California Department of Education
Standards and Assessment Division

California Alternate Performance Assessment (CAPA)

Technical Report

Spring 2006 Administration

March 2007

Educational Testing Service (ETS)

Contract no. 5417

Table of Contents

Introduction **1**
Background 1
Education Code 60602 Legislative Intent 1
California Alternate Performance Assessment (CAPA) 2
Chapter 1. Test Overview **3**
Test Content 3
Target Population 3
Scores for Analysis and Reporting 3
Chapter 2. Content Validity **5**
CAPA Assessment Review Panel 5
Purpose 5
Composition 5
CAPA Item Writers 7
CAPA Development Procedures 8
Test Assembly 8
Test Specifications 8
Task Development 9
Internal Reviews 10
ARP Meetings for Review of CAPA Items 11
SPAR Review Panel 12
Chapter 3. Analysis **13**
Classical Item Analyses 14
Differential Item Functioning (DIF) Analyses 15
Reliability 17
Test Score Reliability 17
Standard Error of Measurement 18
Inter-Rater Reliability 18
Reliability of Classification and Decision Accuracy 19
Scaling and Production of Scoring Tables 19
Item Calibration and Equating 19
Equating Samples 22
Statistical Analysis Results 22
Chapter 4. Statewide Assessment Results **60**
Participation 60
Test Results 63
References 108
Appendix A. Individual Item Statistics **110**

List of Tables

Table 1.1 Summary of CAPA Assessment Levels 3
Table 1.2 Base Rubrics for CAPA Scoring 4
Table 2.1 CAPA ARP Member Education, by Subject and Total 6
Table 2.2 Statistical Specifications for CAPA 9
Table 3.1 Student Subgroups for DIF Analysis 15
Table 3.2 DIF Flags based on the ETS DIF Classification Scheme 17
Table 3.3 Operational Test and Associated Item Summary Statistics 24
Table 3.4 Frequency of Operational Item Scores: ELA 25
Table 3.5 Frequency of Operational Item Scores: Mathematics 26
Table 3.6 Raw Score Intercorrelations by Section: Level I 27
Table 3.7 Raw Score Intercorrelations by Section: Level II 27
Table 3.8 Raw Score Intercorrelations by Section: Level III 27
Table 3.9 Raw Score Intercorrelations by Section: Level IV 27
Table 3.10 Raw Score Intercorrelations by Section: Level V 28

Table 3.11 Ethnic Group DIF Statistics for C Category Items....28
Table 3.12 Disability Group DIF Statistics for C Category Items....28
Table 3.13 Double Rater Summary for Operational Items: Level I....30
Table 3.14 Double Rater Summary for Operational Items: Level II....31
Table 3.15 Double Rater Summary for Operational Items: Level III....32
Table 3.16 Double Rater Summary for Operational Items: Level IV....33
Table 3.17 Double Rater Summary for Operational Items: Level V....34
Table 3.18 Reliability of Classification and Decision Accuracy: Level I English–Language Arts....35
Table 3.19 Reliability of Classification and Decision Accuracy: Level I Mathematics....35
Table 3.20 Reliability of Classification and Decision Accuracy: Level II English–Language Arts....36
Table 3.21 Reliability of Classification and Decision Accuracy: Level II Mathematics....36
Table 3.22 Reliability of Classification and Decision Accuracy: Level III English–Language Arts....37
Table 3.23 Reliability of Classification and Decision Accuracy: Level III Mathematics....37
Table 3.24 Reliability of Classification and Decision Accuracy: Level IV English–Language Arts....38
Table 3.25 Reliability of Classification and Decision Accuracy: Level IV Mathematics....38
Table 3.26 Reliability of Classification and Decision Accuracy: Level V English–Language Arts....39
Table 3.27 Reliability of Classification and Decision Accuracy: Level V Mathematics....39
Table 3.28 CAPA 2006 Raw Score Means and Standard Deviations: Equating Sample vs. Total....40
Table 3.29 Evaluation of Common Items between New and Reference Test Forms....40
Table 3.30 Score Conversions: ELA Level I....41
Table 3.31 Score Conversions: Mathematics Level I....42
Table 3.32 Score Conversions: ELA Level II....43
Table 3.33 Score Conversions: Mathematics Level II....44
Table 3.34 Score Conversions: ELA Level III....45
Table 3.35 Score Conversions: Mathematics Level III....46
Table 3.36 Score Conversions: ELA Level IV....47
Table 3.37 Score Conversions: Mathematics Level IV....48
Table 3.38 Score Conversions: ELA Level V....49
Table 3.39 Score Conversions: Mathematics Level V....50
Table 3.40 Scale Score Frequency Distributions: Level I ELA and Mathematics....51
Table 3.41 Scale Score Frequency Distributions: Level II ELA and Mathematics....52
Table 3.42 Scale Score Frequency Distributions: Level III ELA and Mathematics....53
Table 3.43 Scale Score Frequency Distributions: Level IV ELA and Mathematics....54
Table 3.44 Scale Score Frequency Distributions: Level V ELA and Mathematics....55
Table 3.45 Raw Score Frequency Distributions: Level I Science....56
Table 3.46 Raw Score Frequency Distributions: Level III Science....57
Table 3.47 Raw Score Frequency Distributions: Level IV Science....58
Table 3.48 Raw Score Frequency Distributions: Level V Science....59
Table 4.1 Distribution of Students Across Test Levels....60
Table 4.2 Disability Distributions across All Levels....60
Table 4.3 Level I Disability Distributions....61
Table 4.4 Level II Disability Distributions....61
Table 4.5 Level III Disability Distributions....62
Table 4.6 Level IV Disability Distributions....62
Table 4.7 Level V Disability Distributions....63
Table 4.8 Performance Score Distributions for All Examinees*: English–Language Arts 2006....64
Table 4.9 Performance Score Distributions for All Examinees*: Mathematics 2006....66
Table 4.10 Performance Score Distributions*: Level I English Language Arts....68
Table 4.11 Performance Score Distributions*: Level I Mathematics....70
Table 4.12 Performance Score Distributions*: Level II English Language Arts....72
Table 4.13 Performance Score Distributions*: Level II Mathematics....74
Table 4.14 Performance Score Distributions*: Level III English Language Arts....76
Table 4.15 Performance Score Distributions*: Level III Mathematics....78
Table 4.16 Performance Score Distributions*: Level IV English–Language Arts....80
Table 4.17 Performance Score Distributions*: Level IV Mathematics....82
Table 4.18 Performance Score Distributions*: Level V English–Language Arts....84
Table 4.19 Performance Score Distributions*: Level V Mathematics....86
Table 4.20 Scale Score Distributions*: Level I English–Language Arts....88

Table 4.21 Scale Score Distributions*: Level I Mathematics 90
Table 4.22 Scale Score Distributions*: Level II English–Language Arts 92
Table 4.23 Scale Score Distributions*: Level II Mathematics 94
Table 4.24 Scale Score Distributions*: Level III English–Language Arts 96
Table 4.25 Scale Score Distributions*: Level III Mathematics 98
Table 4.26 Scale Score Distributions*: Level IV English–Language Arts 100
Table 4.27 Scale Score Distributions*: Level IV Mathematics 102
Table 4.28 Scale Score Distributions*: Level V English–Language Arts 104
Table 4.29 Scale Score Distributions*: Level V Mathematics 106
Table A.1 2006 CAPA Item Statistics: Level I 110
Table A.2 2006 CAPA Item Statistics: Level II 112
Table A.3 2006 CAPA Item Statistics: Level III 113
Table A.4 2006 CAPA Item Statistics: Level IV 115
Table A.5 2006 CAPA Item Statistics: Level V 117

Introduction

Background

In 1997 and 1998, the California State Board of Education adopted rigorous content standards in four major content areas: English–language arts, mathematics, history–social science, 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.

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 been extended to the 2006–07 school year (spring 2007 administration).

The primary goal of the STAR Program is to help measure how well students are learning required academic skills. The STAR Program has four components:

- California Standards Tests (CSTs) produced for California public schools
- California Achievement Tests, Sixth Edition Survey (CAT/6 Survey) for grades 3 and 7, published by CTB/McGraw-Hill
- California Alternate Performance Assessment (CAPA), an assessment produced for students with significant cognitive disabilities who are not able to take the CSTs or the CAT/6 Survey
- Aprenda: La prueba de logros en español, Tercera edición (Aprenda 3), published by Harcourt Assessment Inc.[1]

Education Code 60602 Legislative Intent

The results for tests within the STAR Program are used for three primary purposes:

1. Communicating students' progress in attaining proficiency on the state's academic standards to students, parents/guardians, and teachers. In developing the legislation for the STAR Program, the Legislature recognized that school districts will conduct their own, ongoing diagnostic assessments and provide information of the results of these assessments to parents and teachers on a regular basis. The Legislature also recognized that local diagnostic assessment is the primary way in which to identify academic strengths and weaknesses (*Education Code* Section 60602).
2. Informing decisions, along with local assessment data, that teachers and administrators make about helping students improve achievement and about improving the educational program.
3. Providing data for state and federal accountability programs that are used to calculate each school and district's Academic Performance Index (API) and Adequate Yearly Progress (AYP) to meet the requirements of the *No Child Left Behind Act of 2001*.

[1] Beginning in spring 2006, the State Board of Education designated the Aprenda: La prueba de logros en español, Tercera edición (Aprenda 3), to replace the SABE/2 as the designated primary language test (DPLT) for the STAR Program.

California Alternate Performance Assessment (CAPA)

The California Alternate Performance Assessment (CAPA) is an individually administered performance assessment that is used to measure the achievement of students with significant cognitive disabilities on California's Content Standards for English–language arts, mathematics, and science. The CAPA is part of the Standardized Testing and Reporting (STAR) Program. Administered annually in the spring, the STAR Program was authorized in 1997 by state law to measure how well students are learning the knowledge and skills identified in the California Content Standards. The CAPA was added to the STAR Program in 2003 to meet the requirements of the *Individuals with Disabilities Educational Act* of 1997 (IDEA) and the *No Child Left Behind Act* of 2001 (NCLB) that an alternate assessment be in place for those students with significant cognitive disabilities who are unable to take the general STAR California Standards Tests (CSTs) even with accommodations or modifications. Eligibility for participation in CAPA is determined by the student's individualized education program (IEP) team.

In 2006, CAPA was administered in each school during a 21-day window comprised of the ten (10) days before and ten (10) days after the day on which 85% of all instructional days for the school year were completed. This same window was used for the CSTs. Across the state, a total of 41,118 students in grades 2–11 participated in the CAPA.

This technical report outlines the statistical analyses that were carried out in support of the 2006 CAPA. Chapter 1 provides an overview of the test content, target population, and scoring procedures. Chapter 2 describes the procedures required to substantiate the content validity of the CAPA test. Chapter 3 details the statistical procedures that were carried out in support of the CAPA. These procedures include preliminary task analyses, differential task functioning analyses, equating and scaling, and various miscellaneous analyses. Chapter 4 presents statewide test results.

Chapter 1. Test Overview

Test Content

The CAPA is a standards-based, on-demand assessment designed to measure the progress of students with significant cognitive disabilities in meeting the California Content Standards. The CAPA assesses English–language arts (ELA) and mathematics for students in grades 2–11 and has a field-test section in science for students in grades 5, 8, and 10. CAPA has five assessment levels: Level I is for those students in grades 2–11 with the most significant cognitive disabilities who are functioning at or below the 24-months level of development. Levels II–V are age/grade appropriate. For students in ungraded educational settings, grade is determined by the formula: age on December 2 of the school year minus 5 equals grade. Table 1.1 summarizes the grades and content areas assessed by each CAPA level.

Table 1.1 Summary of CAPA Assessment Levels

Test Level	I	II	III	IV	V
Grades	2–11	2–3	4–5	6–8	9–11
Content Area	ELA	ELA	ELA	ELA	ELA
	Mathematics	Mathematics	Mathematics	Mathematics	Mathematics
	Science*		Science**	Science***	Science****
	* Grades 5, 8 and 10 only		** Grade 5 only	*** Grade 8 only	**** Grade 10 only

Target Population

Students with significant cognitive disabilities in grades two through eleven who are unable to take the STAR CSTs even with accommodations or modifications take the CAPA. Participation in CAPA and eligibility for Level I assessment is determined by the student's IEP team. Only students whose parents/guardians have submitted written requests to exempt them from STAR Program testing do not take the tests.

Scores for Analysis and Reporting

In 2006, each ELA and Mathematics test consisted of eight operational tasks and two field test tasks. An additional science section[2] was administered in grades 5, 8, and 10 for field-testing but no scores were reported. Student performance on each task is scored by one primary examiner, usually the child's teacher or other licensed or certificated staff member who is familiar to the student and who has completed the CAPA training. To establish scoring reliability, approximately 10% of students receive a second independent rating by a trained observer who is also a licensed or certificated staff member. The Level I assessment is scored using a 5-point rubric based on the level of independence with which the student completes a task. The Level II–V assessments are scored with a 4-point rubric based on the degree to which the student completes the task and includes task-specific qualifiers to aid in the objective scoring of each task. Table 1.2 provides the general rubrics applied to the CAPA tasks.

[2] This was the third year of science field tests at Levels I, III, and V and the second year at Level IV. In the 2006 science field test there were six field-test forms per level. Like the operational ELA and mathematics forms, each field test form had eight common or core items and two unique items per version, for a total of 20 items per level.

Table 1.2 Base Rubrics for CAPA Scoring

Level I		Levels II–V	
Score Points	**Description**	**Score Points**	**Description**
5	Complete task without prompts	4	Completes task with 100% accuracy
4	Completes task with a verbal or gestural prompt	3	Partially completes task (*scoring criteria specific to the task*)
3	Completes task with a physical or modeled prompt	2	Minimally completes task (*scoring criteria specific to the task*)
2	Attempts task	1	Attempts task
1	Orients to task	NR	No Response
NR	No Response		

For test scoring purposes, No Response (NR) ratings were assigned a task raw score of zero. Thus, CAPA raw scores range from 0 to 40 for Level I and from 0 to 32 for Levels II–V. Total raw scores for each content area on CAPA are converted from raw scores to scale scores. For CAPA, raw scores are converted to scale scores ranging from 15 to 60. Scale scores are also converted to the following performance levels: Far Below Basic, Below Basic, Basic, Proficient, and Advanced. The Basic and Proficient cut points are at scale scores of 30 and 35, respectively. The cut points for Below Basic and Advanced vary by CAPA assessment level and content area. (For information on the standard setting procedures used to establish cut points, see the *California Alternate Performance Assessment (CAPA) Standard Setting Technical Report*, submitted to the California Department of Education [CDE] on July 8, 2003.)

Chapter 2. Content Validity

Content validity refers to the degree to which the content of a test is congruent with the purpose of the testing. CAPA items were developed to align with the content standards that are representative of the broader content domains: English–language arts, mathematics, and science. Thus, the content-related evidence of validity concerns the extent to which the test items represent these specified content domains and cognitive dimensions.

Content validity also provides information about how well a performance task measures its intended construct. Such validity is determined by a critical review of the performance tasks by experts in the field. For the CAPA, these reviews are conducted by a number of experts in their designated areas from both the CDE and ETS. For these reviews, ETS senior content staff worked directly with CDE content consultants. The CDE content consultants each have extensive experience in K–12 assessments, particularly in their subject of expertise, and many are former teachers. At minimum, each CDE content consultant holds a bachelor's degree; most have advanced degrees in their area of expertise. All ETS content and test development staff have extensive experience with K–12 assessments, experience in teaching students with a broad range of disabilities, and understanding of the California standards. They each hold, at minimum, bachelor's degrees; most have advanced degrees within their areas of expertise.

CAPA Assessment Review Panel

In addition to the thorough content reviews completed by ETS content-area experts and the content staff at the CDE, all CAPA performance tasks are reviewed by a content-area Assessment Review Panel (ARP). The ARPs are advisory panels to ETS on areas related to task development for the CAPA. Their credentials are presented later in this document.

Purpose

ETS is responsible for working with ARPs as items are developed for the CAPA tests. For the 2006 development cycle, the ARPs were responsible for reviewing all newly developed items for alignment to the California content standards. The ARPs also reviewed the items for accuracy of item content, clarity of phrasing, and item quality. ETS provided 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. The ARPs may raise concerns in their examination of test items related to age/grade appropriateness and to gender, racial/ethnic, and socioeconomic bias.

The ARPs are also responsible for reviewing all newly developed performance tasks for alignment to the California content standards. Specifically, they determine if performance tasks are:

- Measuring the California standards as appropriate for the CAPA testing population
- Free from bias to the degree possible
- Interesting and appropriate to students tested at a particular grade/course level

Composition

Every effort is made to ensure that ARP committees include representation of gender and of the geographic regions and ethnic groups in California. For CAPA, efforts are also

made to ensure representation by members with experience serving California's diverse special education population.

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

- Experience with students who have severe cognitive disabilities;
- Experience with more than one type of disability;
- Three or more years of general teaching experience in grades kindergarten through twelve and in the content areas (English–language arts, mathematics, or science); or special education teaching experience;
- Bachelor's or higher degree in special education or in a subject area related to English–language arts, mathematics, or science; and
- Knowledge and experience with the California content standards in English–language arts, mathematics, or science.

