1.40	0.64	1.48	0.55	1.55	0.55	1.56
2. Reading for Understanding	16	0.66	1.85	0.50	1.89	0.68	1.83	0.56	1.87	0.40	1.90	0.49	1.89
3. Language	21	0.68	2.11	0.55	2.15	0.73	2.12	0.65	2.09	0.50	2.18	0.38	2.19
ELA Grade 5													
1. Vocabulary	8	0.57	1.19	0.44	1.33	0.62	1.11	0.54	1.21	0.51	1.30	0.53	1.22
2. Reading for Understanding	18	0.62	1.99	0.53	1.99	0.68	1.97	0.61	1.97	0.40	2.00	0.41	2.03
3. Language	22	0.74	2.04	0.73	2.10	0.78	2.01	0.75	1.95	0.66	2.20	0.76	2.09
Mathematics Grade 3													
1. Number Sense	24	0.81	2.16	0.79	2.17	0.82	2.16	0.77	2.14	0.69	2.25	0.82	2.20
2. Algebra & Data Analysis	13	0.70	1.56	0.68	1.58	0.75	1.53	0.60	1.55	0.60	1.67	0.42	1.67
3. Measurement & Geometry	11	0.63	1.44	0.59	1.50	0.66	1.42	0.60	1.39	0.53	1.56	0.48	1.51
Mathematics Grade 4													
1. Number Sense	23	0.79	2.08	0.72	2.10	0.77	2.12	0.75	2.02	0.60	2.25	0.59	2.26
2. Algebra & Data Analysis	15	0.63	1.81	0.67	1.77	0.61	1.82	0.55	1.82	0.22	1.85	0.53	1.84
3. Measurement & Geometry	10	0.44	1.46	0.45	1.44	0.39	1.47	0.39	1.44	0.16	1.51	0.32	1.47
Mathematics Grade 5													
1. Number Sense	21	0.79	1.99	0.79	1.93	0.74	2.02	0.76	1.88	0.63	2.12	0.74	2.05
2. Algebra & Data Analysis	17	0.76	1.85	0.81	1.80	0.76	1.83	0.76	1.76	0.43	1.95	0.69	1.90
3. Measurement & Geometry	10	0.49	1.49	0.51	1.46	0.51	1.47	0.63	1.40	0.24	1.51	0.25	1.54
Science Grade 5													
1. Physical Sciences	16	0.60	1.73	0.58	1.80	0.64	1.71	0.42	1.74	0.51	1.87	0.39	1.78
2. Life Sciences	16	0.69	1.78	0.72	1.80	0.73	1.73	0.59	1.79	0.59	1.87	0.62	1.83
3. Earth Sciences	16	0.70	1.79	0.65	1.84	0.73	1.78	0.61	1.80	0.50	1.89	0.44	1.91

Table 8.B.29 Subscore Reliabilities and SEM by Disability (continued)

Subscore Area	# of Items	Orthopedic Impairment		Other Health Impairment		Specific Learning Disability		Speech or Language Impairment		Traumatic Brain Injury		Visual Impairment	
		Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM	Reliab.	SEM
ELA Grade 3													
1. Vocabulary	14	0.70	1.66	0.72	1.59	0.73	1.63	0.72	1.63	0.77	1.62	0.65	1.62
2. Reading for Understanding	17	0.74	1.83	0.72	1.83	0.72	1.85	0.69	1.87	0.68	1.90	0.69	1.84
3. Language	17	0.69	1.88	0.66	1.91	0.66	1.92	0.66	1.91	0.56	1.95	0.65	1.90
ELA Grade 4													
1. Vocabulary	11	0.69	1.46	0.66	1.40	0.66	1.47	0.65	1.45	0.63	1.49	0.69	1.48
2. Reading for Understanding	16	0.66	1.86	0.65	1.85	0.61	1.86	0.57	1.87	0.46	1.92	0.58	1.88
3. Language	21	0.72	2.11	0.72	2.10	0.68	2.14	0.66	2.12	0.72	2.08	0.76	2.11
ELA Grade 5													
1. Vocabulary	8	0.63	1.18	0.54	1.11	0.55	1.17	0.54	1.16	0.47	1.19	0.48	1.20
2. Reading for Understanding	18	0.63	1.99	0.65	1.97	0.64	1.98	0.60	1.99	0.68	1.96	0.55	2.02
3. Language	22	0.77	2.03	0.76	1.97	0.73	2.01	0.73	1.99	0.83	1.98	0.82	2.03
Mathematics Grade 3													
1. Number Sense	24	0.82	2.16	0.78	2.18	0.80	2.14	0.81	2.13	0.81	2.14	0.85	2.13
2. Algebra & Data Analysis	13	0.75	1.53	0.71	1.53	0.72	1.50	0.71	1.52	0.74	1.55	0.74	1.50
3. Measurement & Geometry	11	0.68	1.44	0.65	1.41	0.63	1.40	0.65	1.40	0.64	1.39	0.70	1.38
Mathematics Grade 4													
1. Number Sense	23	0.77	2.15	0.75	2.08	0.73	2.08	0.75	2.06	0.69	2.11	0.74	2.17
2. Algebra & Data Analysis	15	0.59	1.83	0.56	1.83	0.56	1.82	0.58	1.82	0.59	1.81	0.61	1.82
3. Measurement & Geometry	10	0.48	1.45	0.44	1.44	0.43	1.45	0.45	1.45	0.56	1.46	0.65	1.38
Mathematics Grade 5													
1. Number Sense	21	0.76	2.05	0.76	1.97	0.73	1.96	0.76	1.94	0.82	1.95	0.76	1.99
2. Algebra & Data Analysis	17	0.76	1.86	0.76	1.80	0.73	1.81	0.74	1.82	0.81	1.80	0.71	1.85
3. Measurement & Geometry	10	0.50	1.46	0.47	1.47	0.50	1.47	0.48	1.47	0.28	1.51	0.56	1.45
Science Grade 5													
1. Physical Sciences	16	0.50	1.81	0.55	1.72	0.54	1.73	0.50	1.75	0.59	1.81	0.54	1.77
2. Life Sciences	16	0.68	1.80	0.66	1.72	0.66	1.74	0.63	1.77	0.68	1.81	0.80	1.70
3. Earth Sciences	16	0.65	1.84	0.68	1.79	0.65	1.81	0.63	1.82	0.77	1.78	0.56	1.88

Table 8.B.30 Reliability of Classification for ELA, Grade Three

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	40 – 48	0.06	0.03	0.00	0.00	0.00	0.09
	35 – 39	0.01	0.07	0.05	0.00	0.00	0.14
	28 – 34	0.00	0.03	0.17	0.06	0.00	0.25
	18 – 27	0.00	0.00	0.05	0.29	0.03	0.37
	0 – 17	0.00	0.00	0.00	0.05	0.10	0.16
Estimated Proportion Correctly Classified: Total = 0.69, Proficient & Above =0.92							
Decision Consistency	40 – 48	0.06	0.02	0.00	0.00	0.00	0.09
	35 – 39	0.03	0.06	0.05	0.01	0.00	0.14
	28 – 34	0.00	0.04	0.13	0.07	0.00	0.25
	18 – 27	0.00	0.00	0.07	0.25	0.05	0.37
	0 – 17	0.00	0.00	0.00	0.06	0.09	0.16
Estimated Proportion Consistently Classified: Total = 0.59, Proficient & Above = 0.89							

Table 8.B.31 Reliability of Classification for ELA, Grade Four (Reading only)

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	36 – 48	0.09	0.04	0.00	0.00	0.00	0.13
	30 – 35	0.02	0.11	0.05	0.00	0.00	0.19
	24 – 29	0.00	0.04	0.14	0.06	0.00	0.25
	18 – 23	0.00	0.00	0.05	0.17	0.05	0.27
	0 – 19	0.00	0.00	0.00	0.07	0.10	0.17
Estimated Proportion Correctly Classified: Total = 0.61, Proficient & Above = 0.90							
Decision Consistency	36 – 48	0.09	0.04	0.00	0.00	0.00	0.13
	30 – 35	0.04	0.09	0.05	0.01	0.00	0.19
	24 – 29	0.01	0.05	0.11	0.07	0.01	0.25
	18 – 23	0.00	0.01	0.06	0.13	0.07	0.27
	0 – 19	0.00	0.00	0.01	0.07	0.09	0.17
Estimated Proportion Consistently Classified: Total = 0.50, Proficient & Above = 0.87							

Table 8.B.32 Reliability of Classification for ELA, Grade Four (Reading & Writing)

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	39 – 52	0.08	0.04	0.01	0.00	0.00	0.13
	33 – 38	0.02	0.09	0.06	0.01	0.00	0.19
	27 – 32	0.00	0.04	0.13	0.07	0.00	0.24
	20 – 26	0.00	0.00	0.07	0.18	0.03	0.29
	0 – 19	0.00	0.00	0.01	0.08	0.07	0.16
Estimated Proportion Correctly Classified: Total = 0.55, Proficient & Above = 0.87							
Decision Consistency	39 – 52	0.08	0.04	0.01	0.00	0.00	0.13
	33 – 38	0.04	0.07	0.06	0.02	0.00	0.19
	27 – 32	0.01	0.05	0.09	0.07	0.01	0.24
	20 – 26	0.00	0.02	0.07	0.13	0.06	0.29
	0 – 19	0.00	0.00	0.02	0.07	0.07	0.16
Estimated Proportion Consistently Classified: Total = 0.50, Proficient & Above = 0.82							

Table 8.B.33 Reliability of Classification for ELA, Grade Five

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	38 – 48	0.09	0.04	0.00	0.00	0.00	0.14
	33 – 37	0.03	0.12	0.06	0.00	0.00	0.21
	27 – 32	0.00	0.05	0.16	0.05	0.00	0.27
	17 – 26	0.00	0.00	0.06	0.26	0.01	0.33
	0 – 16	0.00	0.00	0.00	0.03	0.02	0.06
Estimated Proportion Correctly Classified: Total =0.66, Proficient & Above = 0.89							
Decision Consistency	38 – 48	0.09	0.04	0.01	0.00	0.00	0.14
	33 – 37	0.05	0.09	0.06	0.01	0.00	0.21
	27 – 32	0.01	0.06	0.13	0.07	0.00	0.27
	17 – 26	0.00	0.01	0.07	0.22	0.03	0.33
	0 – 16	0.00	0.00	0.00	0.03	0.02	0.06
Estimated Proportion Consistently Classified: Total = 0.55, Proficient & Above = 0.84							

Table 8.B.34 Reliability of Classification for Mathematics, Grade Three

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	42 – 48	0.04	0.03	0.00	0.00	0.00	0.08
	35 – 41	0.02	0.17	0.05	0.00	0.00	0.24
	28 – 34	0.00	0.04	0.16	0.05	0.00	0.25
	17 – 27	0.00	0.00	0.04	0.28	0.02	0.34
	0 – 16	0.00	0.00	0.00	0.04	0.05	0.09
Estimated Proportion Correctly Classified: Total = 0.71, Proficient & Above = 0.91							
Decision Consistency	42 – 48	0.04	0.03	0.00	0.00	0.00	0.08
	35 – 41	0.04	0.14	0.06	0.00	0.00	0.24
	28 – 34	0.00	0.06	0.13	0.06	0.00	0.25
	17 – 27	0.00	0.00	0.06	0.24	0.04	0.34
	0 – 16	0.00	0.00	0.00	0.04	0.05	0.09
Estimated Proportion Consistently Classified: Total = 0.60, Proficient & Above = 0.88							

Table 8.B.35 Reliability of Classification for Mathematics, Grade Four

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	37 – 48	0.07	0.04	0.00	0.00	0.00	0.11
	30 – 36	0.02	0.16	0.06	0.00	0.00	0.25
	25 – 29	0.00	0.05	0.13	0.06	0.00	0.24
	18 – 24	0.00	0.00	0.06	0.21	0.02	0.30
	0 – 17	0.00	0.00	0.00	0.05	0.05	0.10
Estimated Proportion Correctly Classified: Total = 0.62, Proficient & Above = 0.88							
Decision Consistency	37 – 48	0.07	0.04	0.00	0.00	0.00	0.11
	30 – 36	0.04	0.13	0.06	0.02	0.00	0.25
	25 – 29	0.00	0.06	0.10	0.07	0.01	0.24
	18 – 24	0.00	0.02	0.07	0.16	0.05	0.30
	0 – 17	0.00	0.00	0.01	0.05	0.05	0.10
Estimated Proportion Consistently Classified: Total = 0.51, Proficient & Above = 0.83							