School administrators, district/county content/program specialists, or university educators serving on the CAPA 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 the area of special education or areas related to English–language arts, mathematics, or science;
- Bachelor's or higher degree in special education or in a subject area related to English–language arts, mathematics, or science; and
- Knowledge of and experience with the California content standards in English–language arts, mathematics, or science.

For CAPA, efforts are also made to ensure representation by members with experience serving California's diverse special education population.

Current ARP members were recruited through an application process. Recommendations were solicited from districts and county offices of education in addition to CDE and SBE staff. Applications are received and reviewed throughout the year. ARP applications 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 the CDE and SBE staff for review and final approval. Upon approval, the applicant is notified that he or she has been selected to serve on the ARP committee. Table 2.1 shows the qualifications and background of the current CAPA ARP members.

Table 2.1 CAPA ARP Member Education, by Subject and Total

	ELA	Math	Science	Total
Total	**9**	**6**	**8**	**23**
Occupation				
Teacher or Program Specialist, Elementary/Middle School	2	1	0	3
Teacher or Program Specialist, High School	1	0	3	4
Teacher or Program Specialist, K–12	3	3	2	8
University Personnel	0	0	1	1
Other District Personnel (e.g., Director of Special Services, etc.)	3	2	2	7

	ELA	Math	Science	Total
Total	**9**	**6**	**8**	**23**
Highest Degree Earned				
Bachelor's Degree	2	1	0	3
Master's Degree	7	5	7	19
Doctorate	0	0	1	1
Credential (Members may hold multiple credentials)				
Elementary Teaching (Multiple Subjects)	6	3	1	10
Secondary Teaching (Single Subject)	0	1	4	5
Special Education	4	5	6	15
Reading Specialist	2	1	0	3
English Learner (CLAD,BCLAD)	1	0	1	2
Administrative	1	2	2	5
Other	1	0	1	2
None (teaching at university level)	0	1	1	2

Currently, there are no term limits for ARP members. While most members participate in the ARP meetings for only one STAR testing program, some members serve on more than one panel to encourage consistency in decisions among the STAR testing programs. ETS and CDE annually review the ARP membership for active participation. Members who have not attended a meeting within the last two years are notified that their invitation to participate may be withdrawn due to lack of attendance at meetings. In addition, ETS and CDE regularly review concerns about members whose conduct may be unprofessional and not conducive to the purpose of the ARP. If the concerns are determined to be valid, membership is revoked immediately.

CAPA Item Writers

The tasks selected for each CAPA test are developed by special panels of California teachers. Applicants for task writing were screened by senior ETS content staff; only those with strong content and special education backgrounds were approved for inclusion in the training. Thus, the participants were particularly experienced in writing to the standards assessed on CAPA. All item writers met the following minimum qualifications:

- Bachelor's or master's degree in special education or in a specified content area being tested
- Three or more years of general education teaching experience in the content areas (English–language arts, mathematics, and science); or special education teaching experience; or experience as a school administrator or program specialist
- Experience with students who have severe cognitive disabilities
- Experience with more than one type of disability
- Knowledge about the capabilities of the students taking these tests
- Knowledge and experience with California content standards in English–language arts, math, or science

Participants attended a general CAPA task development training session, and then were given specific subject-area training. After viewing multiple examples of previously written

CAPA tasks, participants were given task writing assignments. ETS facilitators provided feedback, and peer review methods were employed.

CAPA Development Procedures

The CAPA exams were constructed to measure the California content standards as well as to meet psychometric criteria for difficulty and reliability. The psychometric criteria were evaluated using projections based on item statistics from field-testing or previous operational administrations.

Test Assembly

Test blueprints for the components of the STAR Program (which includes CAPA) were proposed by ETS, reviewed and approved by the respective ARPs, 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 CAPA tests. For each test, the California content standards were used as the basis for choosing items. Additional technical targets (e.g., difficulty and discrimination) for test construction were established based on past characteristics of the tests, with the goal of maintaining parallel forms to the greatest extent possible.

Test Specifications

Statistical Specifications

The primary statistical targets used for CAPA test assembly in 2006 were the test information functions based on the item response theory (IRT) item parameters (one-parameter model) and average biserial correlation. 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 biserial correlation is a measure of how well the items discriminate among test takers and is related to the overall reliability of the test.

These specifications were developed from the analyses of test forms administered in the years from 2002 to 2005; the target values and ranges for the specifications are presented in Table 2.2 on page 9. The minimum target value for an item polyserial was set at 0.55 for each test.

Content Specifications

ETS developed all CAPA test items to conform to the SBE-approved content standards and test blueprints. The content blueprints for the CAPA can be found on the CDE Web site, at www.cde.ca.gov/ta/tg/sr/capablueprints.asp.

Table 2.2 Statistical Specifications for CAPA

Subject	CAPA Level	Mean AIS*	Min AIS*	Max AIS	Mean Polyserial	Min Polyserial
English–Language Arts	I	2.50–2.80	1.00	4.50	0.80–0.85	0.55
	II	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	III	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	IV	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	V	2.00–2.40	0.75	3.70	0.80–0.85	0.55
Mathematics	I	2.50–2.80	1.00	4.50	0.80–0.85	0.55
	II	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	III	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	IV	2.00–2.40	0.75	3.70	0.80–0.85	0.55
	V	2.00–2.40	0.75	3.70	0.80–0.85	0.55

*AIS=Average Item Score

Task Development

ETS senior content staff led the item writers in the task development and review process. In addition, experienced ETS content specialists and assessment editors reviewed each task during development and the forms construction processes. The lead assessment specialist for each content area worked directly with the other ETS assessment specialists to carefully review and edit each item for technical characteristics like quality (for example, one right answer, clearly stated stem, absence of clueing, plausibility of, distractors), match to standard, and conformity with California-approved item-writing practices.

ETS has maintained item specifications for each CAPA test. ETS followed the SBE-approved Item Utilization Plan to guide the development of the quantity of items for each subject area.

Item specification documents included the constructs to be measured and the California content standards included in the test blueprints. The item specifications help ensure that the CAPA tests consistently match the content standards from year to year. Item writing emphasis is determined in consultation with the CDE. The item specifications also provide specific and important guidance to item writers, and ensure that items are consistent in approach and written to measure the standards. The item specifications describe the general characteristics of the tasks for each content standard or content to be avoided, and define the content limits for the items. In summary, the specifications included the following:

- A statement of the strand or topic for the standard
- A full statement of the academic content standard, as found in each CAPA blueprint
- The expected cognitive level(s) of items written for the standard (low, medium, or high), as defined by ETS and approved by CDE
- The construct(s) appropriately measured by the standard
- A description of the kinds of tasks appropriate for the standard
- A description of specific kinds of items to be avoided, if any

- A description of appropriate stimuli (e.g., stimulus cards or manipulatives) for items
- A description of observable behaviors for science

Internal Reviews

After the tasks were written, ETS employed a series of internal reviews. The reviews established the criteria used to judge the content validity of the tasks, making sure that each item was measuring what it was intended to measure. The internal reviews also examined the overall quality of the tasks before they were prepared for presentation to CDE and the 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 CDE and the ARP and SPAR panels.

The ETS review process for the CAPA included the following:

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

Throughout this multi-step review process, the lead content area assessment specialists and development team members continually evaluated the relevance of the information being assessed, its relevance to the California content standards, its match to the test and item specifications, and its appropriateness to the population being assessed. Tasks that are only peripherally related to the test and item specifications, that do not measure core outcomes reflected in the California content standards, or that are not developmentally appropriate were eliminated early in this rigorous review process.

1. Internal Content Review

CAPA tasks materials received two reviews from the content area assessment specialists. These assessment specialists made sure that the tasks and related materials were in compliance with ETS's written guidelines for clarity, style, accuracy, and appropriateness for California students, and in compliance with the approved item specifications. Assessment specialists reviewed each item following the criteria below:

- Relevance of each item as the item relates to the purpose of the test
- Match of each item to the item specifications, including cognitive level
- Match of each item to the principles of quality item development
- Match of each item to the identified standard
- Difficulty of the item
- Accuracy of the content of the item
- CAPA-level appropriateness of the task
- Appropriateness of any stimulus cards or manipulatives

The assessment specialists also checked all tasks against their cluster classification codes, both to evaluate the correctness of the classification and to ensure that a given task is of a type appropriate to the outcome it was intended to measure. The reviewers accepted the item and classification as written, suggested revisions, or recommended that the tasks be discarded. These steps occurred prior to CDE review.

2. Internal Editorial Review

After the content area assessment specialists reviewed each item, a group of specially trained editors reviewed each item in preparation for review by CDE and the ARPs. The editors checked questions for clarity, correctness of language, appropriateness of language, adherence to the style guidelines, and conformity with accepted 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, conducted the next level of review. These trained staff members reviewed every item before it was prepared for CDE and ARP review. In addition, the review process promoted 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 groups, individuals with disabilities, and women) to the history and culture of the United States and the achievements of individuals within these groups

ARP Meetings for Review of CAPA Items

The ETS content area assessment specialists facilitated the CAPA ARP meetings. Each meeting began with a brief training session on how to review items. ETS provided this training, which consisted of the following steps:

- An overview of the purpose and scope of the CAPA
- An overview of CAPA test design specifications and blueprints
- An analysis of CAPA item specifications
- An overview of criteria for reviewing constructed-response tasks
- Review and evaluation of items for bias and sensitivity issues

The criteria for evaluating constructed-response 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(s)
- plausibility of the distracters
- bias and sensitivity factors

The committee was also trained on how to make recommendations for revising items. Guidelines for reviewing items were provided by ETS and approved by CDE. The guidelines for reviewing items are summarized below:

Item Guidelines: Does the item…

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

Stimulus Guidelines: Is the stimulus (if any) for the item…

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

As the first step of the item review process, panel members reviewed a set of tasks independently and recorded their individual comments. The next step in the review process was for the group to discuss each task. The content area assessment specialists facilitated the discussion and recorded all recommendations. These recommendations were recorded in a master item review booklet. Item review binders and other item evaluation materials also identified potential bias and sensitivity factors the ARP considers as part of its item reviews.

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

SPAR Review Panel

The statewide pupil assessment review (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, grades 2–11. At the SPAR panel meetings, all new items were presented in binders for review. The SPAR panel representatives ensured that the test items conformed to the requirements of *Education Code* Section 60614. The constructed-response tasks were also presented for review. If the SPAR panel rejected specific items and/or constructed-response tasks, the items and/or tasks were replaced. For the SPAR panel meeting, the item development coordinator or an ETS content specialist who had been requested in advance by CDE attended the opening session and remained at a nearby location or near a telephone to be available to respond to any questions during the course of the meeting.

Chapter 3. Analysis

Following the scanning of answer documents, student demographic and item response data were transmitted to the Educational Testing Service (ETS) statistical analysis division. ETS research and statistical analysis staff had primary responsibility for analyzing CAPA operational and field test data to ensure accuracy and validity of scoring. Most of the psychometric work was carried out using Generalized Analysis System (GENASYS), proprietary statistical analysis software developed by ETS. The GENASYS system includes components for establishing testing program statistical information, processing scores for students (including case sampling and scoring of multiple-choice items), traditional item analyses, and item response theory (IRT) analyses. The proprietary version of PARSCALE (Muraki & Bock, 1999) that is contained within GENASYS allows for estimation of IRT item parameters for polytomously scored items. It has been thoroughly tested and is currently utilized by several high-stakes testing programs administered by ETS, including the California STAR assessments and the California High School Exit Examinations (CAHSEE), as well as the National Assessment of Educational Progress (NAEP), and the Test of English as a Foreign Language (TOEFL). All technical support and analyses were carried out in accordance with both the *Standards for Educational and Psychological Testing* (1999), issued jointly by the American Educational Research Association, the American Psychological Association, the National Council on Measurement in Education, and the *ETS Standards for Quality and Fairness* (2002).

ETS staff verified the output from the scoring programs to ensure the accuracy of the scoring process. After the operational administration, ETS analysis staff ran a set of preliminary item analyses based on a sample of the early answer document receipts. The preliminary item analyses were used to assure the accuracy of the scoring and to get an initial indication of how items were functioning. A minimum of approximately 1,500 answer sheets at each assessment level from a heterogeneous sample of different schools (that is, diverse in geography and demographic characteristics) were used. ETS instituted a set of flags that automatically identified items with questionable performance characteristics.

Content specialists examined all flagged items to verify that the items in the published test books were correct and unambiguous. In addition to preliminary item analyses derived statistically, ETS compared hand scoring for a small sample of student answer sheets to the scanned results to confirm the accuracy of scanning and scoring.

After scoring, ETS subjected all test items to extensive statistical analyses. These analyses showed which items were at an appropriate difficulty level for the testing population and screened for differential item difficulty for subgroups of the state's population. Additionally, ETS content specialists confirmed the item-to-standard match for each of the content areas.

The analysis of the test data can be broken down into several components:

1. Classical item analyses
2. Differential item functioning (DIF) analyses
3. Reliability analyses; and
4. Scaling and production of scoring tables.

In the sections that follow, the analysis procedures for each component are described in detail. Tables summarizing the analyses are provided at the end of the chapter.

Classical Item Analyses

Classical item analyses involve computing, for every item in each form, a set of statistics based on classical test theory. Each statistic is designed to provide some key information about the quality of the item from an empirical perspective. The statistics calculated for CAPA operational and field test analyses are described below.

- **Average Item Score (AIS):** For polytomously scored items, this statistic indicates the average rating earned on the item. Desired values generally fall within the range of 30–80% of the maximum item score. Occasionally, items that fall outside this range can be justified for inclusion in an item bank or a test form based upon the quality and educational importance of the item content or to better measure students with very high or low achievement. CAPA rubrics range from 0 to 4 or from 0 to 5 depending on the test level. As a result, the average item score for a CAPA item falls between 0 and either 4 or 5 corresponding to the rubric in use. For Level I items, which are scored on a 0–5 point rubric, 30% is represented by the value 1.50 and 80% is represented by the value 4.00. For Levels II–V items, which are scored on a 0–4 point rubric, 30% is represented by the value 1.20 and 80% is represented by the value 3.20.
- **Polyserial correlation of the item score with the total test score:** This statistic describes the relationship between performance on the specific item and performance on the total test. It is sometimes referred to as a discrimination index because it is an indicator of the degree to which students who do well on the total test also do well on this item. Items with negative or extremely low correlations ($\rho < 0.05$) can indicate serious problems with the item itself or can indicate that students have not been taught the content. Due to the small number (8) and similarity of items, CAPA item-total correlations tend to be higher than seen on longer tests with more heterogeneous items. Based on the range of polyserials produced in field test analyses, an indicator of poor discrimination was set to .60, a relatively low polyserial for CAPA.
- **Pearson Product Moment Correlation:** For all items, this statistic was also computed and is included in Appendix A. The formula for this correlation between two variables, X and Y, is

$$\rho_{xy} = \frac{\mathrm{cov}(x, y)}{\sqrt{\mathrm{var}(x)\,\mathrm{var}(y)}} \qquad [3.1]$$

For the CAPA analyses, flags were defined in order to identify items with extreme values. Flagged items were subject to additional scrutiny by statistical analysis and test development staff. The following flagging criteria were applied to all items tested in Spring 2006:

- **Difficulty Flags**:
  - A: Low average item score (e.g. below 1.5 at Level I; below 1.2 at Levels II–V)
  - H: High average item score (e.g. above 4.0 at Level I; above 3.2 at Levels II–V)

- **Discrimination Flag:**
  - R: Polyserial correlation less than .60
- **Omit/Nonresponse/Flag:**
  - O: Omit/Nonresponse rates greater than 5%

Results of these analyses are presented in Appendix A on page 110.

Differential Item Functioning (DIF) Analyses

One of the goals of test development is to assemble a set of items that provides an estimate of a student's ability that is as fair and accurate as possible for all groups within the population. DIF statistics are used to identify those items for which identifiable groups of students with the same underlying level of ability have different probabilities of answering correctly.

If the item is differentially more difficult for an identifiable subgroup when conditioned on ability, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might be related to actual differences in relevant knowledge or skills (item impact) or statistical Type 1 error. As a result, DIF statistics are used to identify potential sources of item bias. Subsequent review by content experts and bias/sensitivity committees are required to determine the source and meaning of performance differences.

In the CAPA DIF analyses, DIF statistics were estimated for all major subgroups with sufficient sample size. These groups were identified by CDE and are listed in Table 3.1. Items with statistically significant differences in performance were flagged so that items could be carefully examined for possible biased or unfair content that was undetected in earlier fairness and bias content review meetings held prior to form construction.

Table 3.1 Student Subgroups for DIF Analysis

DIF Type	Reference Group	Focal Group
Gender	Male	Female
Race/Ethnicity	White	African American
		Hispanic/Latin American
		American Indian
		Asian
		Pacific Islander
		Filipino
		Combined Asian Group (Asian/Pacific Islander/Filipino)
Disability	Mental Retardation	Hard of Hearing
		Deafness
		Speech or Language Impairment
		Visual Impairment
		Emotional Disturbance
		Orthopedic Impairment
		Other Health Impairment
		Specific Learning Disability
		Deaf-Blindness

DIF Type	Reference Group	Focal Group
		Multiple Disabilities
		Autism
		Traumatic Brain Injury

DIF analyses of the CAPA polytomously scored items were completed using two procedures. The first is the Mantel-Haenszel (MH) ordinal procedure, which is based on the Mantel procedure (Mantel, 1963; Mantel & Haenszel, 1959), compares the proportions of matched examinees from each group in each polytomous item-response category—that is, the probability of a given item score for the studied groups of interest after matching on total test score. As with dichotomously scored items, the common odds ratio is estimated across all categories of matched examinee ability. The resulting estimate is interpreted as the relative likelihood of a given item score for members of two groups when matched on ability. As such, the common odds ratio provides an estimated effect size where a value of unity indicates equal odds, and thus no DIF (Dorans & Holland, 1993). The corresponding statistical test is Ho: $\alpha = 1$, where α is a common odds ratio assumed equal for all matched score categories $s = 1$ to S. Values less than unity indicate DIF in favor of the focal group, a value of unity indicates the null condition, and a value greater than one indicates DIF in favor of the reference group. The associated $MH\chi^2$ is distributed as a chi-square random variable with 1 degree of freedom.

The Mantel chi-square statistic is used in conjunction with a second procedure, the standardization procedure (Dorans & Schmitt, 1993). This procedure produces a DIF statistic based on the standardized mean difference (SMD) in average item scores between members of two groups who have been matched on their overall test score. The SMD compares the item means of the two studied groups after adjusting for differences in the distribution of members across the values of the matching variable (total test score). A negative SMD value means that, conditional on the matching variable, the focal group has a lower mean item score than the reference group. In contrast, a positive SMD value means that, conditional on the matching variable, the reference group has a lower mean item score than the focal group. The SMD is divided by the standard deviation (SD) of the total group item score in its original metric to produce an effect-size measure of differential performance.

The ETS classification scheme puts items into three DIF categories on the basis of a combination of statistical significance of the Mantel chi-square statistic and the magnitude of the SMD effect-size:

- *A items or negligible DIF:* The Mantel chi-square statistic is not statistically significant (at the 0.05 level) or $|SMD/SD| < 0.17$.
- *B items or intermediate DIF:* The Mantel chi-square statistic is statistically significant (at the 0.05 level) and $0.17 \leq |SMD/SD| < 0.25$
- *C items or large DIF:* The Mantel chi-square statistic is statistically significant (at the 0.05 level) and $|SMD/SD| > 0.25$.

Items classified as B+ or C+ tend to be easier for members of the focal group than for members of the reference group with comparable total scores. Items classified as B- or C- tend to be more difficult for members of the focal group than for members of the reference group whose total scores on the test are like those of the focal group. (See Table 3.2.)

Table 3.2 DIF Flags based on the ETS DIF Classification Scheme

Flag	Descriptor
A-	Low DIF favoring members of the reference group
B-	Moderate DIF favoring members of the reference group
C-	High DIF favoring members of the reference group
A+	Low DIF favoring members of the focal group
B+	Moderate DIF favoring members of the focal group
C+	High DIF favoring members of the focal group

Following standard ETS procedure, items classified in Category C were sent for review by test development staff and/or content review committees to consider any identifiable characteristics that may have contributed to the differential item functioning. These items might be revised for additional field testing or removed from the item pool.

Reliability

Reliability is used to measure the extent to which an assessment will yield the same results when administered in different occasions, locations, or populations, when the two administrations do not differ in relevant variables. Reliability coefficients are usually forms of correlation coefficients. The forms of reliability below measure different dimensions of reliability and thus any or all might be used in assessing the reliability of CAPA.

Test Score Reliability

Reliability focuses on the extent to which differences in test scores reflect true differences in the knowledge, ability, or skills being tested rather than fluctuations due to chance or factors other than those tested. The variance in the distributions of test scores—essentially, the differences among individuals—is partly due to real differences in the knowledge, skills, or ability being tested (true variance) and partly due to random errors in the measurement process (error variance). The number used to describe reliability is an estimate of the proportion of the total variance that is true score variance. Several different ways of estimating this proportion exist.

When the goal is to estimate the precision of a set of test scores from a single administration, a measure of internal consistency is frequently used to estimate reliability. For the CAPA, a measure of internal consistency called coefficient alpha (α) was used for estimating the reliability of the test scores. The formula for coefficient alpha, given by

$$\rho_{XX'} \geq \frac{k}{k-1}\left(1-\frac{\sum \sigma_i^2}{\sigma_X^2}\right), \qquad [3.2]$$

where k is the number of items on the test, $\sum \sigma_i^2$ is item score variance summed over all items, and σ_X^2 is observed-score variance, reflects the fact that the reliability of a set of test scores is influenced by the observed-score variance. Coefficient alpha can be thought of as a lower bound to a theoretical reliability coefficient known as the “coefficient of precision,” as well as the lower bound of the proportion of variance in the test scores explained by common factors underlying item performance.

Internal consistency measures apply only to the test form being analyzed. They do not take into account form-to-form variation due to equating limitations or lack of parallelism,

nor are they responsive to day-to-day variation due, for example, to the state of each examinee's 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 upon repeated testing occasions with parallel forms.

Standard Error of Measurement

The squared standard error of measurement (SEM) is an estimate of error score variance, σ_E^2. In Classical Test theory (CTT), the SEM is assumed equal along the measurement scale and is estimated as a function of the standard deviation of observed scores and test reliability coefficient:

$$\mathrm{SEM} = s_{\mathrm{x}} \sqrt{1 - \mathrm{r}_{\mathrm{xx'}}} \ , \qquad [3.3]$$

where

SEM = standard error of measurement,

s_x = standard deviation of observed scores, and

$r_{xx'}$ = coefficient of reliability (alpha).

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

In contrast to CTT, in IRT framework the CSEM is estimated as a function of measured ability. It is typically smaller in scale score units towards the center of the scale 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:

$$\mathrm{CSEM}(\hat{\theta}) = \frac{1}{\sqrt{\mathrm{I}(\theta)}} a, \qquad [3.4]$$

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 the scaling factor, a, in order to place it on the scale score metric. CSEMs are provided for each raw score point of the operational CAPA tests.

Inter-Rater Reliability

Inter-rater reliability addresses the consistency of the implementation of a rating system. For the CAPA, approximately 10% of students received two ratings, one by the primary examiner and a second independent rating by a trained observer. Consistency between the two ratings was evaluated with the following statistics:

- Number and percentage of exact agreement between raters
- Number and percentage of adjacent agreement between raters
- Number and percentage of nonadjacent scores between raters

- Mean absolute difference between ratings for the examiner and the observer
- Correlation between ratings for the examiner and the observer

The Inter-Rater Reliabilities are presented by level in Tables 3.13–3.17 on pages 30–34.

Reliability of Classification and Decision Accuracy

The methodology used for estimating the reliability of performance-level classification decisions as described in Livingston and Lewis (1995) provides estimates of decision accuracy and classification consistency.

> The term *accuracy*…refers to the extent to which the actual classifications of test takers (on the basis of their single-form scores) agree with those that would be made on the basis of their true scores, if their true scores could somehow be known. The term *consistency* refers to the agreement between the classifications based on two non-overlapping, equally difficult forms of the test. (Livingston & Lewis, 1995, p.178)

For CAPA, it is implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.12). For each level and test, RELCLASS-COMP estimates true scores and single-form scores on forms parallel to the one actually given. RELCLASS-COMP estimates decision accuracy using an estimated joint distribution of reported performance level classifications on the current form of the exam and the performance level classifications based on an all-forms average (true score). RELCLASS-COMP estimates decision consistency using an estimated joint distribution of reported performance level classifications on the current form of the exam and performance level classifications on the alternate (parallel) form.

In each case, the proportion of performance level classifications with exact agreement is the sum of the entries in the diagonal of the contingency table representing the joint distribution. Reliability of classification at each performance level cut score is estimated by collapsing the joint distribution at the passing score boundary into a 2-by-2 table and summing the two entries in the diagonal. RELCLASS COMP also computes the effective length of the test. The Reliability of Classification and Decision Accuracies are presented by level in Tables 3.18–3.27, pages 35–39.

Scaling and Production of Scoring Tables

Item Calibration and Equating

The purpose of item calibration and equating is to create a common scale for expressing the difficulty estimates of all the items across versions within a test. When first established, the scale commonly has a mean score of 0 and a standard deviation of 1. It should be noted that this scale is often referred to as the “theta” metric and is not used for reporting purposes because the values typically range from -3 to +3. Therefore, following calibration and equating, the scale is usually transformed to a reporting scale (also known as a scale score; see scaling section below), which can be more meaningfully interpreted by students, teachers, and other stakeholders.

The IRT model used to calibrate the CAPA test items was the 1-parameter partial credit (1PPC) model, a more restrictive model of the generalized partial-credit model (Muraki, 1992) where all items are assumed to be equally discriminating. The fundamental equation of this model is the probability that a person with proficiency θ_k on scale k will

have, for the jth item, a response x_j that is scored in the ith of m_j ordered score categories:

$$P(x_j = i \mid \theta_k, a_j, b_j, d_{j0}, ..., d_{jm_j-1}) = \frac{\exp \sum_{v=0}^{i} 1.7 a_j (\theta_k - b_j + d_{jv})}{\sum_{g=0}^{m_j-1} \exp \sum_{v=0}^{g} 1.7 a_j (\theta_k - b_j + d_{jv})} \equiv P_{ji}(\theta_k), \qquad [3.5]$$

where:

m_j is the number of categories in the response to item j;

x_j is the response to item j, with possible values 0, 1, …, $m_j - 1$;

a_j is the slope parameter;

b_j is the item location parameter characterizing overall difficulty; and

d_{jv} is the category v threshold parameter

All IRT analyses were conducted using the proprietary version of PARSCALE (Muraki & Bock, 1999) that is contained within GENASYS. In IRT-based equating, once two forms have been placed on the same IRT scale through their common items, raw scores on a new form can be converted to raw scores on an old form. These converted raw scores can then be transformed to scale scores through table lookup and linear interpolation. The "base" or "reference" calibrations for the CAPA tests were established by calibrating samples of data from the 2003 administration. This established a scale to which subsequent item calibrations could be linked. For the purpose of linking to the base year, each 2006 CAPA form was constructed to include a set of five items that had been administered operationally in 2005. The 2006 items were placed on the base scale through the set of common items from the 2005 forms. (See Table 3.29 on p. 40.)

The procedures used for equating the CAPA tests involved three steps:

1. Item calibration
2. Item parameter scaling, and
3. True score equating.

These steps are described below.

Step 1: For the item calibrations, the PARSCALE program was constrained by setting a common discrimination value for all items equal to 1.0 / 1.7 (or 0.588). The resulting estimation was equivalent to the Rasch partial credit model for polytomously scored items.

The PARSCALE calibrations were 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 were used as starting values for a second PARSCALE run, in which the subject prior distribution was updated after each expectation maximization (EM) cycle with no constraints. For both stages, the metric of the scale was controlled by the constant discrimination parameters. This approach was used to obtain unequated 2006 item parameter estimates. Once these estimates were obtained, each task was evaluated using fit statistics in conjunction with plots of model-data fit that were generated by the GENASYS system. Items flagged for

potential misfit were evaluated with respect to their impact on test specifications, psychometric quality, and coverage of academic content standards.

Step 2: Next, parameter estimates were transformed to the CAPA base scales using the Stocking and Lord (1983) procedure, with an embedded anchor set drawn from the 2005 forms. In the case of the 1-parameter model, this procedure is equivalent to setting the mean of the new item parameter estimates for the common items equal to the mean of the previously scaled estimates. As commonly done in this approach, the linking process was carried out iteratively by inspecting differences between the transformed new and old (reference) estimates for the linking items, and, if necessary, removing items for which the item difficulty estimates changed significantly. The differences were calculated using the following formula:

$$\mathrm{WRMSD} = \sqrt{\sum_{j=1}^{61} w_j \left[P_n\left(\theta_j\right) - P_r\left(\theta_j\right)\right]^2} \quad , \qquad [3.6]$$

where θ_j ranges from –3.0 to 3.0 by 0.1, w_j is a weight equal to the proportion of estimated abilities from the transformed new form in interval j, $P_n(\theta_j)$ is the probability of a given score for the transformed new form item at ability level j, and $P_r(\theta_j)$ is the probability of the same score for the old (reference) form item.

Simply put, transformed new and old parameter estimates were evaluated using weighted (based on the reference form abilities) root mean square difference (WRMSD) statistics that summarize differences in item characteristic curves (ICCs).

The linking criteria required removing items with an WRMSD greater than 0.625 for Level I and 0.500 for Levels II–V. For the 2005 CAPA tests, no linking items were eliminated.

Step 3: Once the new calibrations for each test were linked to the Rasch scale, defined by the reference calibrations, IRT true score equating procedures were utilized to transform the new form number-correct scores to their respective reference form scale scores. The true score equating procedure is based on the relationship between raw scores and ability. For tests consisting entirely of multiple-choice items, this is the well-known relationship defined in Lord (1980; eq. 4–5):

$$\xi(\theta) = \sum_{i=1}^{n} P_i(\theta), \qquad [3.7]$$

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

For all CAPA tests, ξ(θ) is based on polytomously scored performance (constructed response) items[3], and the relationship can be defined as:

$$\xi(\theta) = \sum_{j=1}^{ncr} \sum_{x=1}^{m} s_x P_{xj}(\theta), \qquad [3.8]$$

where *ncr* is the number of constructed response items in the test, *m* is the number of score categories in each polytomously scored item, s_x is the score value for category x,

[3] See Chapter 1 for the scoring rubric.

and $P_{xj}(\theta)$ is the probability of a score in category x at ability θ (defined by the Rasch partial credit model). For Level I there are six possible scores per item: 0, 1, 2, 3, 4, and 5. For Levels II–V there are five possible scores: 0, 1, 2, 3, and 4. A score of zero is only assigned for students who fail to respond to the prompt.

For each integer score ξ_n on the new form, the true score equating procedure first solved for the corresponding ability level using equation 7. Next, the procedure used that ability level to find the corresponding score ξ_b on the base or reference form. Finally, each score ξ_b was transformed to the appropriate CAPA scale score scale using the reference form CAPA raw-score-to-scale-score conversion tables and linear interpolation. In particular, the theta scale was linearly transformed onto the 15–60 scale by holding the raw scores for the Proficient and Basic cuts obtained in the standard setting fixed at scale scores of 35 and 30, respectively. Remaining scale score cuts for Advanced, Below Basic, and Far Below Basic were allowed to fall along the scale and identified by matching with the corresponding raw score identified in the standard setting process.

Equating Samples

The 2006 equating samples were selected from available student records in a data file obtained in early June. These data consisted of approximately 17 to 25 percent of the total CAPA testing data that were eventually available when all testing was completed. The use of partial student samples for equating was necessitated by score reporting deadlines, and was approved by the CDE. Only students with valid results on the CAPA tests were included in the equating samples.

Statistical Analysis Results

This section contains the tabled results of the analyses described above. Individual classical item statistics (AIS and polyserial correlations) and associated flags are provided in Appendix A on page 110.

Table 3.3 provides a general statistical summary of each CAPA ELA and mathematics test by level. For each level and content area, test level statistics include the scale score mean, standard deviation, median, and range, as well as the internal consistency reliability (coefficient alpha) and standard error of measurement. Item/task summary statistics include the means and standard deviations for the average item scores, polyserial correlations, and Rasch difficulties. Reliability coefficients ranged from 0.88 (Mathematics Level II) to 0.92 (Mathematics Level I). Mean equated Rasch difficulties ranged from -0.62 (ELA Level II) to 0.10 (ELA Level I).

Tables 3.4 and 3.5 provide the frequency and percent of operational item scores for each level of ELA and Mathematics. More than 50% of examinees were awarded the highest item score on eight of the 40 ELA items, four of which are on the Level II test. In Mathematics, more than 50% of examinees received the highest rating on 10 of the 40 items, with six of the items from the Level II test.

Tables 3.6–3.10 provide the content area raw score means and standard deviations and the raw score intercorrelations for each CAPA level. Intercorrelations ranged from 0.80 (Level IV ELA and Science) to 0.90 (Level I Mathematics and Science) indicating a moderate to high degree of correlation between performance on the content areas. Given the functional nature of many of the standards being assessed on CAPA, this degree of correlation is not surprising.

DIF SMD statistics for operational and field test items flagged for C category ethnic and disability DIF are provided in Tables 3.11 and 3.12, respectively. Five ELA items and one science item were flagged for ethnic DIF. No items were flagged for gender DIF. A number of items in all three content areas were flagged for disability DIF, with the majority involving the Mental Retardation/Autism comparison in ELA.

Tables 3.13–3.17 summarize the results of examiner and observer double ratings of each operational ELA and Mathematics item. Included are the mean and standard deviation of assigned ratings, the percentage of exact and adjacent ratings and percentage of ratings that differ by more than one score point (neither exact nor adjacent), as well as the mean absolute difference and the correlation between ratings for the examiner and the observer. Mean absolute differences (MAD) range from 0.04–0.22. The correlations between examiner and observer ratings range from 0.91–0.99. In general, the examiners were not any more or less stringent, on average, than the observers.

The classification reliability for both accuracy and consistency are reported in Tables 3.18–3.27. The decision accuracy for ELA ranges from 0.71–0.76 across all performance levels and from 0.90–0.93 for the proficient and above classification. The decision accuracy for mathematics ranges from 0.67–0.76 across all performance levels and from 0.91–0.92 for the proficient and above classification. The decision consistency for ELA ranges from 0.62–0.69 across all performance levels and from 0.87–0.90 for the proficient and above classification. The decision consistency for Mathematics ranges from 0.58–0.68 across all performance levels and from 0.87–0.89 for the proficient and above classification.