Table 8.B.36 Reliability of Classification for Mathematics, Grade Five

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	39 – 48	0.09	0.03	0.00	0.00	0.00	0.12
	32 – 38	0.02	0.16	0.06	0.00	0.00	0.25
	26 – 31	0.00	0.04	0.15	0.05	0.00	0.25
	18 – 25	0.00	0.00	0.06	0.22	0.02	0.30
	0 – 17	0.00	0.00	0.00	0.05	0.05	0.09
Estimated Proportion Correctly Classified: Total = 0.66, Proficient & Above = 0.90							
Decision Consistency	39 – 48	0.08	0.04	0.00	0.00	0.00	0.12
	32 – 38	0.04	0.13	0.06	0.01	0.00	0.25
	26 – 31	0.00	0.06	0.12	0.07	0.00	0.25
	18 – 25	0.00	0.01	0.07	0.17	0.05	0.30
	0 – 17	0.00	0.00	0.00	0.04	0.04	0.09
Estimated Proportion Consistently Classified: Total = 0.55, Proficient & Above = 0.86							

Table 8.B.37 Reliability of Classification for Science, Grade Five

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total
Decision Accuracy	38 – 48	0.09	0.05	0.00	0.00	0.00	0.14
	32 – 37	0.03	0.19	0.07	0.00	0.00	0.29
	25 – 31	0.00	0.06	0.21	0.04	0.00	0.31
	18 – 24	0.00	0.00	0.06	0.13	0.01	0.20
	0 – 17	0.00	0.00	0.00	0.03	0.03	0.06
Estimated Proportion Correctly Classified: Total = 0.64, Proficient & Above = 0.87							
Decision Consistency	38 – 48	0.08	0.05	0.01	0.00	0.00	0.14
	32 – 37	0.06	0.14	0.08	0.01	0.00	0.29
	25 – 31	0.01	0.07	0.17	0.06	0.00	0.31
	18 – 24	0.00	0.01	0.06	0.10	0.03	0.20
	0 – 17	0.00	0.00	0.00	0.03	0.03	0.06
Estimated Proportion Consistently Classified: Total = 0.52, Proficient & Above = 0.82							

Table 8.B.38 Inter-Rater Analyses for ELA, Grade Four

		Rater 1								
		1	**%**	**2**	**%**	**3**	**%**	**4**	**%**	**Total**
	1	114	8.93	45	3.53	4	0.31	0	0.00	163
	2	52	4.08	429	33.62	90	7.05	2	0.16	573
	3	2	0.16	128	10.03	275	21.55	45	3.53	450
Rater 2	**4**	0	0.00	4	0.31	41	3.21	45	3.53	90
	Total	168	13.17	606	47.49	410	32.13	92	7.21	1,276

Percent Exact = 67.63%

Percent Adjacent = 31.43%

Percent Exact + Adjacent = 99.06%

Kappa = 0.50

Weighted Kappa = 0.60

Table 8.B.39 Inter-Rater Analyses for ELA, Grade Seven

		Rater 1								
		1	**%**	**2**	**%**	**3**	**%**	**4**	**%**	**Total**
	1	61	5.43	23	2.05	1	0.09	0	0.00	85
	2	21	1.87	382	33.99	120	10.68	0	0.00	523
	3	1	0.09	93	8.27	309	27.49	35	3.11	438
Rater 2	**4**	0	0.00	5	0.44	31	2.76	42	3.74	78
	Total	83	7.38	503	44.75	461	41.01	77	6.85	1,124

Percent Exact = 70.64%
Percent Adjacent = 28.74%
Percent Exact + Adjacent = 99.38%
Kappa = 0.53
Weighted Kappa = 0.61

Table 8.B.40 Descriptive Statistics for the Ratings by the Two Raters

Grade	N*	Rater 1		Rater 2		
		Mean	**Std**	**Mean**	**Std**	***Pearson Correlation***
4	1,276	2.33	0.79	2.37	0.79	0.72
7	1,124	2.47	0.73	2.45	0.73	0.71

* Number of students who received valid ratings of 1–4

Table 8.B.41 Generalizability Analyses for Grade Four Essay—[(Person: Essay) x Rater)]

Effect	Degrees of Freedom	Sum of Squares	Mean Squares	Estimated Variance Components	Percentage of Total Variance
E	1	0.1490	0.14897	0.00000	0.00
P:E	1,274	1,381.8710	1.08467	0.45454	72.02
R	1	0.6587	0.65870	0.00099	0.16
ER	1	0.1344	0.13441	0.00000	0.00
PR:E	1,274	223.7069	0.17559	0.17559	27.82
Note: E = Essay Prompt, P = Person, and R = Rater				**G-Coefficient= 0.84**	

Table 8.B.42 Generalizability Analyses for Grade Seven Essay—[(Person: Essay) x Rater]

Effect	Degrees of Freedom	Sum of Squares	Mean Squares	Estimated Variance Components	Percentage of Total Variance
E	1	0.0192	0.01921	0.00000	0.00
P:E	1,122	1,029.4163	0.91748	0.38066	70.62
R	1	0.2353	0.23532	0.00218	0.40
ER	1	0.0398	0.03976	0.00000	0.00
PR:E	1,122	175.2249	0.15617	0.15617	28.97
Note: E = Essay Prompt, P = Person, and R = Rater				**G-Coefficient= 0.83**	

Appendix 8.C—Validity Analyses

Table 8.C.1 2009 CMA Content Area Correlations (All Valid Scores)

Grade		ELA	Mathematics	Science
3	**ELA**	**11,321**	0.62	
	Mathematics	11,146	**11,227**	
4	**ELA**	**17,121**	0.54	
	Mathematics	13,772	**14,415**	
5	**ELA**	**17,821**	0.53	0.65
	Mathematics	14,677	**15,656**	0.53
	Science	16,254	14,364	**16,681**

Table 8.C.2 2009 CMA Content Area Correlations (Gender)

		Male			Female			Gender Unknown		
Grade		ELA	Mathematics	Science	ELA	Mathematics	Science	ELA	Mathematics	Science
3	**ELA**	**7,512**	0.62		**3,729**	0.62		**80**	0.67	
	Mathematics	7,399	**7,454**		3,684	**3,706**		63	**67**	
4	**ELA**	**11,444**	0.54		**5,662**	0.54		**15**	0.69	
	Mathematics	8,994	**9,365**		4,767	**5,039**		11	**11**	
5	**ELA**	**11,862**	0.52	0.65	**5,948**	0.54	0.66	**11**	0.64	0.76
	Mathematics	9,639	**10,193**	0.53	5,027	**5,452**	0.53	11	**11**	0.48
	Science	10,795	9,426	**11,052**	5,448	4,927	**5,617**	11	11	**12**

Table 8.C.3 2009 CMA Content Area Correlations (Primary Ethnicity)

		American Indian			Asian American			Pacific Islander		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**135**	0.67		**381**	0.63		**70**	0.49	
	Mathematics	129	**132**		373	**378**		70	**72**	
4	**ELA**	**203**	0.48		**545**	0.56		**114**	0.68	
	Mathematics	173	**183**		421	**432**		95	**99**	
5	**ELA**	**227**	0.53	0.59	**527**	0.51	0.66	**104**	0.5	0.65
	Mathematics	186	**198**	0.5	409	**426**	0.54	82	**89**	0.38
	Science	208	181	**213**	472	397	**484**	98	83	**103**

		Filipino			Hispanic			African American		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**156**	0.7		**6,718**	0.61		**1,282**	0.65	
	Mathematics	156	**157**		6,637	**6,670**		1,264	**1,272**	
4	**ELA**	**230**	0.54		**10,173**	0.52		**1,888**	0.58	
	Mathematics	181	**192**		8,301	**8,593**		1,619	**1,721**	
5	**ELA**	**215**	0.58	0.67	**10,763**	0.52	0.64	**2,069**	0.58	0.64
	Mathematics	166	**178**	0.56	8,925	**9,329**	0.52	1,816	**1,956**	0.56
	Science	196	162	**202**	9,790	8,625	**9,960**	1,908	1,805	**1,979**

		White			Unknown		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**2,294**	0.63		**285**	0.65	
	Mathematics	2,263	**2,278**		254	**268**	
4	**ELA**	**3,800**	0.53		**168**	0.58	
	Mathematics	2,854	**3,065**		128	**130**	
5	**ELA**	**3,769**	0.51	0.64	**147**	0.57	0.76
	Mathematics	2,971	**3,348**	0.53	122	**132**	0.61
	Science	3,448	2,988	**3,598**	134	123	**142**

Table 8.C.4 2009 CMA Content Area Correlations (English Fluency)

		English Only			Initially Fluent English Proficient			English Learner		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**6,153**	0.64		**145**	0.57		**4,682**	0.6	
	Mathematics	6,060	**6,097**		142	**144**		4,632	**4,662**	
4	**ELA**	**9,560**	0.54		**340**	0.5		**7,078**	0.53	
	Mathematics	7,592	**8,044**		252	**267**		5,822	**5,988**	
5	**ELA**	**9,776**	0.53	0.64	**401**	0.45	0.59	**7,445**	0.53	0.64
	Mathematics	8,040	**8,757**	0.54	306	**336**	0.5	6,169	**6,389**	0.53
	Science	8,940	7,950	**9,243**	354	300	**365**	6,774	5,955	**6,881**

		Reclassified Fluent English Proficient			English Fluency unknown		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**37**	0.8		**304**	0.62	
	Mathematics	37	**37**		275	**287**	
4	**ELA**	**100**	0.63		**43**	0.48	
	Mathematics	70	**78**		36	**38**	
5	**ELA**	**156**	0.51	0.6	**43**	0.48	0.66
	Mathematics	122	**132**	0.38	40	**42**	0.56
	Science	145	118	**148**	41	41	**44**

Table 8.C.5 2009 CMA Content Area Correlations (Economic Status)

		Not Economically Disadvantaged			Economically Disadvantaged			Unknown Economic Status		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**2,522**	0.64		**8,313**	0.61		**486**	0.64	
	Mathematics	2,481	**2,500**		8,217	**8,267**		448	**460**	
4	**ELA**	**4,560**	0.55		**12,498**	0.52		**63**	0.64	
	Mathematics	3,443	**3,681**		10,280	**10,684**		49	**50**	
5	**ELA**	**4,560**	0.52	0.65	**13,191**	0.52	0.64	**70**	0.45	0.67
	Mathematics	3,622	**4,030**	0.52	10,990	**11,560**	0.53	65	**66**	0.51
	Science	4,135	3,613	**4,300**	12,053	10,688	**12,314**	66	63	**67**

Table 8.C.6 2009 CMA Content Area Correlations (Primary Disability)

		Autism			Mental Retardation			Specific Learning Disability		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**652**	0.65		**305**	0.59		**6,039**	0.6	
	Mathematics	645	**650**		300	**300**		5,951	**5,990**	
4	**ELA**	**792**	0.57		**284**	0.53		**9,899**	0.53	
	Mathematics	676	**708**		277	**283**		7,882	**8,221**	
5	**ELA**	**699**	0.57	0.71	**383**	0.44	0.62	**11,138**	0.5	0.63
	Mathematics	614	**664**	0.53	370	**371**	0.45	9,017	**9,584**	0.5
	Science	658	614	**679**	375	365	**376**	10,102	8,785	**10,346**

		Speech or Language Impairment			Other Health Impairment		
Grade		**ELA**	**Mathematics**	**Science**	**ELA**	**Mathematics**	**Science**
3	**ELA**	**2,302**	0.64		**780**	0.65	
	Mathematics	2,280	**2,295**		773	**781**	
4	**ELA**	**2,873**	0.56		**1,252**	0.54	
	Mathematics	2,264	**2,360**		995	**1,063**	
5	**ELA**	**2,417**	0.54	0.65	**1,313**	0.55	0.64
	Mathematics	1,971	**2,066**	0.55	1,094	**1,226**	0.56
	Science	2,212	1,932	**2,263**	1,185	1,090	**1,239**