Table 3.28 provides the means and standard deviations of the 2006 CAPA equating samples, as well as the 2006 total examinee population. The equating samples were very similar to the total examinee population.

Table 3.29 presents, for each CAPA test, the number of common items between the 2006 (new) and 2005 (reference) test forms, the number of items removed from the common item sets, the final correlations between the new and reference difficulty estimates, and the average WRMSD statistic (see equation 1) across the final set of common items. These results indicate that the new and reference difficulty estimates were highly correlated (0.96 or higher) and similar in magnitude (average WRMSD values of 0.11 and lower).

The raw-to-scale-to-performance level conversions and conditional standard errors of measurement (CSEM) are presented in Tables 3.30–3.39 for ELA and Mathematics. Tables 3.40–3.44 present the scale score frequency distributions for ELA and Mathematics by level. In Level I ELA approximately 15% of students obtained the highest scale score. Note that gaps in the table indicate scale scores that were not in the 2006 raw-to-scale conversion tables. Tables 3.45–3.48 present the raw score frequency distributions for Science field tests in Levels I, III, IV, and V.

Table 3.3 Operational Test and Associated Item Summary Statistics

Level/Content	I		II		III		IV		V	
	ELA	Math	ELA	Math	ELA	Math	ELA	Math	ELA	Math
Scale Score Information										
Number of Examinees	8923	8898	6206	6189	6269	6259	10353	10337	9367	9347
Mean Score	43	33	38	39	38	41	37	35	38	35
SD*	13.03	11.50	8.13	8.32	10.24	10.92	10.66	11.01	10.57	10.09
Possible Range	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60
Obtained Range	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60	15–60
Median	45	33	37	39	37	41	36	34	38	35
Reliability	0.90	0.92	0.89	0.88	0.90	0.91	0.91	0.90	0.91	0.91
SEM**	4.12	3.25	2.70	2.88	3.24	3.28	3.20	3.48	3.17	3.03
Item Information										
Number of Items	8	8	8	8	8	8	8	8	8	8
Mean AIS***	2.90	2.72	2.83	2.63	2.61	2.96	2.64	2.75	2.75	2.44
SD AIS***	0.49	0.24	0.32	0.39	0.26	0.35	0.35	0.24	0.46	0.38
Range AIS***	2.22–3.55	2.28–3.02	2.43–3.23	2.04–3.04	2.17–2.87	2.43–3.51	2.23–3.13	2.49–3.19	2.02–3.22	1.77–2.96
Possible Range	0-5	0-5	0-4	0-4	0-4	0-4	0-4	0-4	0-4	0-4
Mean Polyserial	0.81	0.84	0.78	0.78	0.80	0.80	0.82	0.79	0.82	0.83
SD Polyserial	0.04	0.02	0.01	0.05	0.05	0.05	0.06	0.05	0.04	0.03
Range Polyserial	0.71–0.84	0.80–0.87	0.75–0.80	0.67–0.84	0.72–0.85	0.73–0.86	0.74–0.88	0.72–0.87	0.75–0.87	0.76–0.87
Mean Rasch Difficulty	0.10	-0.27	-0.62	-0.10	-0.27	-0.58	-0.39	-0.59	-0.21	-0.29
SD Rasch Difficulty	0.38	0.22	0.37	0.43	0.36	0.52	0.45	0.32	0.63	0.44
Range of Rasch Difficulty	(-0.43)–0.63	(-0.58)–0.17	(-1.16)–(-0.14)	(-0.71)–0.56	(-0.71)–0.15	(-1.45)–0.03	(-1.16)–0.13	(-1.23)–(-0.26)	(-0.94)–0.74	(-0.97)–0.45

*Standard Deviation
**Standard Error of Measurement
***AIS = Average Item Score

Table 3.4 Frequency of Operational Item Scores: ELA

ELA	Score/Item	1		2		3		4		5		6		7		8	
Level		Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
I	0	738	7.97	918	9.91	1176	12.70	1516	16.37	646	6.98	1134	12.25	1564	16.89	923	9.97
	1	1193	12.88	1096	11.84	1513	16.34	1893	20.44	1005	10.85	1314	14.19	1766	19.07	1229	13.27
	2	1515	16.36	1529	16.51	2322	25.08	2013	21.74	1034	11.17	1781	19.24	2236	24.15	1004	10.84
	3	1022	11.04	682	7.37	962	10.39	939	10.14	625	6.75	836	9.03	1013	10.94	813	8.78
	4	1043	11.26	1024	11.06	685	7.40	788	8.51	1092	11.79	840	9.07	868	9.37	1026	11.08
	5	3404	36.76	3654	39.46	2242	24.21	1734	18.73	4461	48.18	2984	32.23	1414	15.27	3879	41.89
II	0	325	4.89	188	2.83	271	4.08	219	3.29	138	2.08	137	2.06	216	3.25	164	2.47
	1	777	11.69	1524	22.92	374	5.62	1654	24.88	1189	17.88	407	6.12	994	14.95	1035	15.57
	2	1420	21.36	1844	27.73	739	11.11	1507	22.67	558	8.39	952	14.32	1950	29.33	434	6.53
	3	1794	26.98	719	10.81	1619	24.35	1163	17.49	1064	16.00	1202	18.08	1262	18.98	1349	20.29
	4	2141	32.20	2192	32.97	3456	51.98	1901	28.59	3491	52.50	3743	56.29	2026	30.47	3459	52.02
III	0	265	3.91	134	1.98	237	3.50	120	1.77	126	1.86	537	7.93	330	4.87	411	6.07
	1	1110	16.38	1063	15.69	2630	38.81	761	11.23	1172	17.30	1024	15.11	2023	29.86	2118	31.26
	2	1264	18.65	1300	19.19	1066	15.73	1174	17.33	1393	20.56	1136	16.77	1082	15.97	564	8.32
	3	1986	29.31	1051	15.51	935	13.80	2298	33.91	617	9.11	1094	16.15	973	14.36	968	14.29
	4	1929	28.47	3007	44.38	1679	24.78	2187	32.28	3234	47.73	2764	40.79	2142	31.61	2468	36.42
IV	0	159	1.46	439	4.04	277	2.55	337	3.10	443	4.07	515	4.73	510	4.69	205	1.88
	1	910	8.37	3493	32.11	3205	29.46	2259	20.77	3549	32.63	2339	21.50	1714	15.76	1608	14.78
	2	1397	12.84	2307	21.21	2649	24.35	3130	28.77	1580	14.52	788	7.24	1560	14.34	1599	14.70
	3	2986	27.45	1754	16.12	1304	11.99	2645	24.32	1640	15.08	1486	13.66	2066	18.99	1316	12.10
	4	5095	46.84	2549	23.43	3099	28.49	2147	19.74	3312	30.45	5406	49.70	4677	43.00	5796	53.28
V	0	360	3.62	166	1.67	241	2.43	362	3.64	450	4.53	472	4.75	168	1.69	162	1.63
	1	4054	40.81	1219	12.27	1667	16.78	2677	26.95	1935	19.48	1214	12.22	935	9.41	1205	12.13
	2	1667	16.78	1656	16.67	2275	22.90	3095	31.16	785	7.90	1259	12.67	1342	13.51	1167	11.75
	3	1554	15.64	2024	20.37	2118	21.32	1803	18.15	1307	13.16	1243	12.51	1105	11.12	943	9.49
	4	1779	17.91	4349	43.78	3070	30.90	1459	14.69	4897	49.30	5220	52.55	5858	58.97	5925	59.64

Table 3.5 Frequency of Operational Item Scores: Mathematics

Math Level	Score/Item	1		2		3		4		5		6		7		8	
		Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent	Count	Percent
I	0	891	9.62	1084	11.71	1255	13.55	2158	23.31	1126	12.16	1200	12.96	1109	11.98	1143	12.34
	1	1265	13.66	1312	14.17	1215	13.12	1866	20.15	1521	16.43	1392	15.03	1376	14.86	1587	17.14
	2	1823	19.69	2155	23.27	1659	17.92	1263	13.64	2617	28.26	1619	17.49	2300	24.84	1683	18.18
	3	781	8.44	930	10.04	740	7.99	558	6.03	855	9.23	781	8.44	805	8.69	860	9.29
	4	920	9.94	814	8.79	903	9.75	687	7.42	865	9.34	778	8.40	833	9.00	897	9.69
	5	3209	34.66	2577	27.83	3082	33.29	2309	24.94	1845	19.93	3096	33.44	2443	26.39	2686	29.01
II	0	130	1.96	255	3.84	191	2.87	230	3.46	234	3.52	245	3.68	315	4.74	257	3.87
	1	1030	15.49	2625	39.48	1091	16.41	1489	22.39	1598	24.03	2497	37.55	1137	17.10	1748	26.29
	2	761	11.45	1094	16.45	767	11.54	866	13.02	989	14.87	1126	16.93	293	4.41	602	9.05
	3	873	13.13	1282	19.28	559	8.41	1131	17.01	1235	18.57	1768	26.59	1653	24.86	746	11.22
	4	3657	55.00	1188	17.87	3827	57.56	2713	40.80	2373	35.69	809	12.17	3040	45.72	3083	46.37
III	0	81	1.20	180	2.66	159	2.35	253	3.73	83	1.22	139	2.05	242	3.57	160	2.36
	1	314	4.63	1946	28.72	1661	24.51	1435	21.18	468	6.91	944	13.93	1161	17.13	1297	19.14
	2	665	9.81	902	13.31	1290	19.04	537	7.93	1383	20.41	755	11.14	449	6.63	496	7.32
	3	601	8.87	1360	20.07	1858	27.42	890	13.13	1074	15.85	786	11.60	683	10.08	694	10.24
	4	4878	71.99	2145	31.66	1572	23.20	3410	50.32	3521	51.96	3916	57.79	4008	59.15	3873	57.16
IV	0	436	4.01	177	1.63	227	2.09	430	3.95	486	4.47	241	2.22	332	3.05	324	2.98
	1	2935	26.98	2141	19.68	2130	19.58	2642	24.29	2192	20.15	3003	27.61	1880	17.28	1460	13.42
	2	1641	15.09	555	5.10	1817	16.70	1438	13.22	1467	13.49	1679	15.43	825	7.58	3254	29.91
	3	1954	17.96	381	3.50	2778	25.54	1603	14.74	3204	29.45	851	7.82	1325	12.18	2219	20.40
	4	3556	32.69	7251	66.66	3557	32.70	4392	40.38	3149	28.95	4746	43.63	6140	56.44	3225	29.65
V	0	287	2.89	237	2.39	220	2.21	403	4.06	309	3.11	435	4.38	423	4.26	394	3.97
	1	2035	20.49	3337	33.59	1616	16.27	1953	19.66	1351	13.60	5389	54.25	2901	29.2	3882	39.08
	2	1190	11.98	1850	18.62	1173	11.81	1812	18.24	2848	28.67	814	8.19	1070	10.77	1425	14.34
	3	1741	17.53	1939	19.52	1639	16.50	2859	28.78	1712	17.23	1339	13.48	1321	13.30	1412	14.21
	4	4148	41.76	2023	20.36	4728	47.59	2338	23.54	3166	31.87	1407	14.16	3662	36.86	2194	22.09

Table 3.6 Raw Score Intercorrelations by Section: Level I

	ELA	Mathematics	Science*
ELA	–	0.88	0.88
Mathematics		–	0.90
Science			–
N	8923	8898	3210
Raw Score Mean	23.18	21.77	22.48
Raw Score SD	10.94	11.69	12.27

*Science raw scores are based on 8 items common across field-test forms.

Table 3.7 Raw Score Intercorrelations by Section: Level II

	ELA	Mathematics
ELA	–	0.87
Mathematics		–
N	6206	6189
Raw Score Mean	22.73	21.18
Raw Score Std	7.11	7.46

Table 3.8 Raw Score Intercorrelations by Section: Level III

	ELA	Mathematics	Science*
ELA	–	0.84	0.81
Mathematics		–	0.82
Science			–
N	6269	6259	3415
Raw Score Mean	20.85	23.68	23.45
Raw Score Std	7.82	7.57	6.34

*Science raw scores are based on 8 items common across field-test forms.

Table 3.9 Raw Score Intercorrelations by Section: Level IV

	ELA	Mathematics	Science*
ELA	–	0.83	0.80
Mathematics		–	0.86
Science			–
N	10353	10337	3850
Raw Score Mean	21.12	21.91	21.34
Raw Score Std	7.74	7.74	6.74

* Science raw scores are based on 8 items common across field-test forms.

Table 3.10 Raw Score Intercorrelations by Section: Level V

	ELA	Mathematics	Science*
ELA	–	0.85	0.84
Mathematics		–	0.84
Science			–
N	9367	9347	3868
Raw Score Mean	22.04	19.48	22.91
Raw Score Std	7.62	7.82	7.32

*Science raw scores are based on 8 items common across field-test forms.

Table 3.11 Ethnic Group DIF Statistics for C Category Items

Content Area	Item No.	Level	Item#	Version	SMD	Comparison	Disadvantaged
ELA	VB543103	2	7	OP	-0.294	White/Asian	Asian
	VB487173	3	6	OP	-0.38	White/Asian	Asian
	VB487173	3	6	OP	-0.369	White/Filipino	Filipino
	VB397806	4	7	OP	-0.449	White/Asian	Asian
	VB397806	4	7	OP	-0.365	White/CombAsian	CombAsian
Mathematics	–	–	–	–	–	–	–
Science	VB541832	5	30	01	0.372	White/Black	White

Table 3.12 Disability Group DIF Statistics for C Category Items

Disability Group DIF Statistics for C Category Items							
Content Area	**Item No.**	**Level**	**Item#**	**Version**	**SMD**	**Comparison**	**Disadvantaged**
ELA OP Items	VB397808	1	5	OP	0.627	MR/VI	MR
	VB397808	1	5	OP	0.596	MR/OI	MR
	VB486738	2	5	OP	-0.446	MR/OI	OI
	VB540182	2	1	OP	-0.384	MR/AU	AU
	VB541371	2	2	OP	0.477	MR/AU	MR
	VB487040	2	4	OP	0.399	MR/AU	MR
	VB486738	2	5	OP	0.344	MR/AU	MR
	VB543105	2	6	OP	-0.283	MR/AU	AU
	VB540186	3	1	OP	-0.340	MR/AU	AU
	VB487040	3	2	OP	0.341	MR/AU	MR
	VB541371	3	5	OP	0.473	MR/AU	MR
	VB487173	3	6	OP	-0.684	MR/AU	AU
	VB539450	3	7	OP	0.481	MR/AU	MR
	VB539454	4	2	OP	0.488	MR/AU	MR
	VB543121	4	5	OP	0.490	MR/AU	MR
	VB397806	4	7	OP	-0.669	MR/AU	AU

Disability Group DIF Statistics for C Category Items							
Content Area	**Item No.**	**Level**	**Item#**	**Version**	**SMD**	**Comparison**	**Disadvantaged**
	VB540151	5	2	OP	0.394	MR/AU	MR
	VB411735	5	6	OP	-0.376	MR/AU	AU
ELA FT Items	VB487174	2	10	2	-0.444	MR/AU	AU
	VB539173	3	9	3	-0.789	MR/AU	AU
	VB540979	4	9	01	-0.265	MR/AU	AU
	VB541379	4	9	02	-0.424	MR/AU	AU
	VB583594	4	10	02	-0.369	MR/AU	AU
	VB487081	4	10	03,05	-0.419	MR/AU	AU
	VB583588	4	10	04	-0.462	MR/AU	AU
	VB539170	4	9,10	05,06	-0.348	MR/AU	AU
	VB540978	5	10	03	-0.267	MR/AU	AU
Math OP Items	VB487027	2	12	OP	0.418	MR/AU	MR
	VB616067	2	13	OP	0.344	MR/AU	MR
	VB541178	4	18	OP	0.298	MR/SL	MR
	VB540414	5	17	OP	-0.333	MR/ED	ED
Math FT Items	VB616069	2	20	3	0.342	MR/AU	MR
	VB616070	3	19	4	-0.35	MR/AU	AU
	VB615612	4	19	05	0.307	MR/SL	MR
	VB542551	4	20	02	0.528	MR/AU	MR
	VB541854	4	19	03,04	0.420	MR/AU	MR
Science OP Items	VB541832	4	24	op	-0.264	MR/SL	SL
	VB555478	5	25	OP	-0.337	MR/AU	AU
Science FT Items	VB541840	5	29	01,02,05	-0.322	MR/AU	AU

Table 3.13 Double Rater Summary for Operational Items: Level I

Level I		First Rating			Second Rating			% Agreement			MAD*	Corr**
Subject	Item	N	Mean	SD	N	Mean	SD	Exact	Adjacent	Neither		
ELA	1	824	3.20	1.71	824	3.18	1.72	89.06	7.41	3.53	0.17	0.94
	2	824	3.43	1.71	824	3.39	1.72	89.56	6.80	3.64	0.17	0.94
	3	824	2.69	1.76	824	2.65	1.74	87.24	9.23	3.53	0.19	0.94
	4	824	2.44	1.75	824	2.42	1.75	87.24	8.63	4.13	0.19	0.93
	5	824	3.87	1.63	824	3.84	1.64	88.96	7.40	3.64	0.19	0.91
	6	824	3.03	1.79	824	2.98	1.81	89.79	6.32	3.89	0.18	0.93
	7	824	2.42	1.69	824	2.43	1.70	88.47	8.13	3.40	0.17	0.94
	8	824	3.47	1.78	824	3.42	1.78	87.24	8.63	4.13	0.20	0.93
Mathematics	11	804	3.18	1.73	804	3.15	1.75	88.18	7.84	3.98	0.20	0.92
	12	804	3.03	1.73	804	2.99	1.73	88.68	7.84	3.48	0.18	0.93
	13	804	3.19	1.78	804	3.13	1.80	88.31	8.08	3.61	0.19	0.93
	14	804	2.49	1.94	804	2.45	1.92	86.44	9.33	4.23	0.22	0.93
	15	804	2.63	1.64	804	2.60	1.64	88.31	7.46	4.23	0.19	0.92
	16	804	3.25	1.78	804	3.24	1.79	89.05	7.34	3.61	0.19	0.93
	17	804	2.78	1.70	804	2.74	1.70	87.80	9.09	3.11	0.18	0.94
	18	804	2.99	1.78	804	2.98	1.80	86.82	10.07	3.11	0.20	0.93

* Mean absolute difference between first and second ratings, ** Pearson correlation between first and second ratings

Table 3.14 Double Rater Summary for Operational Items: Level II

Level II		First Rating			Second Rating			% Agreement			MAD*	Corr**
Subject	Item	N	Mean	SD	N	Mean	SD	Exact	Adjacent	Neither		
ELA	1	693	2.67	1.12	693	2.69	1.15	93.65	5.63	0.72	0.07	0.96
	2	693	2.54	1.24	693	2.54	1.25	96.10	3.32	0.58	0.05	0.98
	3	693	3.17	1.06	693	3.15	1.10	94.81	3.90	1.29	0.07	0.95
	4	693	2.48	1.26	693	2.49	1.26	95.24	3.46	1.30	0.07	0.96
	5	693	3.05	1.23	693	3.04	1.25	93.94	4.76	1.30	0.09	0.94
	6	693	3.29	1.01	693	3.28	1.04	95.96	2.89	1.15	0.06	0.94
	7	693	2.59	1.15	693	2.57	1.19	92.94	6.65	0.41	0.10	0.94
	8	693	2.98	1.25	693	2.99	1.26	95.96	3.03	1.01	0.06	0.96
Mathematics	11	685	3.08	1.20	685	3.08	1.21	96.64	2.63	0.73	0.04	0.98
	12	685	2.03	1.21	685	2.04	1.22	95.33	3.80	0.87	0.06	0.97
	13	685	3.13	1.25	685	3.13	1.24	96.63	2.20	1.17	0.07	0.95
	14	685	2.73	1.29	685	2.73	1.31	94.01	4.53	1.46	0.08	0.95
	15	685	2.64	1.26	685	2.63	1.28	92.41	5.69	1.90	0.11	0.94
	16	685	2.03	1.12	685	2.03	1.13	94.88	3.51	1.61	0.08	0.94
	17	685	3.01	1.26	685	3.00	1.27	97.08	1.75	1.17	0.05	0.97
	18	685	2.65	1.40	685	2.64	1.40	95.77	2.48	1.75	0.08	0.95

* Mean absolute difference between first and second ratings, ** Pearson correlation between first and second ratings

Table 3.15 Double Rater Summary for Operational Items: Level III

Level III		First Rating			Second Rating			% Agreement			MAD*	Corr**
Subject	Item	N	Mean	SD	N	Mean	SD	Exact	Adjacent	Neither		
ELA	1	721	2.65	1.16	721	2.64	1.16	93.62	5.55	0.83	0.08	0.96
	2	721	2.99	1.18	721	2.98	1.19	97.09	2.64	0.27	0.03	0.98
	3	721	2.08	1.25	721	2.07	1.24	94.59	4.16	1.25	0.08	0.95
	4	721	2.93	1.01	721	2.92	1.02	95.69	3.47	0.84	0.06	0.94
	5	721	2.97	1.23	721	2.96	1.25	96.81	2.36	0.83	0.05	0.97
	6	721	2.76	1.35	721	2.75	1.36	93.34	5.41	1.25	0.08	0.97
	7	721	2.46	1.32	721	2.48	1.33	95.15	4.02	0.83	0.06	0.97
	8	721	2.52	1.40	721	2.49	1.41	94.17	4.58	1.25	0.08	0.96
Mathematics	11	719	3.62	0.84	719	3.60	0.87	98.47	0.97	0.56	0.03	0.94
	12	719	2.56	1.26	719	2.55	1.27	96.52	2.92	0.56	0.04	0.98
	13	719	2.54	1.11	719	2.53	1.12	95.69	3.20	1.11	0.06	0.96
	14	719	2.85	1.34	719	2.84	1.34	97.08	2.23	0.69	0.04	0.97
	15	719	3.26	0.99	719	3.23	1.02	97.08	2.09	0.83	0.04	0.96
	16	719	3.27	1.09	719	3.26	1.11	98.05	1.53	0.42	0.03	0.98
	17	719	3.15	1.28	719	3.15	1.29	97.08	2.23	0.69	0.05	0.97
	18	719	3.08	1.27	719	3.07	1.28	97.64	1.81	0.55	0.03	0.98

* Mean absolute difference between first and second ratings, ** Pearson correlation between first and second ratings

Table 3.16 Double Rater Summary for Operational Items: Level IV

Level IV		First Rating			Second Rating			% Agreement			MAD*	Corr**
Subject	Item	N	Mean	SD	N	Mean	SD	Exact	Adjacent	Neither		
ELA	1	939	3.15	1.00	939	3.16	1.00	95.74	3.09	1.17	0.06	0.95
	2	939	2.20	1.24	939	2.23	1.23	93.40	5.64	0.96	0.08	0.96
	3	939	2.33	1.25	939	2.33	1.25	92.97	5.64	1.39	0.09	0.96
	4	939	2.31	1.07	939	2.32	1.06	92.74	6.41	0.85	0.09	0.95
	5	939	2.34	1.34	939	2.34	1.33	94.99	3.94	1.07	0.07	0.97
	6	939	2.88	1.37	939	2.89	1.36	93.40	5.43	1.17	0.08	0.97
	7	939	2.92	1.22	939	2.94	1.21	91.40	7.57	1.03	0.11	0.94
	8	939	3.08	1.21	939	3.09	1.19	95.94	2.77	1.29	0.07	0.95
Mathematics	11	947	2.45	1.28	947	2.46	1.29	92.61	5.60	1.79	0.10	0.95
	12	947	3.17	1.31	947	3.18	1.29	95.24	1.69	3.07	0.11	0.91
	13	947	2.67	1.13	947	2.67	1.14	93.23	5.81	0.96	0.09	0.95
	14	947	2.65	1.33	947	2.67	1.33	95.66	3.17	1.17	0.07	0.95
	15	947	2.62	1.21	947	2.60	1.21	93.76	4.86	1.38	0.09	0.95
	16	947	2.49	1.37	947	2.49	1.38	93.45	4.86	1.69	0.09	0.95
	17	947	3.03	1.26	947	3.05	1.24	93.88	4.54	1.58	0.08	0.95
	18	947	2.66	1.08	947	2.66	1.09	93.34	5.60	1.06	0.09	0.94

* Mean absolute difference between first and second ratings, ** Pearson correlation between first and second ratings

Table 3.17 Double Rater Summary for Operational Items: Level V

Level V		First Rating			Second Rating			% Agreement			MAD*	Corr**
Subject	Item	N	Mean	SD	N	Mean	SD	Exact	Adjacent	Neither		
ELA	1	860	1.89	1.15	860	1.88	1.16	92.78	6.52	0.70	0.08	0.96
	2	860	3.03	1.09	860	3.01	1.11	95.23	4.19	0.58	0.05	0.97
	3	860	2.62	1.16	860	2.63	1.16	92.78	6.17	1.05	0.09	0.95
	4	860	2.10	1.10	860	2.09	1.12	90.57	8.50	0.93	0.11	0.94
	5	860	2.87	1.35	860	2.86	1.35	91.16	7.56	1.28	0.11	0.96
	6	860	3.06	1.29	860	3.04	1.31	93.02	5.12	1.86	0.10	0.95
	7	860	3.24	1.13	860	3.21	1.16	96.39	2.79	0.82	0.05	0.96
	8	860	3.24	1.13	860	3.21	1.16	96.39	2.56	1.05	0.06	0.95
Mathematics	11	846	2.76	1.29	846	2.75	1.30	93.62	4.85	1.53	0.08	0.96
	12	846	2.24	1.19	846	2.25	1.18	93.62	5.08	1.30	0.08	0.96
	13	846	2.92	1.22	846	2.92	1.22	93.50	5.44	1.53	0.08	0.96
	14	846	2.63	1.15	846	2.60	1.16	92.32	6.38	1.30	0.10	0.93
	15	846	2.65	1.17	846	2.63	1.18	94.68	3.90	1.06	0.07	0.95
	16	846	1.66	1.15	846	1.65	1.16	94.80	4.14	1.30	0.07	0.96
	17	846	2.49	1.36	846	2.48	1.38	95.39	3.66	1.42	0.06	0.97
	18	846	2.14	1.30	846	2.13	1.30	93.50	5.32	1.18	0.08	0.96

* Mean absolute difference between first and second ratings, ** Pearson correlation between first and second ratings

Table 3.18 Reliability of Classification and Decision Accuracy: Level I English–Language Arts

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	26–40	0.41	0.05	0.00	0.00	0.00	0.47
	15–25	0.04	0.22	0.03	0.01	0.00	0.30
	12–14	0.00	0.02	0.02	0.01	0.00	0.06
	9–11	0.00	0.01	0.01	0.02	0.01	0.05
	0–8	0.00	0.00	0.01	0.02	0.09	0.13
	Estimated Proportion Correctly Classified : Total = 0.76, Proficient & Above = 0.93						
Decision Consistency Alternate Form	26–40	0.40	0.07	0.00	0.00	0.00	0.47
	15–25	0.06	0.18	0.03	0.02	0.01	0.30
	12–14	0.00	0.02	0.01	0.01	0.01	0.06
	9–11	0.00	0.01	0.01	0.01	0.02	0.05
	0–8	0.00	0.01	0.01	0.02	0.09	0.13
	Estimated Proportion Correctly Classified : Total =0.69, Proficient & Above = 0.90						

*Inconsistencies with category cell entries are due to rounding.

Table 3.19 Reliability of Classification and Decision Accuracy: Level I Mathematics

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	34–40	0.16	0.03	0.00	0	0	0.2
	25–33	0.03	0.18	0.04	0	0	0.26
	19–24	0	0.04	0.08	0.04	0	0.17
	8–18	0	0	0.03	0.18	0.02	0.23
	0–7	0	0	0	0.03	0.12	0.15
	Estimated Proportion Correctly Classified : Total = 0.72, Proficient & Above = 0.92						
Decision Consistency Alternate Form	34–40	0.16	0.04	0	0	0	0.2
	25–33	0.05	0.15	0.05	0.01	0	0.26
	19–24	0	0.05	0.06	0.05	0	0.17
	8–18	0	0	0.04	0.16	0.03	0.23
	0–7	0	0	0	0.04	0.11	0.15
	Estimated Proportion Correctly Classified : Total = 0.64, Proficient & Above = 0.89						

*Inconsistencies with category cell entries are due to rounding.

Table 3.20 Reliability of Classification and Decision Accuracy: Level II English–Language Arts

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	28–32	0.25	0.06	0.00	0.00	0.00	0.31
	21–27	0.05	0.26	0.05	0.00	0.00	0.36
	13–20	0.00	0.05	0.17	0.02	0.00	0.23
	8–12	0.00	0.00	0.02	0.04	0.00	0.06
	0–7	0.00	0.00	0.00	0.01	0.02	0.04
	Estimated Proportion Correctly Classified : Total = 0.74 , Proficient & Above = 0.90						
Decision Consistency Alternate Form	28–32	0.24	0.07	0.00	0.00	0.00	0.31
	21–27	0.07	0.22	0.06	0.00	0.00	0.36
	13–20	0.00	0.06	0.14	0.03	0.00	0.23
	8–12	0.00	0.00	0.02	0.03	0.01	0.06
	0–7	0.00	0.00	0.00	0.01	0.02	0.04
	Estimated Proportion Correctly Classified : Total =0.65, Proficient & Above = 0.88						

*Inconsistencies with category cell entries are due to rounding.

Table 3.21 Reliability of Classification and Decision Accuracy: Level II Mathematics

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	25–32	0.34	0.06	0.00	0.00	0.00	0.40
	16–24	0.05	0.28	0.03	0.00	0.00	0.36
	10–15	0.00	0.05	0.09	0.01	0.00	0.16
	8–9	0.00	0.00	0.02	0.01	0.01	0.04
	0–7	0.00	0.00	0.00	0.01	0.03	0.04
	Estimated Proportion Correctly Classified : Total = 0.75, Proficient & Above = 0.92						
Decision Consistency Alternate Form	25–32	0.32	0.08	0.00	0.00	0.00	0.40
	16–24	0.07	0.24	0.05	0.00	0.00	0.36
	10–15	0.00	0.05	0.07	0.02	0.01	0.16
	8–9	0.00	0.00	0.02	0.01	0.01	0.04
	0–7	0.00	0.00	0.01	0.01	0.03	0.04
	Estimated Proportion Correctly Classified : Total =0.67, Proficient & Above =0.90						

*Inconsistencies with category cell entries are due to rounding.

Table 3.22 Reliability of Classification and Decision Accuracy: Level III English–Language Arts

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	26–32	0.30	0.05	0.00	0.00	0.00	0.36
	19–25	0.03	0.19	0.03	0.00	0.00	0.25
	14–18	0.01	0.05	0.10	0.03	0.00	0.19
	9–13	0.00	0.00	0.04	0.08	0.02	0.14
	0–8	0.00	0.00	0.00	0.02	0.04	0.06
	Estimated Proportion Correctly Classified : Total = 0.71, Proficient & Above = 0.91						
Decision Consistency Alternate Form	26–32	0.29	0.06	0.01	0.00	0.00	0.36
	19–25	0.05	0.15	0.04	0.00	0.00	0.25
	14–18	0.01	0.06	0.08	0.04	0.01	0.19
	9–13	0.00	0.01	0.04	0.06	0.03	0.14
	0–8	0.00	0.00	0.00	0.02	0.04	0.06
	Estimated Proportion Correctly Classified : Total = 0.62, Proficient & Above = 0.87						

*Inconsistencies with category cell entries are due to rounding.

Table 3.23 Reliability of Classification and Decision Accuracy: Level III Mathematics

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	28–32	0.36	0.06	0.01	0.00	0.00	0.43
	20–27	0.03	0.24	0.02	0.00	0.00	0.29
	15–19	0.00	0.04	0.07	0.02	0.00	0.12
	10–14	0.00	0.01	0.03	0.05	0.01	0.11
	0–9	0.00	0.00	0.00	0.01	0.04	0.05
	Estimated Proportion Correctly Classified : Total =0.76 , Proficient & Above = 0.92						
Decision Consistency Alternate Form	28–32	0.35	0.07	0.01	0.00	0.00	0.43
	20–27	0.06	0.20	0.04	0.00	0.00	0.29
	15–19	0.00	0.04	0.05	0.03	0.00	0.12
	10–14	0.01	0.01	0.03	0.04	0.02	0.11
	0–9	0.00	0.00	0.00	0.01	0.04	0.05
	Estimated Proportion Correctly Classified : Total = 0.68, Proficient & Above = 0.89						

*Inconsistencies with category cell entries are due to rounding.

Table 3.24 Reliability of Classification and Decision Accuracy: Level IV English–Language Arts

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	26–32	0.31	0.05	0.00	0.00	0.00	0.36
	21–25	0.04	0.13	0.04	0.00	0.00	0.21
	16–20	0.00	0.04	0.10	0.03	0.00	0.17
	12–15	0.00	0.01	0.04	0.06	0.02	0.13
	0–11	0.00	0.00	0.00	0.03	0.11	0.13
	Estimated Proportion Correctly Classified : Total = 0.71, Proficient & Above = 0.91						
Decision Consistency Alternate Form	26–32	0.30	0.05	0.01	0.00	0.00	0.36
	21–25	0.05	0.10	0.05	0.01	0.00	0.21
	16–20	0.00	0.05	0.08	0.03	0.01	0.17
	12–15	0.00	0.01	0.04	0.05	0.03	0.13
	0–11	0.00	0.00	0.00	0.03	0.10	0.13
	Estimated Proportion Correctly Classified : Total = 0.63, Proficient & Above = 0.87						

*Inconsistencies with category cell entries are due to rounding.

Table 3.25 Reliability of Classification and Decision Accuracy: Level IV Mathematics

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	29–32	0.19	0.06	0.00	0.00	0.00	0.25
	24–28	0.03	0.16	0.05	0.00	0.00	0.25
	19–23	0.00	0.04	0.11	0.03	0.00	0.18
	14–18	0.00	0.00	0.04	0.08	0.02	0.14
	0–13	0.00	0.00	0.00	0.03	0.13	0.17
	Estimated Proportion Correctly Classified : Total =0.67, Proficient & Above = 0.91						
Decision Consistency Alternate Form	29–32	0.18	0.06	0.01	0.00	0.00	0.25
	24–28	0.06	0.13	0.05	0.01	0.00	0.25
	19–23	0.00	0.05	0.08	0.04	0.01	0.18
	14–18	0.00	0.01	0.04	0.06	0.03	0.14
	0–13	0.00	0.00	0.01	0.04	0.13	0.17
	Estimated Proportion Correctly Classified : Total = 0.58, Proficient & Above = 0.87						

*Inconsistencies with category cell entries are due to rounding.