Table 8.C.7 2009 CMA Content Area Correlations (Raw Score: All Valid Scores)

Grade		**ELA**	**Mathematics**
6	**ELA**	**15,900**	0.52
	Mathematics	13,862	**15,278**
7	**ELA**	**13,060**	0.51
	Mathematics	11,670	**13,145**
Grade		**ELA**	**Science**
8	**ELA**	**12,434**	0.59
	Science	10,940	**11,384**

Table 8.C.8 2009 CMA Content Area Correlations (Raw Score: Gender)

		Male		Female		Gender Unknown	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**10,587**	0.52	**5,305**	0.53	**8**	--
	Mathematics	9,121	**9,980**	4,733	**5,289**	8	**9**
7	**ELA**	**8,648**	0.51	**4,402**	0.51	**10**	--
	Mathematics	7,675	**8,548**	3,985	**4,587**	10	**10**
Grade		**ELA**	**Science**	**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**8,218**	0.61	**4,212**	0.57	**4**	--
	Science	7,184	**7,447**	3,753	**3,933**	3	**4**

Table 8.C.9 2009 CMA Content Area Correlations (Raw Score: Primary Ethnicity)

		American Indian		Asian American		Pacific Islander	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**195**	0.55	**502**	0.53	**95**	0.61
	Mathematics	171	**192**	401	**442**	79	**84**
7	**ELA**	**157**	0.57	**395**	0.5	**46**	0.38
	Mathematics	140	**154**	345	**383**	44	**50**
Grade		**ELA**	**Science**	**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**175**	0.54	**9,573**	0.52	**1,818**	0.5
	Science	155	**161**	8,348	**8,964**	1,675	**1,824**
		Filipino		**Hispanic**		**African American**	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**166**	0.54	**9,573**	0.52	**1,818**	0.5
	Mathematics	146	**162**	8,348	**8,964**	1,675	**1,824**
7	**ELA**	**135**	0.49	**7,631**	0.5	**1,589**	0.48
	Mathematics	124	**148**	6,834	**7,537**	1,479	**1,626**
Grade		**ELA**	**Science**	**ELA**	**Science**		
8	**ELA**	**125**	0.67	**7,185**	0.58		
	Science	101	**109**	6,303	**6,511**		

		White		Gender Unknown	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**3,417**	0.52	**134**	0.49
	Mathematics	2,921	**3,476**	121	**134**
7	**ELA**	**3,003**	0.53	**104**	0.51
	Mathematics	2,612	**3,141**	92	**106**
Grade		**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**2,970**	0.59	**97**	0.67
	Science	2,606	**2,763**	83	**85**

Table 8.C.10 2009 CMA Content Area Correlations (Raw Score: English Fluency)

		English Only		Initially Fluent English Proficient		English Learner	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**8,621**	0.52	**291**	0.51	**6,741**	0.53
	Mathematics	7,540	**8,548**	249	**288**	5,869	**6,196**
7	**ELA**	**7,380**	0.52	**268**	0.57	**5,111**	0.5
	Mathematics	6,600	**7,615**	237	**271**	4,573	**4,946**
Grade		**ELA**	**Science**	**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**7,139**	0.6	**339**	0.63	**4,692**	0.55
	Science	6,278	**6,599**	278	**283**	4,153	**4,252**

		Reclassified Fluent English Proficient		English Fluency unknown	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**218**	0.47	**29**	0.41
	Mathematics	177	**215**	27	**31**
7	**ELA**	**260**	0.54	**41**	0.49
	Mathematics	224	**273**	36	**40**
Grade		**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**234**	0.64	**30**	0.62
	Science	205	**222**	26	**28**

Table 8.C.11 2009 CMA Content Area Correlations (Raw Score: Economic Status)

		Not Economically Disadvantaged		Economically Disadvantaged		Unknown Economic Status	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**4,035**	0.51	**11,802**	0.52	**63**	0.5
	Mathematics	3,419	**4,022**	10,383	**11,191**	60	**65**
7	**ELA**	**3,592**	0.53	**9,400**	0.49	**68**	0.56
	Mathematics	3,137	**3,739**	8,473	**9,341**	60	**65**
Grade		**ELA**	**Science**	**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**3,635**	0.59	**8,748**	0.58	**51**	0.51
	Science	3,147	**3,352**	7,748	**7,986**	45	**46**

Table 8.C.12 2009 CMA Content Area Correlations (Raw Score: Primary Disability)

		Autism		Mental Retardation		Specific Learning Disability	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**499**	0.48	**387**	0.41	**10,577**	0.51
	Mathematics	455	**540**	382	**391**	9,117	**9,905**
7	**ELA**	**404**	0.53	**347**	0.37	**8,872**	0.5
	Mathematics	362	**415**	337	**357**	7,834	**8,785**
Grade		**ELA**	**Science**	**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**328**	0.65	**340**	0.46	**8,738**	0.58
	Science	296	**313**	315	**319**	7,654	**7,927**

		Speech or Language Impairment		Other Health Impairment	
Grade		**ELA**	**Mathematics**	**ELA**	**Mathematics**
6	**ELA**	**1,661**	0.53	**1,113**	0.52
	Mathematics	1,411	**1,569**	1,007	**1,191**
7	**ELA**	**1,179**	0.54	**929**	0.52
	Mathematics	1,068	**1,182**	845	**1,034**
Grade		**ELA**	**Science**	**ELA**	**Science**
8	**ELA**	**976**	0.56	**862**	0.61
	Science	871	**892**	731	**779**

Appendix 8.D—IRT Analyses

Table 8.D.1 IRT Model Data Fit Distribution for ELA, Grades Three through Eight

Items	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
Flag	**N**	**Pct.**	**N**	Pct.	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**
A	17	35%	16	33%	11	23%	18	33%	19	35%	27	50%
B	11	23%	13	27%	15	31%	15	28%	17	31%	10	19%
C	15	31%	14	29%	16	33%	15	28%	18	33%	15	28%
D	5	10%	6	12%	4	8%	5	9%	1	2%	1	2%
F	0	0%	0	0%	2	4%	1	2%	0	0%	1	2%
Total	48	100%	49	100%	48	100%	54	100%	55	100%	54	100%

Table 8.D.2 IRT Model Data Fit Distribution for Mathematics, Grades Three through Seven

Items	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7	
Flag	**N**	**Pct.**	**N**	Pct.	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**
A	12	25%	15	31%	12	25%	19	35%	31	57%
B	18	38%	16	33%	14	29%	21	39%	14	26%
C	17	35%	16	33%	19	40%	9	17%	8	15%
D	0	0%	1	2%	3	6%	5	9%	1	2%
F	1	2%	0	0%	0	0%	0	0%	0	0%
Total	48	100%	48	100%	48	100%	54	100%	54	100%

Table 8.D.3 IRT Model Data Fit Distribution for Science, Grades Five and Eight

Items	Grade 5		Grade 8	
Flag	**N**	**Pct.**	**N**	**Pct.**
A	14	29%	25	46%
B	15	31%	12	22%
C	11	23%	16	30%
D	7	15%	1	2%
F	1	2%	0	0%
Total	48	100%	54	100%

Table 8.D.4 IRT Model Data Fit Distribution for ELA, Grades Three through Eight (field-test items)

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7		Grade 8	
Flag	**N**	**Pct.**	**N**	Pct.	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**
A	16	30%	14	33%	11	22%	17	47%	10	28%	10	29%
B	12	22%	7	17%	10	20%	9	25%	9	25%	12	34%
C	11	20%	14	33%	20	39%	7	19%	14	39%	11	31%
D	12	22%	4	10%	9	18%	2	6%	1	3%	1	3%
F	3	6%	3	7%	1	2%	1	3%	2	6%	1	3%
Total	54	100%	42	100%	51	100%	36	100%	36	100%	35	100%

Table 8.D.5 IRT Model Data Fit Distribution for Mathematics, Grades Three through Seven (field-test items)

	Grade 3		Grade 4		Grade 5		Grade 6		Grade 7	
Flag	**N**	**Pct.**	**N**	Pct.	**N**	**Pct.**	**N**	**Pct.**	**N**	**Pct.**
A	7	13%	6	13%	8	15%	7	19%	22	61%
B	6	11%	15	33%	7	13%	14	39%	9	25%
C	16	30%	19	42%	16	30%	11	31%	5	14%
D	13	24%	3	7%	15	28%	2	6%	0	0%
F	12	22%	2	4%	8	15%	2	6%	0	0%

Table 8.D.6 IRT Model Data Fit Distribution for Science, Grades Five and Eight (field-test items)

	Grade 5		Grade 8	
Flag	**N**	**Pct.**	**N**	**Pct.**
A	7	13%	16	30%
B	11	20%	17	31%
C	22	41%	13	24%
D	10	19%	3	6%
F	4	7%	5	9%
Total	54	100%	54	100%

Table 8.D.7 IRT *b*-values for ELA, Grade Three

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Vocabulary	14	–0.66	0.48	–1.51	0.32
Reading for Understanding	17	–0.28	0.61	–1.06	1.46
Language	17	–0.17	0.46	–1.16	0.44
All operational items	48	–0.35	0.55	–1.51	1.46
Field-test items	54	0.01	0.66	–1.54	1.23

Table 8.D.8 IRT *b*-values for ELA, Grade Four

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Vocabulary	11	–0.55	0.38	–1.23	–0.08
Reading for Understanding	16	0.11	0.50	–0.63	1.07
Language	21	–0.07	0.49	–1.01	0.80
Writing Applications	1	–0.42	0.00	–0.42	–0.42
All operational items	49	–0.12	0.53	–1.23	1.07
Field-test items	42	0.35	0.43	–0.83	1.03

Table 8.D.9 IRT *b*-values for ELA, Grade Five

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Vocabulary	8	–0.86	0.81	–2.06	0.15
Reading for Understanding	18	0.00	0.49	–0.87	1.04
Language	22	–0.66	0.80	–2.23	0.56
All operational items	48	–0.45	0.77	–2.23	1.04
Field-test items	51	0.09	0.65	–1.87	1.09

Table 8.D.10 IRT *b*-values for Mathematics, Grade Three

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Number Sense	24	–0.29	0.75	–1.63	2.11
Algebra and Data Analysis	13	–0.54	0.81	–1.37	1.41
Measurement and Geometry	11	–0.77	0.59	–1.44	0.62
All operational items	48	–0.47	0.75	–1.63	2.11
Field-test items	54	0.29	0.78	–1.63	2.09

Table 8.D.11 IRT *b*-values for Mathematics, Grade Four

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Number Sense	23	–0.54	0.84	–2.54	0.60
Algebra and Data Analysis	15	0.08	0.43	–0.56	0.87
Measurement and Geometry	10	–0.12	0.78	–1.90	0.76
All operational items	48	–0.26	0.76	–2.54	0.87
Field-test items	45	0.45	0.72	–1.35	1.99

Table 8.D.12 IRT *b*-values for Mathematics, Grade Five

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Number Sense	21	–0.64	0.92	–3.03	0.48
Algebra and Data Analysis	17	–0.32	0.62	–1.28	1.07
Measurement and Geometry	10	0.01	0.54	–0.73	0.74
All operational items	48	–0.39	0.78	–3.03	1.07
Field-test items	54	0.47	0.71	–1.17	2.20

Table 8.D.13 IRT *b*-values for Science, Grade Five

Content Area	Number of items	Mean	Standard Deviation	Minimum	Maximum
Physical Sciences	16	–0.66	0.91	–2.27	0.54
Life Sciences	16	–0.62	0.68	–1.85	0.34
Earth Sciences	16	–0.33	0.59	–1.29	0.68
All operational items	48	–0.54	0.74	–2.27	0.68
Field-test items	54	0.19	0.66	–1.97	1.19

Table 8.D.14 Distribution of IRT *b*-values for ELA

IRT *b*-value	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
>=3.5	0	0	0	0	0	0
3.0 < 3.5	0	0	0	0	0	0
2.5 < 3.0	0	0	0	0	0	0
2.0 < 2.5	0	0	0	0	0	0
1.5 < 2.0	0	0	0	1	0	0
1.0 < 1.5	1	1	1	1	0	0
0.5 < 1.0	0	6	3	7	3	8
0.0 < 0.5	14	15	10	13	19	16
–0.5 < 0.0	13	14	16	19	19	16
–1.0 < –0.5	15	11	6	9	10	11
–1.5 < –1.0	4	2	7	3	4	2
–2.0 < –1.5	1	0	2	1	0	0
–2.5 < –2.0	0	0	3	0	0	1
–3.0 < –2.5	0	0	0	0	0	0
–3.5 < –3.0	0	0	0	0	0	0
< –3.5	0	0	0	0	0	0
Total	48	49	48	54	55	54

Table 8.D.15 Distribution of IRT *b*-values for ELA (field-test items)

IRT *b*-value	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
>=3.5	0	0	0	0	0	0
3.0 < 3.5	0	0	0	0	0	0
2.5 < 3.0	0	0	0	0	0	0
2.0 < 2.5	0	0	0	0	0	0
1.5 < 2.0	0	0	0	0	1	0
1.0 < 1.5	2	1	3	1	2	1
0.5 < 1.0	14	17	11	6	12	5
0.0 < 0.5	14	17	17	16	9	9
–0.5 < 0.0	13	5	11	10	8	11
–1.0 < –0.5	7	2	6	3	3	6
–1.5 < –1.0	2	0	2	0	1	2
–2.0 < –1.5	2	0	1	0	0	1
–2.5 < –2.0	0	0	0	0	0	0
–3.0 < –2.5	0	0	0	0	0	0
–3.5 < –3.0	0	0	0	0	0	0
< –3.5	0	0	0	0	0	0
Total	54	42	51	36	36	35

Table 8.D.16 Distribution of IRT *b*-values for Mathematics

IRT *b*-value	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
>=3.5	0	0	0	0	0
3.0 < 3.5	0	0	0	0	0
2.5 < 3.0	0	0	0	0	0
2.0 < 2.5	1	0	0	0	0
1.5 < 2.0	0	0	0	0	0
1.0 < 1.5	1	0	1	0	2
0.5 < 1.0	3	8	2	4	8
0.0 < 0.5	4	13	13	20	24
–0.5 < 0.0	12	13	12	19	16
–1.0 < –0.5	15	6	12	7	4
–1.5 < –1.0	11	5	5	3	0
–2.0 < –1.5	1	2	1	1	0
–2.5 < –2.0	0	0	0	0	0
–3.0 < –2.5	0	1	1	0	0
–3.5 < –3.0	0	0	1	0	0
< –3.5	0	0	0	0	0
Total	48	48	48	54	54

Table 8.D.17 Distribution of IRT *b*-values for Mathematics (field–test items)

IRT *b*-value	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
>=3.5	0	0	0	0	0
3.0 < 3.5	0	0	0	0	0
2.5 < 3.0	0	0	0	0	0
2.0 < 2.5	1	0	1	0	0
1.5 < 2.0	2	1	0	1	1
1.0 < 1.5	8	9	10	4	3
0.5 < 1.0	9	17	20	7	16
0.0 < 0.5	13	10	8	14	12
–0.5 < 0.0	11	3	9	10	4
–1.0 < –0.5	8	2	4	0	0
–1.5 < –1.0	1	3	2	0	0
–2.0 < –1.5	1	0	0	0	0
–2.5 < –2.0	0	0	0	0	0
–3.0 < –2.5	0	0	0	0	0
–3.5 < –3.0	0	0	0	0	0
< –3.5	0	0	0	0	0
Total	54	45	54	36	36

Table 8.D.18 Distribution of IRT *b*-values for Science

IRT *b*-value	Grade 5	Grade 8
>=3.5	0	0
3.0 < 3.5	0	0
2.5 < 3.0	0	0
2.0 < 2.5	0	0
1.5 < 2.0	0	0
1.0 < 1.5	0	0
0.5 < 1.0	2	6
0.0 < 0.5	11	21
–0.5 < 0.0	11	17
–1.0 < –0.5	11	7
–1.5 < –1.0	7	2
–2.0 < –1.5	5	1
–2.5 < –2.0	1	0
–3.0 < –2.5	0	0
–3.5 < –3.0	0	0
< –3.5	0	0
Total	48	54

Table 8.D.19 Distribution of IRT *b*-values for Science (field-test items)

IRT *b*-value	Grade 5	Grade 8
>=3.5	0	0
3.0 < 3.5	0	0
2.5 < 3.0	0	0
2.0 < 2.5	0	0
1.5 < 2.0	0	1
1.0 < 1.5	4	2
0.5 < 1.0	13	15
0.0 < 0.5	23	21
–0.5 < 0.0	7	11
–1.0 < –0.5	3	3
–1.5 < –1.0	3	1
–2.0 < –1.5	1	0
–2.5 < –2.0	0	0
–3.0 < –2.5	0	0
–3.5 < –3.0	0	0
< –3.5	0	0
Total	54	54

Table 8.D.20 2009 New Conversions for ELA, Grade Three—Regular Standard Form

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	–40	150		25	–0.2693	281	
1	–4.3452	150		26	–0.1794	287	**Below Basic**
2	–3.6249	150		27	–0.0889	293	
3	–3.1917	150		28	0.0026	300	
4	–2.8756	150		29	0.0954	307	
5	–2.6236	150		30	0.1899	313	
6	–2.4118	150		31	0.2864	320	**Basic**
7	–2.2277	150		32	0.3854	327	
8	–2.0635	152	**Far Below Basic**	33	0.4875	335	
9	–1.9145	163		34	0.5933	342	
10	–1.7773	173		35	0.7034	350	
11	–1.6495	182		36	0.8189	359	
12	–1.5292	190		37	0.9407	367	**Proficient**
13	–1.4151	199		38	1.0703	377	
14	–1.3061	206		39	1.2094	386	
15	–1.2013	214		40	1.3607	397	
16	–1.1001	221		41	1.5273	409	
17	–1.0018	228		42	1.7143	423	
18	–0.9059	235		43	1.9293	438	
19	–0.8120	242		44	2.1853	456	**Advanced**
20	–0.7196	248		45	2.5057	479	
21	–0.6284	255	**Below Basic**	46	2.9441	511	
22	–0.5381	261		47	3.6702	563	
23	–0.4484	268		48	20	600	
24	–0.3589	274					

Table 8.D.21 2009 New Conversions for ELA, Grade 3—Braille Version

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	-40	150	Far Below Basic	25	–0.22	284	Below Basic
1	–4.3253	150		26	–0.1278	291	
2	–3.6043	150		27	–0.0347	297	
3	–3.1703	150		28	0.0597	304	Basic
4	–2.8535	150		29	0.1556	311	
5	–2.6006	150		30	0.2534	318	
6	–2.3881	150		31	0.3538	325	
7	–2.2031	150		32	0.4571	333	
8	–2.0381	154		33	0.5641	340	
9	–1.8881	165		34	0.6753	348	
10	–1.75	175		35	0.7918	357	Proficient
11	–1.6211	184		36	0.9147	365	
12	–1.4998	192		37	1.0452	375	
13	–1.3846	201		38	1.1853	385	
14	–1.2745	209		39	1.3375	396	
15	–1.1686	216		40	1.505	408	Advanced
16	–1.0661	224		41	1.693	421	
17	–0.9665	231		42	1.9088	437	
18	–0.8692	238	Below Basic	43	2.1656	455	
19	–0.7738	244		44	2.4868	478	
20	–0.6799	251		45	2.9261	509	
21	–0.587	258		46	3.653	561	
22	–0.495	264		47	20	600	
23	–0.4033	271					
24	–0.3118	278					

Table 8.D.22 2009 New Conversions for ELA, Grade Four—Multiple Choice Only

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	–40	150		25	–0.0290	308	
1	–4.0873	150		26	0.0599	316	
2	–3.3677	150		27	0.1494	325	**Basic**
3	–2.9352	150		28	0.2398	333	
4	–2.6199	150		29	0.3315	341	
5	–2.3685	150		30	0.4247	350	
6	–2.1574	150		31	0.5200	359	
7	–1.9740	150		32	0.6177	368	**Proficient**
8	–1.8105	150	**Far Below Basic**	33	0.7183	377	
9	–1.6622	157		34	0.8225	387	
10	–1.5256	170		35	0.9310	397	
11	–1.3984	182		36	1.0446	407	
12	–1.2788	193		37	1.1644	418	
13	–1.1653	203		38	1.2918	430	
14	–1.0570	213		39	1.4286	443	
15	–0.9530	223		40	1.5771	456	
16	–0.8525	232		41	1.7409	472	
17	–0.7549	241		42	1.9246	489	**Advanced**
18	–0.6597	250		43	2.1360	508	
19	–0.5665	259		44	2.3877	531	
20	–0.4750	267	**Below Basic**	45	2.7033	560	
21	–0.3846	275		46	3.1363	600	
22	–0.2951	284		47	3.8563	600	
23	–0.2062	292		48	20	600	
24	–0.1176	300	**Basic**				

Table 8.D.23 2009 New Conversions for ELA, Grade Four—Composite Version (multiple choice and essay)

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	–40	150	Far Below Basic	27	–0.0533	306	Basic
1	–4.2667	150		28	0.0314	314	
2	–3.5486	150		29	0.1165	322	
3	–3.1166	150		30	0.2023	330	
4	–2.8011	150		31	0.289	338	
5	–2.5491	150		32	0.3771	346	
6	–2.3371	150		33	0.4666	354	Proficient
7	–2.1526	150		34	0.5581	362	
8	–1.988	150		35	0.6519	371	
9	–1.8386	150		36	0.7483	380	
10	–1.7011	154		37	0.848	389	
11	–1.573	166		38	0.9515	399	
12	–1.4528	177		39	1.0596	409	Advanced
13	–1.339	187		40	1.173	419	
14	–1.2305	197		41	1.293	430	
15	–1.1265	207		42	1.4208	442	
16	–1.0264	216		43	1.5583	455	
17	–0.9295	225		44	1.708	469	
18	–0.8353	234		45	1.8731	484	
19	–0.7435	242		46	2.0588	501	
20	–0.6535	251	Below Basic	47	2.2725	521	
21	–0.5652	259		48	2.527	544	
22	–0.4782	267		49	2.846	574	
23	–0.3922	275		50	3.2832	600	
24	–0.307	283		51	4.0081	600	
25	–0.222	290		52	20	600	
26	–0.1378	298					

Table 8.D.24 2009 New Conversions for ELA, Grade Five

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	-40	150	Far Below Basic	25	–0.3354	285	Below Basic
1	–4.5785	150		26	–0.2415	293	
2	–3.8478	150		27	–0.1471	301	Basic
3	–3.4045	150		28	–0.0520	309	
4	–3.0787	150		29	0.0443	317	
5	–2.8173	150		30	0.1420	325	
6	–2.5965	150		31	0.2415	334	
7	–2.4037	150		32	0.3434	343	
8	–2.2312	150		33	0.4482	351	Proficient
9	–2.0742	150		34	0.5564	361	
10	–1.9292	150		35	0.6688	370	
11	–1.7939	161		36	0.7862	380	
12	–1.6665	172		37	0.9097	390	
13	–1.5455	182		38	1.0408	402	Advanced
14	–1.4299	192		39	1.1811	413	
15	–1.3187	202		40	1.3332	426	
16	–1.2114	211		41	1.5003	441	
17	–1.1072	220	Below Basic	42	1.6875	456	
18	–1.0056	228		43	1.9021	475	
19	–0.9062	237		44	2.1571	496	
20	–0.8087	245		45	2.4760	523	
21	–0.7125	253		46	2.9121	560	
22	–0.6174	261		47	3.6350	600	
23	–0.5230	269		48	20	600	
24	–0.4291	277					