Table 3.26 Reliability of Classification and Decision Accuracy: Level V English–Language Arts

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	27–32	0.31	0.05	0.00	0.00	0.00	0.36
	21–26	0.04	0.19	0.04	0.00	0.00	0.28
	16–20	0.00	0.03	0.09	0.02	0.00	0.14
	12–15	0.00	0.00	0.03	0.05	0.01	0.10
	0–11	0.00	0.00	0.01	0.03	0.09	0.13
	Estimated Proportion Correctly Classified : Total = 0.73 , Proficient & Above = 0.93						
Decision Consistency Alternate Form	27–32	0.29	0.06	0.00	0.00	0.00	0.36
	21–26	0.06	0.16	0.05	0.01	0.00	0.28
	16–20	0.00	0.04	0.07	0.03	0.00	0.14
	12–15	0.00	0.01	0.03	0.04	0.02	0.10
	0–11	0.00	0.00	0.01	0.02	0.09	0.13
	Estimated Proportion Correctly Classified : Total = 0.65 , Proficient & Above = 0.89						

*Inconsistencies with category cell entries are due to rounding.

Table 3.27 Reliability of Classification and Decision Accuracy: Level V Mathematics

	Placement Score	Advanced	Proficient	Basic	Below Basic	Far Below Basic	Category Total*
Decision Accuracy All-forms Average	26–32	0.23	0.04	0.00	0.00	0.00	0.27
	20–25	0.03	0.17	0.04	0.00	0.00	0.25
	15–19	0.00	0.04	0.11	0.02	0.01	0.19
	13–14	0.00	0.00	0.03	0.02	0.02	0.07
	0–12	0.00	0.00	0.02	0.03	0.18	0.23
	Estimated Proportion Correctly Classified : Total = 0.71 , Proficient & Above = 0.92						
Decision Consistency Alternate Form	26–32	0.22	0.05	0.00	0.00	0.00	0.27
	20–25	0.05	0.14	0.05	0.00	0.00	0.25
	15–19	0.00	0.05	0.09	0.02	0.02	0.19
	13–14	0.00	0.01	0.02	0.01	0.02	0.07
	0–12	0.00	0.00	0.02	0.02	0.17	0.23
	Estimated Proportion Correctly Classified : Total = 0.63 , Proficient & Above = 0.89						

*Inconsistencies with category cell entries are due to rounding.

Table 3.28 CAPA 2006 Raw Score Means and Standard Deviations: Equating Sample vs. Total

Group	Level	ELA				Mathematics			
		N	% Total N	Mean RS	SD RS	N	% Total N	Mean RS	SD RS
Total	I	8923		23.18	10.94	8898		21.77	11.69
	II	6206		22.73	7.11	6189		21.18	7.46
	III	6269		20.85	7.82	6259		23.68	7.57
	IV	10353		21.12	7.74	10337		21.91	7.74
	V	9367		22.04	7.62	9347		19.48	7.82
Equating Sample	I	1527	17%	22.02	10.75	1521	17%	20.29	11.45
	II	1493	24%	22.23	7.54	1487	24%	20.73	7.81
	III	1569	25%	20.49	8.04	1566	25%	23.30	7.82
	IV	1876	18%	20.20	7.68	1869	18%	21.19	7.79
	V	2293	24%	21.64	7.73	2289	24%	19.18	7.85

Table 3.29 Evaluation of Common Items between New and Reference Test Forms

Subject	Level	N Common Items	N Items Removed	Common Item Correlation	WRMSD*
ELA	I	5	0	0.99	0.10
	II	5	0	0.99	0.08
	III	5	0	0.99	0.04
	IV	5	0	0.99	0.03
	V	5	0	0.99	0.05
Mathematics	I	5	0	0.98	0.10
	II	5	0	0.96	0.11
	III	5	0	0.97	0.10
	IV	5	0	0.99	0.04
	V	5	0	0.99	0.06

*Weighted root mean square difference

Table 3.30 Score Conversions: ELA Level I

Raw Score	Scale Score	CSEM*	Performance Level
40	60	13.8	Advanced
39	60	13.8	
38	60	9.2	
37	60	7.4	
36	60	6.4	
35	59	5.8	
34	57	5.4	
33	55	5.0	
32	54	4.8	
31	52	4.6	
30	51	4.0	
29	50	4.4	
28	49	4.3	
27	48	4.2	
26	46	4.2	
25	45	4.1	Proficient
24	44	4.1	
23	43	4.1	
22	42	4.1	
21	41	4.1	
20	40	4.1	
19	39	4.1	
18	38	4.2	
17	37	4.3	
16	36	4.3	
15	35	4.5	
14	33	4.6	Basic
13	32	4.8	
12	31	5.0	
11	29	5.3	Below Basic
10	27	5.6	
9	26	5.9	
8	25	6.3	Far Below Basic
7	24	6.8	
6	23	7.3	
5	21	8.0	
4	19	8.8	
3	17	10.1	
2	15	12.1	
1	15	16.7	
0	15	16.7	

*Conditional standard error of measurement

Table 3.31 Score Conversions: Mathematics Level I

Raw Score	Scale Score	CSEM	Performance Level
40	60	15.4	Advanced
39	57	12.4	
38	51	7.0	
37	48	5.6	
36	46	4.8	
35	44	4.3	
34	43	3.9	
33	41	3.7	Proficient
32	40	3.5	
31	39	3.4	
30	39	3.3	
29	38	3.2	
28	37	3.1	
27	36	3.1	
26	36	3.0	
25	35	3.0	
24	34	3.0	Basic
23	33	3.0	
22	33	3.1	
21	32	3.1	
20	31	3.2	
19	30	3.3	
18	29	3.3	Below Basic
17	28	3.5	
16	27	3.6	
15	26	3.8	
14	25	4.0	
13	24	4.2	
12	23	4.5	
11	23	4.8	
10	22	5.1	
9	22	5.4	
8	21	5.7	
7	20	5.9	Far Below Basic
6	19	6.3	
5	19	6.7	
4	18	7.2	
3	16	8.1	
2	15	9.6	
1	15	13.1	
0	15	13.1	

Table 3.32 Score Conversions: ELA Level II

Raw Score	Scale Score	CSEM	Performance Level
32	60	17.8	**Advanced**
31	47	4.8	
30	44	3.4	
29	42	2.8	
28	41	2.5	
27	40	2.3	**Proficient**
26	39	2.2	
25	38	2.1	
24	37	2.0	
23	36	1.9	
22	36	1.9	
21	35	1.9	
20	34	1.8	**Basic**
19	34	1.8	
18	33	1.8	
17	32	1.8	
16	32	1.8	
15	31	1.9	
14	30	1.9	
13	30	2.0	
12	29	2.1	**Below Basic**
11	28	2.2	
10	27	2.2	
9	26	2.4	
8	25	2.6	
7	23	2.7	**Far Below Basic**
6	22	2.8	
5	20	3.0	
4	18	3.2	
3	16	3.6	
2	15	4.1	
1	15	5.5	
0	15	5.5	

Table 3.33 Score Conversions: Mathematics Level II

Raw Score	Scale Score	CSEM	Performance Level
32	60	11.5	Advanced
31	55	6.5	
30	49	4.5	
29	46	3.6	
28	45	3.1	
27	43	2.8	
26	42	2.6	
25	41	2.4	
24	40	2.2	Proficient
23	39	2.1	
22	39	2.1	
21	38	2.0	
20	37	2.0	
19	37	1.0	
18	36	2.0	
17	36	2.0	
16	35	2.0	
15	34	2.0	Basic
14	34	2.1	
13	33	2.2	
12	32	2.3	
11	31	2.5	
10	30	2.8	
9	29	3.1	Below Basic
8	27	3.4	
7	25	3.8	Far Below Basic
6	23	4.0	
5	20	4.1	
4	17	4.3	
3	15	4.5	
2	15	5.2	
1	15	6.8	
0	15	6.8	

Table 3.34 Score Conversions: ELA Level III

Raw Score	Scale Score	CSEM	Performance Level
32	60	11.5	**Advanced**
31	56	7.5	
30	49	5.1	
29	46	4.1	
28	44	3.6	
27	43	3.2	
26	41	3.0	
25	40	2.9	**Proficient**
24	39	2.7	
23	38	2.7	
22	37	2.6	
21	37	2.6	
20	36	2.6	
19	35	2.6	
18	34	2.6	**Basic**
17	33	2.7	
16	32	2.7	
15	31	2.8	
14	30	2.9	
13	29	3.2	**Below Basic**
12	28	3.3	
11	26	3.5	
10	25	3.7	
9	23	4.0	
8	22	4.2	**Far Below Basic**
7	21	4.5	
6	21	4.7	
5	20	4.9	
4	19	5.2	
3	18	5.6	
2	16	6.4	
1	15	8.4	
0	15	8.4	

Table 3.35 Score Conversions: Mathematics Level III

Raw Score	Scale Score	CSEM	Performance Level
32	60	12	Advanced
31	56	8.0	
30	49	3.5	
29	46	4.5	
28	44	3.8	
27	42	3.4	Proficient
26	41	3.2	
25	40	3.0	
24	38	2.8	
23	37	2.7	
22	37	2.7	
21	36	2.6	
20	35	2.6	
19	34	2.6	Basic
18	33	2.6	
17	32	2.7	
16	31	2.8	
15	30	2.9	
14	29	3.0	Below Basic
13	28	3.2	
12	27	3.4	
11	26	3.7	
10	25	4.0	
9	24	4.3	Far Below Basic
8	23	4.5	
7	22	4.7	
6	21	4.8	
5	20	4.9	
4	19	5.1	
3	17	5.5	
2	16	6.4	
1	15	8.5	
0	15	8.5	

Table 3.36 Score Conversions: ELA Level IV

Raw Score	Scale Score	CSEM	Performance Level
32	60	13.1	Advanced
31	54	7.1	
30	48	5.0	
29	45	4.1	
28	43	3.5	
27	42	3.2	
26	41	3.0	
25	39	2.9	Proficient
24	38	2.8	
23	37	2.7	
22	36	2.7	
21	35	2.6	
20	34	2.6	Basic
19	33	2.6	
18	32	2.6	
17	31	2.7	
16	30	2.7	
15	29	2.8	Below Basic
14	28	2.9	
13	27	3.1	
12	26	3.3	
11	24	3.5	Far Below Basic
10	22	3.8	
9	21	4.1	
8	20	4.4	
7	19	4.7	
6	19	4.8	
5	18	4.9	
4	17	5.0	
3	16	5.3	
2	15	6.0	
1	15	7.9	
0	15	7.9	

Table 3.37 Score Conversions: Mathematics Level IV

Raw Score	Scale Score	CSEM	Performance Level
32	60	16.3	**Advanced**
31	52	8.3	
30	46	5.8	
29	43	4.6	
28	40	4.0	**Proficient**
27	39	3.6	
26	37	3.3	
25	36	3.1	
24	35	3.0	
23	34	2.9	**Basic**
22	33	2.8	
21	32	2.8	
20	31	2.7	
19	30	2.7	
18	29	2.7	**Below Basic**
17	28	2.8	
16	28	2.8	
15	27	2.9	
14	26	3.0	
13	24	3.2	**Far Below Basic**
12	23	3.4	
11	22	3.7	
10	20	4.1	
9	20	4.7	
8	19	5.2	
7	19	5.5	
6	18	5.6	
5	17	5.7	
4	17	5.8	
3	16	6.2	
2	15	7.0	
1	15	9.2	
0	15	9.2	

Table 3.38 Score Conversions: ELA Level V

Raw Score	Scale Score	CSEM	Performance Level
32	60	9.9	Advanced
31	57	6.9	
30	50	5.0	
29	47	4.2	
28	45	3.8	
27	43	3.5	
26	41	3.3	Proficient
25	40	3.1	
24	38	3.0	
23	37	2.8	
22	36	2.7	
21	35	2.6	
20	34	2.6	Basic
19	33	2.6	
18	32	2.5	
17	31	2.6	
16	30	2.6	
15	29	2.6	Below Basic
14	28	2.7	
13	27	2.9	
12	26	3.0	
11	24	3.2	Far Below Basic
10	23	3.5	
9	23	3.7	
8	22	4.0	
7	21	4.2	
6	21	4.3	
5	20	4.5	
4	19	4.7	
3	18	5.0	
2	16	5.7	
1	15	7.5	
0	15	7.5	

Table 3.39 Score Conversions: Mathematics Level V

Raw Score	Scale Score	CSEM	Performance Level
32	60	10.3	Advanced
31	57	7.3	
30	49	5.0	
29	46	4.1	
28	44	3.5	
27	42	3.2	
26	41	3.0	
25	40	2.9	Proficient
24	39	2.8	
23	38	2.7	
22	37	2.7	
21	36	2.6	
20	35	2.6	
19	34	2.6	Basic
18	33	2.6	
17	32	2.7	
16	31	2.7	
15	30	2.8	
14	29	3.0	Below Basic
13	27	3.1	
12	26	3.4	Far Below Basic
11	25	3.7	
10	24	4.1	
9	23	4.6	
8	22	5.0	
7	21	5.1	
6	20	5.1	
5	19	5.0	
4	18	5.1	
3	17	5.4	
2	15	6.1	
1	15	8.0	
0	15	8.0	

Table 3.40 Scale Score Frequency Distributions: Level I ELA and Mathematics

Scale Score	ELA				Mathematics			
	Frequency	Percent	Cumulative Frequency	Percent Below	Frequency	Percent	Cumulative Frequency	Percent Below
60	1299	14.56	1299	85.44	376	4.23	376	95.77
57–59	602	6.75	1901	78.70	193	2.17	569	93.61
54–56	504	5.65	2405	73.05				
51–53	582	6.52	2987	66.52	200	2.25	769	91.36
48–50	899	10.08	3886	56.45	278	3.12	1047	88.23
45–47	585	6.56	4471	49.89	203	2.28	1250	85.95
42–44	844	9.46	5315	40.43	504	5.66	1754	80.29
39–41	696	7.80	6011	32.63	965	10.85	2719	69.44
36–38	621	6.96	6632	25.68	1017	11.43	3736	58.01
33–35	358	4.01	6990	21.66	1100	12.36	4836	45.65
30–32	336	3.77	7326	17.90	686	7.71	5522	37.94
27–29	325	3.64	7651	14.26	657	7.38	6179	30.56
24–26	432	4.84	8083	9.41	514	5.78	6693	24.78
21–23	244	2.73	8327	6.68	872	9.80	7565	14.98
18–20	103	1.15	8430	5.53	553	6.21	8118	8.77
15–17	493	5.53	8923	0.00	780	8.77	8898	0.00

Note: Gaps indicate scale scores that were not in the 2006 raw-to-scale conversions.

Table 3.41 Scale Score Frequency Distributions: Level II ELA and Mathematics

Scale Score	ELA				Mathematics			
	Frequency	Percent	Cumulative Frequency	Percent Below	Frequency	Percent	Cumulative Frequency	Percent Below
60	345	5.56	345	94.44	210	3.39	210	96.61
57–59								
54–56					304	4.91	514	91.69
51–53								
48–50					305	4.93	819	86.77
45–47	413	6.65	758	87.79	729	11.78	1548	74.99
42–44	781	12.58	1539	75.20	613	9.90	2161	65.08
39–41	1097	17.68	2636	57.52	1154	18.65	3315	46.44
36–38	1266	20.40	3902	37.13	1172	18.94	4487	27.50
33–35	937	15.10	4839	22.03	796	12.86	5283	14.64
30–32	746	12.02	5585	10.01	422	6.82	5705	7.82
27–29	263	4.24	5848	5.77	222	3.59	5927	4.23
24–26	127	2.05	5975	3.72	58	0.94	5985	3.30
21–23	95	1.53	6070	2.19	33	0.53	6018	2.76
18–20	51	0.82	6121	1.37	35	0.57	6053	2.20
15–17	85	1.37	6206	0.00	136	2.20	6189	0.00

Note: Gaps indicate scale scores that were not in the 2006 raw-to-scale conversions.

Table 3.42 Scale Score Frequency Distributions: Level III ELA and Mathematics

Scale Score	ELA				Mathematics			
	Frequency	Percent	Cumulative Frequency	Percent Below	Frequency	Percent	Cumulative Frequency	Percent Below
60	332	5.30	332	94.70	608	9.71	608	90.29
57–59								
54–56	341	5.44	673	89.26	577	9.22	1185	81.07
51–53								
48–50	328	5.23	1001	84.03	568	9.07	1753	71.99
45–47	345	5.50	1346	78.53	534	8.53	2287	63.46
42–44	603	9.62	1949	68.91	797	12.73	3084	50.73
39–41	754	12.03	2703	56.88	554	8.85	3638	41.88
36–38	913	14.56	3616	42.32	728	11.63	4366	30.24
33–35	653	10.42	4269	31.90	449	7.17	4815	23.07
30–32	714	11.39	4983	20.51	461	7.37	5276	15.71
27–29	418	6.67	5401	13.85	426	6.81	5702	8.90
24–26	371	5.92	5772	7.93	312	4.98	6014	3.91
21–23	322	5.14	6094	2.79	130	2.08	6144	1.84
18–20	90	1.44	6184	1.36	41	0.66	6185	1.18
15–17	85	1.36	6269	0.00	74	1.18	6259	0.00

Note: Gaps indicate scale scores that were not in the 2006 raw-to-scale conversions.