Table 8.D.25 2009 New Conversions for Mathematics, Grade Three

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	-40	150	Far Below Basic	25	–0.4023	281	Below Basic
1	–4.5241	150		26	–0.3087	288	
2	–3.8031	150		27	–0.2142	294	
3	–3.3691	150		28	–0.1186	301	Basic
4	–3.0522	150		29	–0.0214	308	
5	–2.7992	150		30	0.0778	314	
6	–2.5864	150		31	0.1792	321	
7	–2.4011	150		32	0.2836	328	
8	–2.2358	156		33	0.3913	336	
9	–2.0854	166		34	0.5030	344	
10	–1.9468	175		35	0.6196	352	Proficient
11	–1.8174	184		36	0.7419	360	
12	–1.6954	193		37	0.8711	369	
13	–1.5795	201		38	1.0087	378	
14	–1.4687	208		39	1.1565	388	
15	–1.3619	216		40	1.3172	399	
16	–1.2586	223		41	1.4941	412	
17	–1.1581	230	Below Basic	42	1.6925	425	Advanced
18	–1.0598	236		43	1.9202	441	
19	–0.9633	243		44	2.1902	459	
20	–0.8683	249		45	2.5266	482	
21	–0.7743	256		46	2.9831	514	
22	–0.6811	262		47	3.7302	565	
23	–0.5882	269		48	20	600	
24	–0.4953	275					

Table 8.D.26 2009 New Conversions for Mathematics, Grade Four

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	-40	150		25	–0.1387	301	
1	–4.4537	150		26	–0.0448	310	
2	–3.7130	150		27	0.0494	320	**Basic**
3	–3.2611	150		28	0.1443	330	
4	–2.9280	150		29	0.2402	341	
5	–2.6602	150		30	0.3375	351	
6	–2.4338	150		31	0.4367	361	
7	–2.2361	150		32	0.5380	372	
8	–2.0594	150	**Far Below Basic**	33	0.6422	383	**Proficient**
9	–1.8986	150		34	0.7497	394	
10	–1.7504	150		35	0.8613	406	
11	–1.6122	150		36	0.9779	418	
12	–1.4822	159		37	1.1006	431	
13	–1.3590	172		38	1.2307	445	
14	–1.2415	184		39	1.3700	460	
15	–1.1287	196		40	1.5210	476	
16	–1.0199	208		41	1.6870	493	
17	–0.9145	219		42	1.8728	513	**Advanced**
18	–0.8120	229		43	2.0862	535	
19	–0.7117	240		44	2.3397	562	
20	–0.6135	250		45	2.6571	596	
21	–0.5167	261	**Below Basic**	46	3.0917	600	
22	–0.4212	271		47	3.8127	600	
23	–0.3266	281		48	20	600	
24	–0.2325	291					

Table 8.D.27 2009 New Conversions for Mathematics, Grade Five

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	–40	150		25	–0.2717	292	Below Basic
1	–4.6259	150		26	–0.1786	300	
2	–3.8700	150		27	–0.0851	308	
3	–3.4065	150		28	0.0092	316	**Basic**
4	–3.0647	150		29	0.1046	325	
5	–2.7906	150		30	0.2015	333	
6	–2.5597	150		31	0.3003	342	
7	–2.3590	150		32	0.4014	350	
8	–2.1803	150	**Far Below Basic**	33	0.5054	359	
9	–2.0185	150		34	0.6129	369	
10	–1.8698	154		35	0.7245	378	**Proficient**
11	–1.7317	166		36	0.8412	388	
12	–1.6022	177		37	0.9641	399	
13	–1.4797	187		38	1.0945	410	
14	–1.3630	197		39	1.2343	422	
15	–1.2513	207		40	1.3858	436	
16	–1.1436	216		41	1.5525	450	
17	–1.0393	226		42	1.7392	466	
18	–0.9379	234		43	1.9534	485	**Advanced**
19	–0.8388	243		44	2.2081	507	
20	–0.7417	251		45	2.5267	534	
21	–0.6460	260	**Below Basic**	46	2.9626	572	
22	–0.5516	268		47	3.6854	600	
23	–0.4579	276		48	20	600	
24	–0.3647	284					

Table 8.D.28 2009 New Conversions for Science, Grade Five

Raw score	Theta	Scale Score	Performance Level	Raw score	Theta	Scale Score	Performance Level
0	–40	150		25	–0.4325	300	
1	–4.6709	150		26	–0.3380	307	
2	–3.9416	150		27	–0.2430	314	
3	–3.4995	150		28	–0.1473	321	**Basic**
4	–3.1749	150		29	–0.0504	329	
5	–2.9145	150		30	0.0479	336	
6	–2.6946	150		31	0.1482	343	
7	–2.5025	150		32	0.2507	351	
8	–2.3307	159	**Far Below Basic**	33	0.3560	359	
9	–2.1742	170		34	0.4649	367	**Proficient**
10	–2.0296	181		35	0.5778	375	
11	–1.8946	191		36	0.6958	384	
12	–1.7673	201		37	0.8199	393	
13	–1.6465	210		38	0.9515	403	
14	–1.5309	218		39	1.0923	414	
15	–1.4197	227		40	1.2449	425	
16	–1.3122	235		41	1.4124	437	
17	–1.2078	242		42	1.5999	451	
18	–1.1059	250		43	1.8149	467	**Advanced**
19	–1.0062	257		44	2.0701	486	
20	–0.9083	265		45	2.3891	510	
21	–0.8117	272	**Below Basic**	46	2.8252	543	
22	–0.7161	279		47	3.5484	596	
23	–0.6212	286		48	20	600	
24	–0.5268	293					

Appendix 8.E—DIF Analyses

Table 8.E.1 Operational Items Exhibiting Significant DIF

Content Area	Accession No.	Grade	Item Seq#	MHD-DIF	Comparison	In Favor Of
English–Language Arts	VC326880	3	3	1.91	SLD/AUT	AUT
	VC326734	3	9	2.17	SLD/AUT	AUT
	VC326880	3	3	1.72	SLD/MR	MR
	VC326726	3	12	1.72	SLD/MR	MR
	VC325577	4	39	3.03	WHITE/FILIPINO	FILIPINO
	VC325577	4	39	1.9	WHITE/COMBASIAN	COMBASIAN
	VC326607	4	3	–1.65	SLD/AUT	SLD
	VC744644	4	58	SMD: –0.426 SE SMD: 0.040	SLD/AUT	SLD
	VC744644	4	58	SMD: –0.441 SE SMD: 0.056	SLD/ED	SLD
	VC744644	4	58	SMD: –0.507 SE SMD: 0.074	SLD/MR	SLD
	VC325593	5	57	2	WHITE / FILIPINO	FILIPINO
	VC325529	5	34	1.53	SLD / AUT	AUT
	VC743453	6	2	–1.62	ENGSPKR /ENGLRN	ENGSPKR
	VC743441	6	1	–2.59	SLD/AUT	SLD
	VC743443	6	4	–2.67	SLD/AUT	SLD
	VC743514	6	40	2.22	SLD/AUT	AUT
	VC743498	6	48	2.01	SLD/AUT	AUT
	VC743471	6	59	1.6	SLD/AUT	AUT
	VC743635	7	50	2.05	WHITE/ASIAN	ASIAN
	VC744649	7	64	SMD: –0.325 SE SMD: 0.053	SLD/ED	SLD
	VC744649	7	64	SMD: –0.429 SE SMD: 0.061	SLD/MR	SLD
	VC743760	7	9	–1.61	SLD/AUT	SLD
	VC743764	7	13	–1.71	SLD/AUT	SLD
	VC743648	7	52	1.57	SLD/AUT	AUT
	VC744649	7	64	SMD: –0.531 SE SMD: 0.047	SLD/AUT	SLD
	VC743916	8	1	–2.16	WHITE/ASIAN	WHITE
	VC743837	8	61	1.52	SLD/SLI	SLI
	VC743883	8	3	–2.56	SLD/MR	SLD
	VC743883	8	3	–2.17	SLD/AUT	SLD
	VC743916	8	11	–1.96	SLD/AUT	SLD
	VC743816	8	62	2.01	SLD/AUT	AUT
Mathematics	VC325953	3	15	2.3	SLD/MR	MR
	VC744107	6	3	–1.68	SLD /ED	SLD
	VC744397	7	9	–1.83	WHITE/ASIAN	WHITE
	VC744425	7	50	–2	WHITE/ASIAN	WHITE
Science	VC744523	8	22	–2.06	WHITE/ASIAN	WHITE
	VC744536	8	12	–1.85	SLD/AUT	SLD

Table 8.E.2 Field Test Items Exhibiting Significant DIF

Content Area	Accession No.	Grade	Item Seq#	MHD-DIF	Comparison	In Favor Of
English–Language-Arts *	–	–	–	–	–	–
Mathematics *	–	–	–	–	–	–
Science *	–	–	–	–	–	–

* No Field Test items exhibited significant DIF

Table 8.E.3 DIF Classifications for ELA, Grade Three Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	3	6
B+	0	0	0	0	1	2	2	4	3	6	0	0	0	0	3	6	0	0	2	4	0	0	2	4	0	0	1	2	0	0	0	0	1	2	0	0	0	0	0	0	6	13
A+	28	58	24	50	22	46	20	42	19	40	24	50	0	0	18	38	27	56	22	46	0	0	26	54	0	0	19	40	0	0	0	0	22	46	22	46	0	0	0	0	13	27
A-	20	42	24	50	24	50	25	52	24	50	24	50	0	0	26	54	21	44	20	42	0	0	17	35	0	0	24	50	0	0	0	0	24	50	26	54	0	0	0	0	18	38
B-	0	0	0	0	1	2	1	2	2	4	0	0	0	0	1	2	0	0	2	4	0	0	3	6	0	0	2	4	0	0	0	0	1	2	0	0	0	0	0	0	8	17
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.4 DIF Classifications for ELA, Grade Four Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	1	2	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
B+	1	2	0	0	1	2	2	4	3	6	0	0	0	0	1	2	0	0	3	6	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	7	14
A+	26	53	24	49	21	43	17	35	24	49	29	59	0	0	22	45	21	43	23	47	0	0	25	51	0	0	23	47	0	0	0	0	25	51	24	49	0	0	0	0	14	29
A-	22	45	25	51	27	55	29	59	19	39	20	41	0	0	24	49	28	57	20	41	0	0	21	43	0	0	19	39	0	0	0	0	24	49	25	51	0	0	0	0	18	37
B-	0	0	0	0	0	0	1	2	2	4	0	0	0	0	1	2	0	0	1	2	0	0	2	4	0	0	4	8	0	0	0	0	0	0	0	0	0	0	0	0	7	14
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	1	2	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	2	4
Small N	0	0	0	0	0	0	0	0	0	0	0	0	49	100	0	0	0	0	0	0	49	100	0	0	49	100	0	0	49	100	49	100	0	0	0	0	49	100	49	100	0	0
TOTAL	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100	49	100

Table 8.E.5 DIF Classifications for ELA, Grade Five Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4
B+	1	2	0	0	0	0	3	6	5	10	0	0	0	0	2	4	0	0	2	4	0	0	1	2	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	6	13
A+	22	46	28	58	26	54	24	50	15	31	25	52	0	0	23	48	25	52	20	42	0	0	23	48	0	0	23	48	0	0	0	0	23	48	26	54	0	0	0	0	12	25
A-	25	52	19	40	21	44	19	40	25	52	23	48	0	0	23	48	22	46	25	52	0	0	23	48	0	0	23	48	0	0	0	0	25	52	22	46	0	0	0	0	20	42
B-	0	0	1	2	1	2	2	4	2	4	0	0	0	0	0	0	1	2	0	0	0	0	1	2	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	8	17
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.6 DIF Classifications for ELA, Grade Six Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6
B+	0	0	0	0	1	2	3	6	3	6	0	0	0	0	2	4	0	0	2	4	0	0	0	0	0	0	1	2	0	0	0	0	0	0	3	6	0	0	0	0	6	11
A+	27	50	29	54	26	48	22	41	23	43	30	56	0	0	25	46	28	52	28	52	0	0	29	54	0	0	25	46	0	0	0	0	27	50	21	39	0	0	0	0	16	30
A-	27	50	25	46	27	50	28	52	26	48	24	44	0	0	27	50	24	44	15	28	0	0	24	44	0	0	26	48	0	0	0	0	27	50	29	54	0	0	0	0	19	35
B-	0	0	0	0	0	0	1	2	2	4	0	0	0	0	0	0	1	2	4	7	0	0	1	2	0	0	2	4	0	0	0	0	0	0	1	2	0	0	0	0	7	13
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6
Small N	0	0	0	0	0	0	0	0	0	0	0	0	54	100	0	0	0	0	0	0	54	100	0	0	54	100	0	0	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.7 DIF Classifications for ELA, Grade Seven Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4
B+	1	2	0	0	1	2	3	5	0	0	0	0	0	0	1	2	0	0	3	5	0	0	1	2	0	0	1	2	0	0	0	0	0	0	2	4	0	0	0	0	8	15
A+	27	49	25	45	23	42	21	38	0	0	32	58	0	0	28	51	30	55	26	47	0	0	25	45	0	0	30	55	0	0	0	0	30	55	24	44	0	0	0	0	15	27
A-	26	47	30	55	31	56	27	49	0	0	23	42	0	0	24	44	24	44	18	33	0	0	27	49	0	0	20	36	0	0	0	0	25	45	28	51	0	0	0	0	17	31
B-	1	2	0	0	0	0	2	4	0	0	0	0	0	0	2	4	1	2	4	7	0	0	1	2	0	0	3	5	0	0	0	0	0	0	1	2	0	0	0	0	10	18
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	5	0	0	1	2	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	3	5
Small N	0	0	0	0	0	0	1	2	55	100	0	0	55	100	0	0	0	0	0	0	55	100	0	0	55	100	0	0	55	100	55	100	0	0	0	0	55	100	55	100	0	0
TOTAL	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100	55	100

Table 8.E.8 DIF Classifications for ELA, Grade Eight Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	2	4
B+	0	0	0	0	0	0	3	6	0	0	0	0	0	0	1	2	0	0	3	6	0	0	1	2	0	0	1	2	0	0	0	0	0	0	1	2	0	0	0	0	6	11
A+	27	50	29	54	30	56	27	50	0	0	27	50	0	0	30	56	29	54	24	44	0	0	22	41	0	0	26	48	0	0	0	0	27	50	26	48	0	0	0	0	22	41
A-	27	50	25	46	24	44	19	35	0	0	27	50	0	0	20	37	25	46	21	39	0	0	30	56	0	0	22	41	0	0	0	0	27	50	24	44	0	0	0	0	13	24
B-	0	0	0	0	0	0	4	7	0	0	0	0	0	0	3	6	0	0	3	6	0	0	1	2	0	0	3	6	0	0	0	0	0	0	2	4	0	0	0	0	8	15
C-	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	3	6
Small N	0	0	0	0	0	0	0	0	54	100	0	0	54	100	0	0	0	0	0	0	54	100	0	0	54	100	0	0	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.9 DIF Classifications for Mathematics, Grade Three Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	2
B+	0	0	0	0	0	0	1	2	4	8	0	0	0	0	2	4	0	0	2	4	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	19
A+	25	52	22	46	27	56	28	58	19	40	25	52	0	0	25	52	25	52	22	46	0	0	24	50	0	0	25	52	0	0	0	0	29	60	25	52	0	0	0	0	14	29
A-	23	48	26	54	20	42	17	35	23	48	23	48	0	0	21	44	23	48	23	48	0	0	21	44	0	0	20	42	0	0	0	0	18	38	23	48	0	0	0	0	17	35
B-	0	0	0	0	1	2	2	4	2	4	0	0	0	0	0	0	0	0	1	2	0	0	1	2	0	0	2	4	0	0	0	0	1	2	0	0	0	0	0	0	7	15
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.10 DIF Classifications for Mathematics, Grade Four Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	1	2	0	0	0	0	4	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	10	0	0	0	0	0	0	0	0	0	0	0	0	10	21
A+	27	56	22	46	25	52	18	38	25	52	23	48	0	0	24	50	29	60	25	52	0	0	22	46	0	0	15	31	0	0	0	0	24	50	22	46	0	0	0	0	19	40
A-	20	42	26	54	22	46	23	48	21	44	25	52	0	0	22	46	19	40	23	48	0	0	25	52	0	0	25	52	0	0	0	0	24	50	26	54	0	0	0	0	12	25
B-	0	0	0	0	1	2	3	6	2	4	0	0	0	0	2	4	0	0	0	0	0	0	1	2	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	7	15
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.11 DIF Classifications for Mathematics, Grade Five Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	2	4	2	4	0	0	0	0	1	2	0	0	0	0	0	0	1	2	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	4	8
A+	27	56	22	46	25	52	23	48	24	50	23	48	0	0	24	50	24	50	28	58	0	0	26	54	0	0	23	48	0	0	0	0	25	52	27	56	0	0	0	0	17	35
A-	21	44	26	54	23	48	22	46	20	42	25	52	0	0	23	48	24	50	18	38	0	0	21	44	0	0	21	44	0	0	0	0	23	48	21	44	0	0	0	0	20	42
B-	0	0	0	0	0	0	1	2	2	4	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	7	15
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.12 DIF Classifications for Mathematics, Grade Six Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	1	2	0	0	0	0	5	9	3	6	0	0	0	0	2	4	0	0	0	0	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	17
A+	27	50	31	57	25	46	22	41	24	44	29	54	0	0	29	54	28	52	31	57	0	0	22	41	0	0	28	52	0	0	0	0	24	44	31	57	0	0	0	0	21	39
A-	26	48	23	43	29	54	23	43	24	44	25	46	0	0	21	39	26	48	22	41	0	0	28	52	0	0	26	48	0	0	0	0	30	56	23	43	0	0	0	0	18	33
B-	0	0	0	0	0	0	4	7	3	6	0	0	0	0	2	4	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	9
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
Small N	0	0	0	0	0	0	0	0	0	0	0	0	54	100	0	0	0	0	0	0	54	100	0	0	54	100	0	0	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.13 DIF Classifications for Mathematics, Grade Seven Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	6	11	1	2	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	7	13
A+	27	50	27	50	26	48	19	35	31	57	23	43	0	0	27	50	24	44	26	48	0	0	27	50	0	0	22	41	0	0	0	0	28	52	30	56	0	0	0	0	22	41
A-	27	50	27	50	28	52	24	44	19	35	31	57	0	0	23	43	30	56	28	52	0	0	26	48	0	0	26	48	0	0	0	0	26	48	24	44	0	0	0	0	16	30
B-	0	0	0	0	0	0	3	6	3	6	0	0	0	0	3	6	0	0	0	0	0	0	1	2	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	7	13
C-	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4
Small N	0	0	0	0	0	0	0	0	0	0	0	0	54	100	0	0	0	0	0	0	54	100	0	0	54	100	0	0	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.14 DIF Classifications for Science, Grade Five Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	1	2	1	2	4	8	0	0	0	0	1	2	0	0	4	8	0	0	1	2	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	8	17
A+	27	56	25	52	22	46	25	52	24	50	26	54	0	0	25	52	27	56	19	40	0	0	26	54	0	0	26	54	0	0	0	0	22	46	26	54	0	0	0	0	16	33
A-	21	44	23	48	25	52	20	42	18	38	22	46	0	0	22	46	21	44	22	46	0	0	20	42	0	0	19	40	0	0	0	0	26	54	22	46	0	0	0	0	16	33
B-	0	0	0	0	0	0	2	4	2	4	0	0	0	0	0	0	0	0	3	6	0	0	1	2	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	8	17
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	48	100	0	0	0	0	0	0	48	100	0	0	48	100	0	0	48	100	48	100	0	0	0	0	48	100	48	100	0	0
TOTAL	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100	48	100

Table 8.E.15 DIF Classifications for Science, Grade Eight Operational Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	2	4	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	4	7
A+	30	56	24	44	27	50	27	50	0	0	28	52	0	0	28	52	28	52	28	52	0	0	32	59	0	0	25	46	0	0	0	0	26	48	30	56	0	0	0	0	21	39
A-	24	44	30	56	27	50	21	39	0	0	26	48	0	0	22	41	26	48	24	44	0	0	20	37	0	0	24	44	0	0	0	0	28	52	24	44	0	0	0	0	21	39
B-	0	0	0	0	0	0	3	6	0	0	0	0	0	0	3	6	0	0	1	2	0	0	2	4	0	0	3	6	0	0	0	0	0	0	0	0	0	0	0	0	6	11
C-	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4
Small N	0	0	0	0	0	0	0	0	54	100	0	0	54	100	0	0	0	0	0	0	54	100	0	0	54	100	0	0	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.16 DIF Classifications for ELA, Grade Three Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	2	4
A+	21	39	21	39	0	0	0	0	0	0	27	50	0	0	0	0	25	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	61	0	0	0	0	20	37
A-	32	59	31	57	0	0	0	0	0	0	24	44	0	0	0	0	29	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	33	0	0	0	0	25	46
B-	1	2	2	4	0	0	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	7	13
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	54	100	54	100	54	100	0	0	54	100	54	100	0	0	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.17 DIF Classifications for ELA, Grade Four Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	2	5	1	2	0	0	0	0	4	10
A+	22	52	13	31	0	0	0	0	0	0	19	45	0	0	3	7	21	50	17	40	0	0	2	5	0	0	0	0	0	0	0	0	23	55	23	55	0	0	0	0	20	48
A-	19	45	29	69	0	0	0	0	0	0	23	55	0	0	0	0	21	50	23	55	0	0	1	2	0	0	0	0	0	0	0	0	17	40	18	43	0	0	0	0	16	38
B-	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	42	100	42	100	42	100	0	0	42	100	39	93	0	0	0	0	42	100	39	93	42	100	42	100	42	100	42	100	0	0	0	0	42	100	42	100	0	0
TOTAL	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100	42	100

Table 8.E.18 DIF Classifications for ELA, Grade Five Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	2	4	1	2	0	0	0	0	4	8
A+	19	37	21	41	0	0	0	0	0	0	20	39	0	0	1	2	16	31	0	0	0	0	2	4	0	0	0	0	0	0	0	0	27	53	20	39	0	0	0	0	16	31
A-	32	63	27	53	0	0	0	0	0	0	30	59	0	0	2	4	34	67	2	4	0	0	0	0	0	0	0	0	0	0	0	0	22	43	30	59	0	0	0	0	28	55
B-	0	0	3	6	0	0	0	0	0	0	1	2	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	51	100	51	100	51	100	0	0	51	100	48	94	0	0	48	94	51	100	49	96	51	100	51	100	51	100	51	100	0	0	0	0	51	100	51	100	0	0
TOTAL	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100	51	100

Table 8.E.19 DIF Classifications for ELA, Grade Six Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	1	3
A+	18	50	12	33	0	0	0	0	0	0	14	39	0	0	0	0	18	50	3	8	0	0	0	0	0	0	0	0	0	0	0	0	19	53	19	53	0	0	0	0	15	42
A-	18	50	24	67	0	0	0	0	0	0	22	61	0	0	0	0	18	50	6	17	0	0	0	0	0	0	0	0	0	0	0	0	17	47	16	44	0	0	0	0	20	56
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	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	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	36	100	36	100	36	100	0	0	36	100	36	100	0	0	27	75	36	100	36	100	36	100	36	100	36	100	36	100	0	0	0	0	36	100	36	100	0	0
TOTAL	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100

Table 8.E.20 DIF Classifications for ELA, Grade Seven Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	1	3	0	0	0	0	1	3
A+	22	61	14	39	0	0	0	0	0	0	17	47	0	0	0	0	14	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	39	18	50	0	0	0	0	12	33
A-	13	36	22	61	0	0	0	0	0	0	19	53	0	0	0	0	21	58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	58	16	44	0	0	0	0	21	58
B-	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	2	6
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	36	100	36	100	36	100	0	0	36	100	36	100	0	0	36	100	36	100	36	100	36	100	36	100	36	100	36	100	0	0	0	0	36	100	36	100	0	0
TOTAL	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100