Table 3.43 Scale Score Frequency Distributions: Level IV ELA and Mathematics

Scale Score	ELA				Mathematics			
	Frequency	Percent	Cumulative Frequency	Percent Below	Frequency	Percent	Cumulative Frequency	Percent Below
60	601	5.81	601	94.19	693	6.70	693	93.30
57–59								
54–56	557	5.38	1158	88.81				
51–53					654	6.33	1347	86.97
48–50	581	5.61	1739	83.20				
45–47	491	4.74	2230	78.46	615	5.95	1962	81.02
42–44	1006	9.72	3236	68.74	626	6.06	2588	74.96
39–41	908	8.77	4144	59.97	1132	10.95	3720	64.01
36–38	1309	12.64	5453	47.33	974	9.42	4694	54.59
33–35	1114	10.76	6567	36.57	1265	12.24	5959	42.35
30–32	1036	10.01	7603	26.56	1097	10.61	7056	31.74
27–29	1048	10.12	8651	16.44	1194	11.55	8250	20.19
24–26	649	6.27	9300	10.17	590	5.71	8840	14.48
21–23	461	4.45	9761	5.72	499	4.83	9339	9.65
18–20	353	3.41	10114	2.31	713	6.90	10052	2.76
15–17	239	2.31	10353	0.00	285	2.76	10337	0.00

Note: Gaps indicate scale scores that were not in the 2006 raw-to-scale conversions.

Table 3.44 Scale Score Frequency Distributions: Level V ELA and Mathematics

Scale Score	ELA				Mathematics			
	Frequency	Percent	Cumulative Frequency	Percent Below	Frequency	Percent	Cumulative Frequency	Percent Below
60	397	4.24	397	95.76	303	3.24	303	96.76
57–59	530	5.66	927	90.10	333	3.56	636	93.20
54–56								
51–53								
48–50	575	6.14	1502	83.96	371	3.97	1007	89.23
45–47	1275	13.61	2777	70.35	382	4.09	1389	85.14
42–44	637	6.80	3414	63.55	800	8.56	2189	76.58
39–41	1098	11.72	4512	51.83	1157	12.38	3346	64.20
36–38	1224	13.07	5736	38.76	1136	12.15	4482	52.05
33–35	838	8.95	6574	29.82	1125	12.04	5607	40.01
30–32	710	7.58	7284	22.24	997	10.67	6604	29.35
27–29	661	7.06	7945	15.18	627	6.71	7231	22.64
24–26	454	4.85	8399	10.33	1014	10.85	8245	11.79
21–23	728	7.77	9127	2.56	687	7.35	8932	4.44
18–20	128	1.37	9255	1.20	207	2.21	9139	2.23
15–17	112	1.20	9367	0.00	208	2.23	9347	0.00

Note: Gaps indicate scale scores that were not in the 2006 raw-to-scale conversions.

Table 3.45 Raw Score Frequency Distributions: Level I Science

Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below	Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below
40	229	7.13	229	92.87	17	63	1.96	2150	33.02
39	87	2.71	316	90.16	16	68	2.12	2218	30.90
38	83	2.59	399	87.57	15	55	1.71	2273	29.19
37	112	3.49	511	84.08	14	61	1.90	2334	27.29
36	87	2.71	598	81.37	13	62	1.93	2396	25.36
35	68	2.12	666	79.25	12	54	1.68	2450	23.68
34	117	3.64	783	75.61	11	55	1.71	2505	21.96
33	83	2.59	866	73.02	10	61	1.90	2566	20.06
32	78	2.43	944	70.59	9	52	1.62	2618	18.44
31	105	3.27	1049	67.32	8	109	3.40	2727	15.05
30	64	1.99	1113	65.33	7	46	1.43	2773	13.61
29	86	2.68	1199	62.65	6	51	1.59	2824	12.02
28	94	2.93	1293	59.72	5	57	1.78	2881	10.25
27	92	2.87	1385	56.85	4	46	1.43	2927	8.82
26	76	2.37	1461	54.49	3	48	1.50	2975	7.32
25	99	3.08	1560	51.40	2	45	1.40	3020	5.92
24	79	2.46	1639	48.94	1	31	0.97	3051	4.95
23	86	2.68	1725	46.26	0	159	4.95	3210	0.00
22	82	2.55	1807	43.71					
21	79	2.46	1886	41.25					
20	66	2.06	1952	39.19					
19	55	1.71	2007	37.48					
18	80	2.49	2087	34.98					

* Level I Science raw scores are based on 8 items common across field-test forms.

Table 3.46 Raw Score Frequency Distributions: Level III Science

Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below	Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below
32	195	5.71	195	94.29	12	33	0.97	3238	5.18
31	150	4.39	345	89.90	11	25	0.73	3263	4.45
30	228	6.68	573	83.22	10	28	0.82	3291	3.63
29	232	6.79	805	76.43	9	18	0.53	3309	3.10
28	246	7.20	1051	69.22	8	36	1.05	3345	2.05
27	257	7.53	1308	61.70	7	10	0.29	3355	1.76
26	213	6.24	1521	55.46	6	15	0.44	3370	1.32
25	202	5.92	1723	49.55	5	4	0.12	3374	1.20
24	225	6.59	1948	42.96	4	5	0.15	3379	1.05
23	183	5.36	2131	37.60	3	6	0.18	3385	0.88
22	167	4.89	2298	32.71	2	5	0.15	3390	0.73
21	153	4.48	2451	28.23	1	4	0.12	3394	0.61
20	125	3.66	2576	24.57	0	21	0.61	3415	0.00
19	139	4.07	2715	20.50					
18	127	3.72	2842	16.78					
17	113	3.31	2955	13.47					
16	97	2.84	3052	10.63					
15	67	1.96	3119	8.67					
14	48	1.41	3167	7.26					
13	38	1.11	3205	6.15					

* Level III Science raw scores are based on 8 items common across field-test forms.

Table 3.47 Raw Score Frequency Distributions: Level IV Science

Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below	Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below
32	116	3.01	116	96.99	11	47	1.22	3564	7.43
31	112	2.91	228	94.08	10	64	1.66	3628	5.77
30	136	3.53	364	90.55	9	35	0.91	3663	4.86
29	161	4.18	525	86.36	8	59	1.53	3722	3.32
28	204	5.30	729	81.06	7	26	0.68	3748	2.65
27	240	6.23	969	74.83	6	5	0.13	3753	2.52
26	231	6.00	1200	68.83	5	17	0.44	3770	2.08
25	217	5.64	1417	63.19	4	15	0.39	3785	1.69
24	206	5.35	1623	57.84	3	10	0.26	3795	1.43
23	241	6.26	1864	51.58	2	14	0.36	3809	1.06
22	216	5.61	2080	45.97	1	12	0.31	3821	0.75
21	229	5.95	2309	40.03	0	29	0.75	3850	0.00
20	176	4.57	2485	35.45					
19	185	4.81	2670	30.65					
18	165	4.29	2835	26.36					
17	150	3.90	2985	22.47					
16	148	3.84	3133	18.62					
15	127	3.30	3260	15.32					
14	93	2.42	3353	12.91					
13	81	2.10	3434	10.81					
12	83	2.16	3517	8.65					

* Level IV Science raw scores are based on 8 items common across field-test forms.

Table 3.48 Raw Score Frequency Distributions: Level V Science

Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below	Raw Score	Frequency	Percent	Cumulative Frequency	Percent Below
32	337	8.71	337	91.29	11	54	1.40	3596	7.03
31	249	6.44	586	84.85	10	43	1.11	3639	5.92
30	227	5.87	813	78.98	9	40	1.03	3679	4.89
29	249	6.44	1062	72.54	8	68	1.76	3747	3.13
28	211	5.46	1273	67.09	7	19	0.49	3766	2.64
27	220	5.69	1493	61.40	6	9	0.23	3775	2.40
26	209	5.40	1702	56.00	5	9	0.23	3784	2.17
25	204	5.27	1906	50.72	4	9	0.23	3793	1.94
24	191	4.94	2097	45.79	3	8	0.21	3801	1.73
23	175	4.52	2272	41.26	2	8	0.21	3809	1.53
22	181	4.68	2453	36.58	1	8	0.21	3817	1.32
21	145	3.75	2598	32.83	0	51	1.32	3868	0.00
20	143	3.70	2741	29.14					
19	152	3.93	2893	25.21					
18	112	2.90	3005	22.31					
17	126	3.26	3131	19.05					
16	117	3.02	3248	16.03					
15	106	2.74	3354	13.29					
14	73	1.89	3427	11.40					
13	66	1.71	3493	9.69					
12	49	1.27	3542	8.43					

* Level V Science raw scores are based on 8 items common across field-test forms.

Chapter 4. Statewide Assessment Results

In 2006, a total of 41,118 students in grades 2–11 participated in the CAPA. This chapter provides detailed information on statewide test results. The first section includes information on the number of participants and their disabilities; the second section includes test results in the aggregate and disaggregated by demographic and disability subgroups.

Participation

The number of students with one or more test sections is presented in Table 4.1 by test level. Level IV had the largest sample size with 10,353 examinees, followed by Levels V and I with 9,367 and 8,923 assessed students, respectively.

Table 4.1 Distribution of Students Across Test Levels

Test Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
I	8923	21.7	8923	21.7
II	6206	15.1	15129	36.8
III	6269	15.2	21398	52.0
IV	10353	25.2	31751	77.2
V	9367	22.8	41118	100.0

Table 4.2 provides the CAPA population disability frequencies for students completing a CAPA test in ELA or mathematics. Across all levels, the largest disability group (45.5%) is Mental Retardation, followed by Autism (18.4%) and Orthopedic Impairment (9.8%). The unknown category comprises those examinees for which no disability type was gridded. This category included 2.2% of the examinees. Tables 4.3–4.7 provide parallel information by test level.

Table 4.2 Disability Distributions across All Levels

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	18712	45.5	18681	45.5
Hard of Hearing	313	0.8	312	0.8
Deafness	493	1.2	493	1.2
Speech or Language Impairment	1344	3.3	1342	3.3
Visual Impairment	525	1.3	522	1.3
Emotional Disturbance	373	0.9	372	0.9
Orthopedic Impairment	4016	9.8	4004	9.8
Other Health Impairment	1359	3.3	1358	3.3
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	2702	6.6	2698	6.6
Deaf-Blindness	43	0.1	42	0.1
Multiple Disabilities	2505	6.1	2499	6.1
Autism	7546	18.4	7532	18.4
Traumatic Brain Injury	283	0.7	282	0.7
Unknown	904	2.2	893	2.2
TOTAL	41118	100.0	41030	100.0

Table 4.3 Level I Disability Distributions

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	3047	34.1	3040	34.2
Hard of Hearing	70	0.8	70	0.8
Deafness	41	0.5	41	0.5
Speech or Language Impairment	53	0.6	53	0.6
Visual Impairment	249	2.8	247	2.8
Emotional Disturbance	19	0.2	19	0.2
Orthopedic Impairment	2159	24.2	2153	24.2
Other Health Impairment	193	2.2	193	2.2
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	86	1.0	86	1.0
Deaf-Blindness	28	0.3	27	0.3
Multiple Disabilities	1185	13.3	1182	13.3
Autism	1562	17.5	1560	17.5
Traumatic Brain Injury	63	0.7	63	0.7
Unknown	168	1.9	164	1.8
TOTAL	8923	100.0	8898	100.0

Table 4.4 Level II Disability Distributions

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	2419	39.0	2413	39.0
Hard of Hearing	44	0.7	44	0.7
Deafness	82	1.3	82	1.3
Speech or Language Impairment	525	8.5	525	8.5
Visual Impairment	53	0.9	53	0.9
Emotional Disturbance	33	0.5	33	0.5
Orthopedic Impairment	347	5.6	345	5.6
Other Health Impairment	248	4.0	248	4.0
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	409	6.6	408	6.6
Deaf-Blindness	3	0.0	3	0.0
Multiple Disabilities	234	3.8	233	3.8
Autism	1609	25.9	1606	26.0
Traumatic Brain Injury	33	0.5	33	0.5
Unknown	167	2.7	163	2.6
TOTAL	6206	100.0	6189	100.0

Table 4.5 Level III Disability Distributions

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	2779	44.3	2775	44.4
Hard of Hearing	46	0.7	46	0.7
Deafness	60	1.0	60	1.0
Speech or Language Impairment	323	5.2	323	5.2
Visual Impairment	44	0.7	44	0.7
Emotional Disturbance	56	0.9	56	0.9
Orthopedic Impairment	392	6.3	392	6.3
Other Health Impairment	260	4.1	259	4.1
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	518	8.3	518	8.3
Deaf-Blindness	2	0.0	2	0.0
Multiple Disabilities	266	4.2	264	4.2
Autism	1352	21.6	1349	21.6
Traumatic Brain Injury	37	0.6	37	0.6
Unknown	134	2.1	134	2.1
TOTAL	6269	100.0	6259	100.0

Table 4.6 Level IV Disability Distributions

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	5163	49.9	5154	49.9
Hard of Hearing	75	0.7	74	0.7
Deafness	175	1.7	175	1.7
Speech or Language Impairment	286	2.8	284	2.7
Visual Impairment	88	0.9	87	0.8
Emotional Disturbance	102	1.0	102	1.0
Orthopedic Impairment	621	6.0	620	6.0
Other Health Impairment	359	3.5	359	3.5
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	877	8.5	877	8.5
Deaf-Blindness	8	0.0	8	0.0
Multiple Disabilities	401	3.9	402	3.9
Autism	1879	18.2	1876	18.2
Traumatic Brain Injury	60	0.6	60	0.6
Unknown	259	2.5	259	2.5
TOTAL	10353	100.0	10337	100.0

Table 4.7 Level V Disability Distributions

Disability	ELA		Mathematics	
	Frequency	Percent	Frequency	Percent
Mental Retardation	5304	56.6	5299	56.7
Hard of Hearing	78	0.8	78	0.8
Deafness	135	1.4	135	1.4
Speech or Language Impairment	157	1.7	157	1.7
Visual Impairment	91	1.0	91	1.0
Emotional Disturbance	163	1.7	162	1.7
Orthopedic Impairment	497	5.3	494	5.3
Other Health Impairment	299	3.2	299	3.2
Established Medical Disability	0	0.0	0	0.0
Specific Learning Disability	812	8.7	809	8.7
Deaf-Blindness	2	0.0	2	0.0
Multiple Disabilities	419	4.5	418	4.5
Autism	1144	12.2	1141	12.2
Traumatic Brain Injury	90	1.0	89	1.0
Unknown	176	1.9	173	1.8
TOTAL	9367	100.0	9347	100.0

Test Results

In this section, test results are provided by level and disaggregated by demographic subgroups. Consistent with CDE policy, information on subgroups with 11 or fewer members is not reported.

Tables 4.8–4.9 summarize the performance score distributions in ELA and mathematics for all CAPA examinees, and disaggregated by grade, demographic subgroup, and primary disability. Provided are the number of students tested and the percent of students scoring in each performance level (Far Below Basic, Below Basic, Basic, Proficient, and Advanced). In the aggregate, there were approximately 4,000 students per grade in grades 2 through 10, and 3,588 in grade 11. In general, students performed better in ELA than in mathematics. Approximately 65% of all students were classified as Proficient or Advanced in ELA, and 58% in mathematics.

Tables 4.10–4.19 provide parallel information by CAPA test level. As expected, the performance distribution is substantially more varied by primary disability than grade or demographic subgroup. For ELA, most frequently, students with Specific Learning Impairment outperformed students with other primary disabilities. Across all levels, the group having the largest percentage of students in the Proficient or Advanced category was Specific Learning Impairment. The second highest was Emotional Disturbance.

Tables 4.20–4.29 summarize the scale score distributions by test level for ELA and mathematics for the population, and disaggregated by grade, demographic subgroup, and primary disability. Statistics include the number of valid scores, the scale score mean and standard deviation, the mean scale score at selected percentile points, as well as alpha reliability and SEM. Subgroup reliabilities ranged from 0.76 (Level III Mathematics for Primary Disability Specific Learning Impairment, Level IV Mathematics for Speech/ Language Impairment and Specific Learning Impairment and Emotional Disturbance) to 0.97 (Level I Mathematics for Specific Learning Impairment) with most closer to 0.90.