Table 8.E.21 DIF Classifications for ELA, Grade Eight Field-test Items

DIF	M-F		W-AfrA		W-Aml		W-Asn		W-Fil		W-His		W-Pacl		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	2	6	0	0	0	0	0	0	1	3	0	0	0	0	0	0	1	3	0	0	0	0	0	0	1	3	0	0	0	0	4	11	1	3	0	0	0	0	8	23
A+	22	63	14	40	0	0	0	0	0	0	18	51	0	0	3	9	18	51	4	11	0	0	4	11	0	0	4	11	0	0	0	0	14	40	4	11	0	0	0	0	15	43
A-	13	37	19	54	0	0	0	0	0	0	16	46	0	0	4	11	17	49	3	9	0	0	6	17	0	0	3	9	0	0	0	0	17	49	5	14	0	0	0	0	11	31
B-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	0	0	1	3	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	1	3
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	35	100	35	100	35	100	0	0	35	100	26	74	0	0	26	74	35	100	25	71	35	100	26	74	35	100	35	100	0	0	25	71	35	100	35	100	0	0
TOTAL	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100	35	100

Table 8.E.22 DIF Classifications for Mathematics, Grade Three Field-test Items

DIF	M-F		W-AfrA		W-Aml		W-Asn		W-Fil		W-His		W-Pacl		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	1	2	3	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6
A+	21	39	22	41	0	0	0	0	0	0	25	46	0	0	0	0	21	39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	9	20	37	0	0	0	0	19	35
A-	31	57	27	50	0	0	0	0	0	0	28	52	0	0	0	0	32	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	7	29	54	0	0	0	0	23	43
B-	1	2	2	4	0	0	0	0	0	0	1	2	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	9	0	0	0	0	9	17
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	54	100	54	100	54	100	0	0	54	100	54	100	0	0	54	100	54	100	54	100	54	100	54	100	54	100	54	100	45	83	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.23 DIF Classifications for Mathematics, Grade Four Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	2	4
A+	22	49	20	44	0	0	0	0	0	0	16	36	0	0	0	0	18	40	9	20	0	0	0	0	0	0	0	0	0	0	0	0	19	42	16	36	0	0	0	0	16	36
A-	22	49	23	51	0	0	0	0	0	0	28	62	0	0	0	0	27	60	8	18	0	0	0	0	0	0	0	0	0	0	0	0	23	51	29	64	0	0	0	0	22	49
B-	1	2	2	4	0	0	0	0	0	0	1	2	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	5	11
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	45	100	45	100	45	100	0	0	45	100	45	100	0	0	27	60	45	100	45	100	45	100	45	100	45	100	45	100	0	0	0	0	45	100	45	100	0	0
TOTAL	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100	45	100

Table 8.E.24 DIF Classifications for Mathematics, Grade Five Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	1	2	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	4	7
A+	24	44	28	52	0	0	0	0	0	0	27	50	0	0	0	0	29	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29	54	27	50	0	0	0	0	19	35
A-	30	56	23	43	0	0	0	0	0	0	26	48	0	0	0	0	25	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	43	25	46	0	0	0	0	27	50
B-	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	4	7
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	54	100	54	100	54	100	0	0	54	100	54	100	0	0	54	100	54	100	54	100	54	100	54	100	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.25 DIF Classifications for Mathematics, Grade Six Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	2	6
A+	15	42	14	39	0	0	0	0	0	0	19	53	0	0	0	0	14	39	14	39	0	0	0	0	0	0	0	0	0	0	0	0	20	56	15	42	0	0	0	0	18	50
A-	18	50	19	53	0	0	0	0	0	0	17	47	0	0	0	0	21	58	12	33	0	0	0	0	0	0	0	0	0	0	0	0	15	42	21	58	0	0	0	0	11	31
B-	3	8	3	8	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	14
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	36	100	36	100	36	100	0	0	36	100	36	100	0	0	9	25	36	100	36	100	36	100	36	100	36	100	36	100	0	0	0	0	36	100	36	100	0	0
TOTAL	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100

Table 8.E.26 DIF Classifications for Mathematics, Grade Seven Field-test Items

DIF	M-F		W-AfrA		W-AmI		W-Asn		W-Fil		W-His		W-PacI		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	2	6
A+	16	44	16	44	0	0	0	0	0	0	18	50	0	0	0	0	18	50	0	0	0	0	2	6	0	0	0	0	0	0	0	0	14	39	14	39	0	0	0	0	12	33
A-	19	53	20	56	0	0	0	0	0	0	17	47	0	0	0	0	18	50	0	0	0	0	6	17	0	0	0	0	0	0	0	0	22	61	21	58	0	0	0	0	20	56
B-	1	3	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	36	100	36	100	36	100	0	0	36	100	36	100	0	0	36	100	36	100	27	75	36	100	36	100	36	100	36	100	0	0	0	0	36	100	36	100	0	0
TOTAL	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100	36	100

Table 8.E.27 DIF Classifications for Science, Grade Five Field-test Items

DIF	M-F		W-AfrA		W-Aml		W-Asn		W-Fil		W-His		W-Pacl		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	3	6
A+	25	46	22	41	0	0	0	0	0	0	16	30	0	0	0	0	23	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	29	54	30	56	0	0	0	0	22	41
A-	27	50	32	59	0	0	0	0	0	0	38	70	0	0	0	0	30	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	41	24	44	0	0	0	0	26	48
B-	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	3	6
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	54	100	54	100	54	100	0	0	54	100	54	100	0	0	54	100	54	100	54	100	54	100	54	100	54	100	54	100	0	0	0	0	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Table 8.E.28 DIF Classifications for Science, Grade Eight Field-test Items

DIF	M-F		W-AfrA		W-Aml		W-Asn		W-Fil		W-His		W-Pacl		W-ComA		E-ELnr		SLD-AUT		SLD-Deaf		SLD-ED		SLD-HH		SLD-MR		SLD-MD		SLD-OI		SLD-OHI		SLD-SLI		SLD-TBI		SLD-VI		Total	
category	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct	N	Pct
C+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0	0	2	4
A+	24	44	21	39	0	0	0	0	0	0	27	50	0	0	0	0	16	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	28	0	0	0	0	0	0	19	35
A-	28	52	32	59	0	0	0	0	0	0	26	48	0	0	0	0	37	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	28	52	0	0	0	0	0	0	30	56
B-	2	4	1	2	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6
C-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small N	0	0	0	0	54	100	54	100	54	100	0	0	54	100	54	100	0	0	54	100	54	100	54	100	54	100	54	100	54	100	54	100	9	17	54	100	54	100	54	100	0	0
TOTAL	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100	54	100

Chapter 9: Quality Control

ETS implements rigorous quality control procedures throughout the test development, administration, scoring, and reporting processes. As part of this effort, ETS maintains an Office of Testing Integrity (OTI) that resides in the ETS legal department. OTI provides quality assurance services for all testing programs administered by ETS. In addition, the Quality Assurance division of ETS publishes and maintains the *ETS Standards for Quality and Fairness,* which supports OTI's goals and activities. The purposes of the *ETS Standards for Quality and Fairness* are to help ETS design, develop, and deliver technically sound, fair, and useful products and services; and to help the public and auditors evaluate those products and services.

In addition, each department at ETS that is involved in the testing cycle designs and implements an independent set of procedures to ensure the quality of their products. In the next sections, these procedures are described.

Quality Control of Item Development

The item development process for the CMA is described in detail in Chapter 3, starting on page 60. The following sections highlight elements of the process devoted specifically to quality control of item development.

Item Specifications

ETS maintains item development specifications for each CMA and has developed an Item Utilization Plan to guide the development of the items for each content area. Item writing emphasis is determined in consultation with the CDE. Adherence to the specifications ensures the maintenance of quality and consistency of the item development process.

Item Writers

The items for each CMA are written by panels of item writers that have a thorough understanding of the California content standards. The item writers are carefully screened and selected by senior ETS content staff and approved by the CDE. Only those with strong content and teaching backgrounds are invited to participate in an extensive training program for item writers.

Internal Contractor Reviews

Once items have been written, ETS assessment specialists make sure that each item goes through an intensive internal review process. Every step of this process is designed to produce items that exceed industry standards for quality. It includes three rounds of content reviews, two rounds of editorial reviews, an internal fairness review, and a high-level review and approval by a content area director. A carefully designed and monitored workflow and detailed checklists help to ensure that all items meet the specifications for the process.

Content Review

ETS assessment specialists make sure that the test items and related materials comply with ETS's written guidelines for clarity, style, accuracy, and appropriateness and with approved item specifications.

The artwork and graphics for the items are created during the internal content review period so assessment specialists can evaluate the correctness and appropriateness of the art early in the item development process. ETS selects visuals that are relevant to the item content

and that are easily understood so students do not struggle to determine the purpose or meaning of the questions.

Editorial Review

Another step in the ETS internal review process involves a team of specially trained editors who check questions for clarity, correctness of language, grade-level appropriateness of language, adherence to style guidelines, and conformity to acceptable item-writing practices. The editorial review also includes rounds of copyediting and proofreading. ETS takes pride in the typographical integrity of the items presented to our clients and strives for error-free items beginning with the initial rounds of review.

Fairness Review

One of the final steps in the ETS internal review process is to have all items and stimuli reviewed for fairness. Only ETS staff members who have participated in the ETS Fairness Training, a rigorous internal training course, conduct this bias and sensitivity review. These staff members have been trained to identify and eliminate test questions that contain content that could be construed as offensive to, or biased against, members of specific ethnic, racial, or gender groups.

Assessment Director Review

As a final quality control step, the content area's assessment director or another senior-level content reviewer will read each item before it is presented to the CDE.

Assessment Review Panel (ARP) Review

The ARPs are panels that advise the CDE and ETS on areas related to item development for the CMA. The ARPs are responsible for reviewing all newly developed items for alignment to the California content standards. The ARPs also review the items for accuracy of content, clarity of phrasing, and quality.

Statewide Pupil Assessment Review (SPAR) Panel Review

The SPAR panel is responsible for reviewing and approving a single achievement test to be used statewide for the testing of students in California public schools in grades two through eleven. The SPAR panel representatives ensure that the test items conform to the requirements of *Education Code* Section 60614. The constructed-response writing tasks are also presented to the SPAR panel for review. If the SPAR panel rejects specific items and/or constructed-response writing tasks, the items and/or tasks are replaced with other items and/or tasks.

Data Review of Field Tested Items

ETS field tests newly developed items to obtain statistical information about item performance. This information is used to evaluate items that are candidates for use in operational test forms. The items and item statistics are examined carefully at data review meetings, which is where content experts discuss items that have poor statistics and do not meet the psychometric criteria for item quality. The CDE defines the criteria for acceptable or unacceptable item statistics. These criteria ensure that the item (1) has an appropriate level of difficulty for the target population; (2) discriminates well between examinees that differ in ability; and (3) conforms well to the statistical model underlying the measurement of the intended constructs.

The panel members also use the results of analyses for differential item functioning (DIF) to make judgments about the appropriateness of items for various subgroups. The panelists respond to questions such as:

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

The panelists make recommendations about whether to accept or reject each item for inclusion in the STAR item bank.

Quality Control of the Item Bank

After the data review meetings, items are placed in the item bank along with their statistics and reviewers' evaluations of their quality. ETS then delivers the items to the CDE through the STAR electronic item bank. The item bank database is maintained by a staff of application systems programmers, led by the Item Bank Manager, at ETS. All processes are logged; all change requests—including item bank updates for item availability status—are tracked; and all output and STAR item bank deliveries are quality-controlled for accuracy.

Quality of the item bank and secure transfer of the STAR item bank to CDE is very important. The ETS internal item bank database resides on a server within the ETS firewall; access to the SQL Server database is strictly controlled by means of system administration. The electronic item banking application includes a login/password system to authorize access to the database or designated portions of the database. In addition, only users authorized to access the specific database are able to use the item bank. Users are authorized by a designated administrator at the CDE and at ETS.

ETS has extensive experience in accurate and secure data transfer of many types including CDs, secure remote hosting, secure Web access, and secure file transfer protocol (SFTP), which is the current method to deliver the STAR electronic item bank to the CDE. In addition, all files posted on the SFTP site by the Item Bank staff are encrypted with a password.

The measures taken for ensuring the accuracy, confidentiality, and security of electronic files are as follows:

- Electronic forms of test content, documentation, and item banks are backed up electronically, with the backup media kept offsite, to prevent loss from system breakdown or a natural disaster.
- The offsite backup files are kept in secure storage, with access limited to authorized personnel only.
- Advanced network security measures are used to prevent unauthorized electronic access to the item bank.

Quality Control of Test Materials

Collecting Test Materials

Once the tests are administered, school districts return scorable and nonscorable materials within five working days after the last selected testing day of each test administration period. School districts return the CMA writing booklets within two working days after the makeup day for each administration. The freight return kits provided to the districts contain color-coded labels identifying scorable and nonscorable materials and labels with bar-coded information identifying the school and district. The school districts apply the appropriate labels and number the cartons prior to returning the materials to the processing center by means of their assigned carrier.

The use of the color-coded labels streamlines the return process. All scorable materials are delivered to the Pearson scanning and scoring facilities in Iowa City, Iowa. The nonscorable materials, including test booklets, are returned to the Security Processing Department in Pearson's Cedar Rapids, Iowa, facility. ETS and Pearson closely monitor the return of materials. The STAR Technical Assistance Center (TAC) at ETS monitors returns and notifies school districts that do not return their materials in a timely manner. STAR TAC contacts the district STAR coordinators and works with them to facilitate the return of the test materials.

Processing Test Materials

Upon receipt of the test materials, Pearson uses precise inventory and test processing systems, in addition to quality assurance procedures, to maintain an up-to-date accounting of all the testing materials within their facilities. The materials are removed carefully from the shipping cartons and examined for a number of conditions, including physical damage, shipping errors, and omissions. A visual inspection to compare the number of students recorded on the School and Grade Identification (SGID) sheets with the number of answer documents in the stack is also conducted.

Pearson's image scanning process captures security information electronically and compares scorable material quantities reported on the SGIDs to actual documents scanned. School districts are contacted by phone if there are any missing shipments or the quantity of materials returned appears to be less than expected.

Quality Control of Scanning

Before any STAR documents are scanned, Pearson conducts a complete check of the scanning system. ETS and Pearson create test decks for every test and form. Each test deck consists of approximately 25 answer documents marked to cover response ranges, demographic data, blanks, double marks, and other responses. Fictitious students are created to verify that each marking possibility is processed correctly by the scanning program. The output file generated as a result of this activity is thoroughly checked against each answer document after each stage to verify that the scanner is capturing marks correctly. When the program output is confirmed to match the expected results, a scan program release form is signed and the scan program is placed in the production environment under configuration management.

The intensity levels of each scanner are constantly monitored for quality control purposes. Intensity diagnostics sheets are run before and during each batch to verify that the scanner is working properly. In the event that a scanner fails to properly pick up items on the diagnostic sheets, the scanner is recalibrated to work properly before being allowed to continue processing student documents.

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

Post-scanning Edits

After scanning, there are three opportunities for demographic data to be edited:

- After scanning, by Pearson online editors
- After paper reporting, by district STAR coordinators (online demographic data corrections)
- After online corrections, by district STAR coordinators (post-P3 edits)

Demographic edits completed by the Pearson editors and by the district STAR coordinator online are included in the data used for the paper reporting and for the technical reports.

Quality Control of Image Editing

Prior to submitting any STAR operational documents through the image editing process, Pearson creates a mock set of documents to test all of the errors listed in the edit specifications. The set of test documents is used to verify that each image of the document is saved so that an editor would be able to review the documents through an interactive interface. The edits are confirmed to show the appropriate error, the correct image to edit the item, and the appropriate problem and resolution text that instructs the editor on the actions that should be taken.

Once the set of mock test documents is created, the image edit system completes the following procedures:

1. Scan the set of test documents.
2. Verify that the images from the documents are saved correctly.
3. Verify that the appropriate problem and resolution text displays for each type of error.
4. Submit the post-edit program.
5. Make changes and resubmit the post-edit program if errors are identified that require correction.

Pearson checks the post file against expected results to ensure the appropriate corrections are made. The post file will have all keyed corrections and any defaults from the edit specifications.

Quality Control of Answer Document Processing and Scoring

Accountability of Answer Documents

In addition to the quality control checks carried out in scanning and image editing, the following manual quality checks are conducted to verify that the answer documents are correctly attributed to the students, schools, districts, and subgroups:

- Grade counts are compared to the District Master File Sheets.
- Document counts are compared to the School Master File Sheets.
- Document counts are compared to the SGIDs.
- All school districts and grades are compared to the CDE County-District-School (CDS) Master File.

Any discrepancies identified in the steps outlined above are followed up by Pearson staff with the school districts for resolution.

Processing of Answer Documents

Prior to processing operational answer sheets and executing subsequent data processing programs, ETS conducts an end-to-end test. As part of this test, ETS prepares approximately 700 test cases covering all tests and many scenarios designed to exercise particular business rule logic. ETS marks answer sheets for those 700 test cases. They are then scanned, scored, and aggregated. The results at various inspection points are checked by psychometricians and Data Quality Services staff. Additionally, a post-scan test file of approximately 50,000 records are scored and aggregated to test a broader range of scoring and aggregation scenarios. These procedures assure that students and school districts get the correct scores when the actual scoring process is carried out.

Scoring and Reporting Specifications

ETS develops standardized scoring procedures and specifications so testing materials are processed and scored accurately. These documents include:

- General Reporting Specifications
- Form Planner Specifications
- Aggregation Rules
- ”What If” . . . List
- Edit Specifications
- Reporting Cluster Names and Item Numbers
- CST and CMA Matching Criteria
- Matching Criteria for Multiple-choice and Writing Answer Documents

Each of these documents is explained in detail in Chapter 7, which starts on page 89. The scoring specifications are reviewed and revised by the CDE, ETS, and Pearson each year. After a version that all parties endorse is finalized, the CDE issues a formal approval of the scoring and reporting specifications.

Matching Information on CMA Answer Documents

Answer documents are designed to produce a single complete record for each student. This record includes demographic data and scanned responses for each student; once computed, the scored responses and the total test scores for a student are also merged into the same record. All scores must comply with the ETS scoring specifications.

All STAR answer documents contain uniquely numbered lithocodes that are both scannable and eye-readable. The lithocodes allow all pages of the document to be linked throughout processing, even after the documents have been slit into single sheets for scanning. For those students using more than one answer document, lithocodes link their demographics and responses within a document while matching criteria are used to create a single record for all of the student’s documents. The documents are matched within grade using the matching criteria approved by the CDE.

Matching Multiple-choice and Writing Scores for ELA Grades Four and Seven

The CMA for ELA multiple-choice scores for all grades four and seven students are matched with their scores on CMA for Writing. The documents are matched within grades using the matching criteria approved by the CDE. Student documents that cannot be matched based on the approved criteria are reported separately. In addition, school districts receive an unmatched report with their reporting package listing the grades four and seven students for whom there are multiple-choice scores but no writing score, and grades four and seven students for whom there is a writing score but no multiple-choice score.

Storing Answer Documents

After the answer documents have been scanned, edited, scored, and have cleared the clean-post process, they are palletized and placed in the secure storage facilities at Pearson. The materials are stored until October 31 of each year, after which ETS requests permission to salvage the materials. After receiving CDE approval, the materials are salvaged in a secure manner.

Quality Control of Psychometric Processes

Score Key Verification Procedures

ETS and Pearson take various necessary measures to ascertain that the scoring keys are applied to the student responses as expected and the student scores are computed accurately. Scoring keys, provided in the form planners, are produced by ETS and verified thoroughly by performing various quality control checks. The form planners contain the information about an assembled test form; other information in the form planner includes the test name, administration year, subscore identification, and standards and statistics associated with each item. Various checks typically are performed before keys are finalized, as listed below:

1. The form planners are checked for accuracy against the Form Planner Specification document and the Score Key and Score Conversion document before the keys are loaded into the score key management system (SKM) at ETS.
2. The printed lists of the scoring keys are checked again once the keys have been loaded into the SKM system.
3. The sequence of items in the form planners are matched with their sequence in the actual test booklets.
4. The demarcations of various sections in the actual test book are checked against the list of demarcations provided by test development staff.
5. Scoring is verified internally at Pearson. ETS independently generates scores and verifies Pearson's scoring of the data by comparing the two results. Any discrepancies are then resolved.
6. The entire scoring system is tested using a test deck that includes typical and extremely atypical responses vectors, as described earlier in "Processing of Answer Documents" on page 222.
7. Classical item analyses are run on an early sample of data to provide an additional check of the keys. Although rare, if an item is found to be problematic, a followup process is performed that will exclude it from further analyses.

Quality Control of Item Analyses, and DIF

The psychometric analyses conducted at ETS undergo comprehensive quality checks by a team of psychometricians and data analysts. Detailed checklists are consulted by members of the team for each of the statistical procedures performed on each CMA. Quality assurance checks also include a review of the current year's statistics to statistics from previous years. The results of preliminary classical item analyses that provide a check on scoring keys are also reviewed by a senior psychometrician. The items that are flagged for questionable statistical attributes are sent to test development staff for their review; their comments are reviewed by the psychometricians.

A few additional checks are performed for each process, as described below:

Calibrations

During the calibration process, checks are made to ascertain that the correct options for the analyses are selected. Checks are also made on the number of items, number of examinees with valid scores, IRT Rasch item difficulties, standard errors for the Rasch item difficulties, and the match of selected statistics to the results on the same statistics obtained during preliminary item analyses. Psychometricians also perform detailed reviews of plots and statistics to investigate if the data fit the model.

Scoring Tables

Once calibrations are complete and raw-to-scale scoring tables are generated, the psychometricians carry out quality control checks on each scoring table. Scoring tables are checked to verify that all raw scores are included in the tables, scale scores increase as raw scores increase, and that the cut points for the performance levels are correctly identified. After all quality control steps are completed and any discrepancies are resolved, a senior psychometrician inspects the scoring tables as the final step in quality control.

Score Verification Process

Pearson utilizes the raw-to-scale scoring tables to compute scale scores for each student. ETS verifies Pearson's scale scores by following procedures, such as:

- Independently generating the scale scores for students in a small number of school districts and comparing these scores with those generated by Pearson; the selection of districts is based on the availability of data on all schools included in those districts, known as "complete districts"
- Reviewing longitudinal data for reasonableness; the results of the analyses are used to look at the tendencies and trends for the complete districts
- Reviewing longitudinal data for reasonableness using 99 percent of the entire testing population; the results are used to look at the trends for the state as well as few large districts.

The results of the longitudinal analyses are provided to the CDE and jointly discussed. Any anomalies in the results are investigated further and jointly discussed. Scores are released after explanations that satisfy both CDE and ETS are obtained.

Offloads to Test Development

The statistics based on classical item analyses and the IRT analyses are obtained at two different times in the testing cycle. The first time, the statistics are obtained on the calibration samples to ensure the quality of scaling and then on larger sample sizes to ensure the stability of the statistics that are to be used for future test assembly. Statistics used to generate DIF flags are also obtained from the larger samples and are provided to test development staff in specially designed Excel spreadsheets called "Statistical Offloads." The offloads are thoroughly checked by the psychometric staff before their release for test development review.

Quality Control of Reporting

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

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

All reports are required to include a single, accurate CDS code, a charter school number (if applicable), a school district name, and a school name. All elements conform to the CDE's official CDS code and naming records. From the start of processing through scoring and

reporting, the CDS Master File is used to verify and confirm accurate codes and names. The CDE Master File is provided by the CDE throughout the year as updates are available.

For students who use more than one answer document, the matching process, as described on page 223, provides for the creation of individual student records from which reports are created.

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

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

Excluding Student Scores from Summary Reports

ETS provides specifications to the CDE that document when to exclude student scores from summary reports. These specifications include the logic for handling answer documents that, for example, indicate the student tested but marked no answers and was absent, was not tested due to parent/guardian request, or did not complete the test due to illness. The methods for handling other anomalies are also covered in the specifications.