Table 4.8 Performance Score Distributions for All Examinees*: English–Language Arts 2006

Performance Score Distributions for All Examinees*: English–Language Arts 2006							
ELA All Levels	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	41118	10	10	15	27	38
Grade	2	4296	6	6	20	36	33
	3	4111	5	6	17	33	39
	4	3963	7	13	17	27	37
	5	4092	7	12	15	26	40
	6	4286	14	13	15	24	33
	7	4405	14	11	15	22	39
	8	4301	12	11	14	23	41
	9	4051	13	9	13	29	36
	10	4025	12	8	12	28	39
	11	3588	13	9	11	27	40
Gender	Male	26011	10	10	15	27	38
	Female	14912	11	10	15	27	37
	Unknown	195	11	8	13	28	41
Race Ethnicity	American Indian or Alaska Native	340	9	7	14	29	42
	Asian	2773	13	11	14	28	33
	Pacific Islander	189	10	10	15	30	35
	Filipino	1162	14	10	14	28	33
	Hispanic or Latino	19556	11	11	16	27	36
	African American	4727	9	8	14	27	41
	White(not Hispanic origin)	11784	10	9	14	27	40
	Unknown	587	9	7	14	29	40

Performance Score Distributions for All Examinees*: English–Language Arts 2006

ELA All Levels	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Language Fluency	English Only	25528	10	9	15	27	39
	Initially-Fluent English Proficient	1181	16	13	15	26	30
	English Learner	12954	10	10	16	27	36
	Reclassified-Fluent English Proficient	618	12	12	16	26	34
	Unknown	837	8	6	12	31	43
Economic Disadvantage	Yes	24006	10	10	16	27	37
	No	15802	11	10	14	27	38
	Unknown	1310	8	7	12	30	43
Primary Disability	Mental Retardation	18712	10	12	18	27	33
	Hard of Hearing	313	11	9	16	29	35
	Deafness	493	6	12	22	35	25
	Speech/Language Impairment	1344	1	3	8	29	58
	Visual Impairment	525	16	9	14	34	28
	Emotional Disturbance	373	3	3	6	18	70
	Orthopedic Impairment	4016	16	10	13	30	31
	Other Health Impairment	1359	7	6	13	29	45
	Specific Learning Impairment	2702	1	2	5	23	69
	Deaf Blindness	43	23	12	19	30	16
	Multiple Group	2505	18	11	14	28	29
	Autism	7546	12	10	14	27	37
	Traumatic Brain Injury	283	9	6	17	27	41
	Unknown	904	8	7	13	29	43

*Results for groups with fewer than 11 members are not reported

Table 4.9 Performance Score Distributions for All Examinees*: Mathematics 2006

Performance Score Distributions for All Examinees*: Mathematics 2006							
Mathematics All Levels Group	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	41030	14	12	17	27	29
Grade	2	4283	7	8	18	36	31
	3	4102	6	8	15	32	39
	4	3954	7	13	14	30	36
	5	4083	7	14	14	26	39
	6	4281	18	17	19	25	21
	7	4395	17	16	17	25	24
	8	4292	16	16	18	24	27
	9	4035	22	10	18	25	24
	10	4021	20	10	18	25	27
	11	3584	22	10	18	24	27
Gender	Male	25953	14	12	16	27	31
	Female	14885	15	13	18	27	26
	Unknown	192	15	10	17	23	36
Race Ethnicity	American Indian or Alaska Native	340	11	12	13	25	39
	Asian	2771	17	14	18	28	23
	Pacific Islander	187	14	10	16	30	29
	Filipino	1158	17	15	17	28	23
	Hispanic or Latino	19520	14	13	17	27	30
	African American	4718	13	11	17	28	31
	White(not Hispanic origin)	11752	14	12	17	27	30
	Unknown	584	13	9	18	29	31

Performance Score Distributions for All Examinees*: Mathematics 2006

Mathematics All Levels Group	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Language Fluency	English Only	25470	15	12	17	27	29
	Initially-Fluent English Proficient	1180	20	16	18	25	21
	English Learner	12934	13	12	16	28	30
	Reclassified-Fluent English Proficient	615	17	13	16	28	25
	Unknown	831	10	8	17	32	32
Economic Disadvantage	Yes	23961	13	12	17	27	31
	No	15765	16	13	17	27	27
	Unknown	1304	11	9	17	28	35
Primary Disability	Mental Retardation	18681	14	12	20	28	25
	Hard of Hearing	312	13	13	17	24	33
	Deafness	493	8	10	17	30	35
	Speech/Language Impairment	1342	2	2	8	29	59
	Visual Impairment	522	22	16	19	25	19
	Emotional Disturbance	372	4	2	7	26	61
	Orthopedic Impairment	4004	21	21	17	23	18
	Other Health Impairment	1358	9	8	14	29	39
	Specific Learning Impairment	2698	2	2	6	25	66
	Deaf Blindness	42	29	38	10	19	5
	Multiple group	2499	22	21	18	22	17
	Autism	7532	16	11	16	29	28
	Traumatic Brain Injury	282	14	10	14	27	35
	Unknown	893	10	9	17	30	35

*Results for groups with fewer than 11 members are not reported

Table 4.10 Performance Score Distributions*: Level I English Language Arts

Performance Score Distributions*: Level I English Language Arts							
ELA Level 1	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	8923	13	5	6	30	47
Grade	2	1227	10	4	4	29	53
	3	974	12	6	5	29	48
	4	850	11	4	5	30	50
	5	936	12	6	6	32	43
	6	899	12	5	5	33	46
	7	831	16	6	6	29	43
	8	909	15	5	6	31	44
	9	845	12	4	8	28	47
	10	775	14	5	7	27	47
	11	677	16	7	5	30	43
Gender	Male	5466	12	5	5	29	48
	Female	3425	14	6	6	31	44
	Unknown	32	19	0	3	41	38
Race Ethnicity	American Indian or Alaska Native	54	13	4	7	26	50
	Asian	672	12	6	5	32	46
	Pacific Islander	35	3	3	9	31	54
	Filipino	278	15	6	5	31	44
	Hispanic or Latino	4391	14	6	6	28	47
	African American	1003	11	4	5	28	51
	White(not Hispanic origin)	2371	11	5	6	33	45
	Unknown	119	15	3	6	29	46

Performance Score Distributions*: Level I English Language Arts

ELA Level 1	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Language Fluency	English Only	5435	12	5	6	31	45
	Initially-Fluent English Proficient	332	16	5	7	33	39
	English Learner	2919	13	5	5	27	50
	Reclassified-Fluent English Proficient	94	19	6	3	20	51
	Unknown	143	9	1	6	31	52
Economic Disadvantage	Yes	5048	12	5	5	28	49
	No	3708	14	5	6	32	43
	Unknown	167	10	2	5	34	49
Primary Disability	Mental Retardation	3047	8	4	4	26	58
	Hard of Hearing	70	9	1	6	31	53
	Deafness	41	7	7	2	29	54
	Speech/Language Impairment	53	2	0	6	25	68
	Visual Impairment	249	17	7	8	36	33
	Emotional Disturbance	19	0	0	5	26	68
	Orthopedic Impairment	2159	20	7	8	33	32
	Other Health Impairment	193	18	7	4	33	39
	Specific Learning Impairment	86	8	3	1	23	64
	Deaf Blindness	28	21	7	18	36	18
	Multiple group	1185	21	7	9	31	32
	Autism	1562	5	3	3	31	58
	Traumatic Brain Injury	63	25	6	13	16	40
	Unknown	168	13	2	5	29	51

*Results for groups with fewer than 11 members are not reported

Table 4.11 Performance Score Distributions*: Level I Mathematics

Performance Score Distributions*: Level I Mathematics							
Mathematics Level 1	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	8898	15	23	17	26	20
Grade	2	1226	12	20	18	26	25
	3	970	15	23	16	27	19
	4	846	13	21	21	29	15
	5	932	15	25	19	25	17
	6	897	14	23	16	26	21
	7	827	18	25	17	22	18
	8	907	17	24	16	23	20
	9	841	14	24	15	26	21
	10	775	16	23	14	25	22
	11	677	18	23	16	25	18
Gender	Male	5452	14	22	17	27	21
	Female	3415	16	25	17	24	18
	Unknown	31	26	23	16	10	26
Race Ethnicity	American Indian or Alaska Native	54	15	28	17	22	19
	Asian	670	13	23	18	28	17
	Pacific Islander	33	6	21	27	24	21
	Filipino	278	17	26	14	26	17
	Hispanic or Latino	4381	16	24	16	24	20
	African American	1002	13	19	17	28	23
	White(not Hispanic origin)	2363	14	23	18	26	18
	Unknown	117	19	17	18	26	21

Performance Score Distributions*: Level I Mathematics

Mathematics Level 1	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Language Fluency	English Only	5420	15	23	18	26	18
	Initially-Fluent English Proficient	331	20	23	18	28	11
	English Learner	2913	14	23	15	25	23
	Reclassified-Fluent English Proficient	94	18	21	15	27	19
	Unknown	140	10	23	16	30	21
Economic Disadvantage	Yes	5038	13	23	16	26	22
	No	3696	18	24	18	24	17
	Unknown	164	15	20	17	29	20
Primary Disability	Mental Retardation	3040	9	19	17	29	27
	Hard of Hearing	70	10	21	27	23	19
	Deafness	41	10	22	17	41	10
	Speech/Language Impairment	53	8	11	11	36	34
	Visual Impairment	247	22	25	19	21	13
	Emotional Disturbance	19	0	16	21	21	42
	Orthopedic Impairment	2153	24	30	16	19	10
	Other Health Impairment	193	24	21	17	21	17
	Specific Learning Impairment	86	12	6	9	27	47
	Deaf Blindness	27	26	48	11	15	0
	Multiple group	1182	24	31	17	18	11
	Autism	1560	5	17	19	34	25
	Traumatic Brain Injury	63	29	24	10	21	17
	Unknown	164	16	18	13	27	26

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

Table 4.12 Performance Score Distributions*: Level II English Language Arts

Performance Score Distributions*: Level II English Language Arts							
ELA Level 1I	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	6206	4	6	23	36	31
Grade	2	3069	4	6	26	38	25
	3	3137	3	6	21	34	36
Gender	Male	4146	4	7	22	36	31
	Female	2020	4	5	26	35	30
	Unknown	40	5	8	25	30	33
Race Ethnicity	American Indian or Alaska Native	65	2	0	9	45	45
	Asian	416	6	7	25	35	27
	Pacific Islander	38	3	3	16	50	29
	Filipino	208	7	12	23	35	24
	Hispanic or Latino	3063	4	6	25	35	30
	African American	647	4	4	23	37	32
	White(not Hispanic origin)	1660	3	7	20	36	33
	Unknown	109	4	6	29	36	25
Language Fluency	English Only	3812	4	6	22	36	32
	Initially-Fluent English Proficient	117	3	9	29	30	29
	English Learner	2084	3	6	25	37	29
	Reclassified-Fluent English Proficient	43	2	9	26	35	28
	Unknown	150	2	7	21	37	33
Economic Disadvantage	Yes	3659	3	5	24	37	31
	No	2308	5	8	23	34	30
	Unknown	239	2	4	20	41	34

Performance Score Distributions*: Level II English Language Arts

ELA Level 1I	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Primary Disability	Mental Retardation	2419	3	7	32	37	21
	Hard of Hearing	44	5	7	16	48	25
	Deafness	82	5	5	18	38	34
	Speech/Language Impairment	525	0	1	10	35	54
	Visual Impairment	53	6	11	26	42	15
	Emotional Disturbance	33	0	3	3	27	67
	Orthopedic Impairment	347	4	4	24	39	29
	Other Health Impairment	248	1	3	15	39	42
	Specific Learning Impairment	409	0	0	7	31	62
	Multiple group	234	7	8	23	40	23
	Autism	1609	7	10	22	33	29
	Traumatic Brain Injury	33	0	0	42	30	27
	Unknown	167	1	5	17	40	38

*Results for groups with fewer than 11 members are not reported

Table 4.13 Performance Score Distributions*: Level II Mathematics

Performance Score Distributions*: Level II Mathematics							
Mathematics Level II	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	6189	4	4	16	36	40
Grade	2	3057	5	4	18	40	34
	3	3132	4	3	15	33	45
Gender	Male	4133	4	4	15	36	41
	Female	2017	4	3	18	37	38
	Unknown	39	3	0	26	31	41
Race Ethnicity	American Indian or Alaska Native	65	0	2	8	38	52
	Asian	416	5	4	19	42	31
	Pacific Islander	38	0	11	16	37	37
	Filipino	206	7	8	17	37	31
	Hispanic or Latino	3057	4	4	17	35	41
	African American	645	4	4	13	37	42
	White(not Hispanic origin)	1653	5	3	15	37	41
	Unknown	109	6	3	21	39	32
Language Fluency	English Only	3800	5	3	16	36	40
	Initially-Fluent English Proficient	117	3	5	21	35	36
	English Learner	2080	3	4	16	37	39
	Reclassified-Fluent English Proficient	43	2	7	19	37	35
	Unknown	149	2	3	18	42	36
Economic Disadvantage	Yes	3649	3	4	16	36	42
	No	2302	6	4	17	37	36
	Unknown	238	2	2	15	38	43

Performance Score Distributions*: Level II Mathematics

Mathematics Level II	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Primary Disability	Mental Retardation	2413	4	4	21	39	31
	Hard of Hearing	44	5	2	11	16	66
	Deafness	82	4	2	5	26	63
	Speech/Language Impairment	525	1	1	5	31	62
	Visual Impairment	53	9	4	19	36	32
	Emotional Disturbance	33	0	0	3	21	76
	Orthopedic Impairment	345	4	2	19	38	38
	Other Health Impairment	248	2	2	12	32	52
	Specific Learning Impairment	408	0	1	5	21	73
	Multiple group	233	6	4	19	36	34
	Autism	1606	8	4	16	38	34
	Traumatic Brain Injury	33	0	9	24	36	30
	Unknown	163	1	2	14	39	44

*Results for groups with fewer than 11 members are not reported

Table 4.14 Performance Score Distributions*: Level III English Language Arts

Performance Score Distributions*: Level III English Language Arts							
ELA Level 1II	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	6269	6	14	19	25	36
Grade	4	3113	6	15	20	26	33
	5	3156	6	14	18	24	38
	Male	4120	6	14	18	25	37
Gender	Female	2120	5	15	20	25	34
	Unknown	29	3	10	17	24	45
	American Indian or Alaska Native	58	5	7	19	31	38
Race	Asian	426	11	17	15	28	28
Ethnicity	Pacific Islander	25	8	8	24	36	24
	Filipino	176	13	15	17	30	26
	Hispanic or Latino	3052	5	16	21	24	34
	African American	678	4	13	17	25	41
	White(not Hispanic origin)	1770	6	11	17	24	40
	Unknown	84	5	8	17	33	37
Language	English Only	3881	6	13	18	25	37
Fluency	Initially-Fluent English Proficient	134	7	27	22	22	22
	English Learner	2073	5	16	20	25	34
	Reclassified-Fluent English Proficient	75	1	19	23	27	31
	Unknown	106	8	8	12	29	42
Economic	Yes	3798	5	15	19	25	35
Disadvantage	No	2301	8	14	19	24	36
	Unknown	170	6	11	10	31	42

Performance Score Distributions*: Level III English Language Arts

ELA Level 1II	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Primary Disability	Mental Retardation	2779	5	19	23	25	27
	Hard of Hearing	46	4	22	15	24	35
	Deafness	60	2	15	30	37	17
	Speech/Language Impairment	323	2	4	9	25	60
	Visual Impairment	44	7	11	23	43	16
	Emotional Disturbance	56	0	4	13	16	68
	Orthopedic Impairment	392	5	15	19	29	31
	Other Health Impairment	260	5	6	17	24	48
	Specific Learning Impairment	518	1	2	6	17	74
	Multiple group	266	7	16	25	21	30
	Autism	1352	12	14	16	26	32
	Traumatic Brain Injury	37	5	11	11	32	41
	Unknown	134	5	12	16	24	43

*Results for groups with fewer than 11 members are not reported

Table 4.15 Performance Score Distributions*: Level III Mathematics

Performance Score Distributions*: Level III Mathematics							
Mathematics Level III	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	6259	5	11	12	29	43
Grade	4	3108	5	11	12	31	41
	5	3151	5	10	12	27	46
Gender	Male	4112	5	10	11	28	45
	Female	2118	5	11	14	31	40
	Unknown	29	7	7	17	21	48
Race Ethnicity	American Indian or Alaska Native	58	0	7	12	24	57
	Asian	426	10	11	11	33	35
	Pacific Islander	25	4	4	16	24	52
	Filipino	176	7	10	19	28	35
	Hispanic or Latino	3046	4	12	13	29	43
	African American	678	5	9	9	28	49
	White(not Hispanic origin)	1766	6	10	11	28	45
	Unknown	84	2	5	12	36	45
Language Fluency	English Only	3874	6	10	12	29	44
	Initially-Fluent English Proficient	134	4	15	23	26	31
	English Learner	2070	4	11	13	29	43
	Reclassified-Fluent English Proficient	75	0	13	13	31	43
	Unknown	106	11	7	5	32	45
Economic Disadvantage	Yes	3793	4	10	12	29	44
	No	2296	7	11	12	29	41
	Unknown	170	6	8	7	26	54

Performance Score Distributions*: Level III Mathematics

Mathematics Level III	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Primary Disability	Mental Retardation	2775	4	13	17	31	35
	Hard of Hearing	46	9	13	9	22	48
	Deafness	60	0	7	8	32	53
	Speech/Language Impairment	323	1	2	3	24	70
	Visual Impairment	44	11	7	16	41	25
	Emotional Disturbance	56	2	2	2	25	70
	Orthopedic Impairment	392	6	14	14	33	34
	Other Health Impairment	259	3	6	8	30	54
	Specific Learning Impairment	518	0	1	2	20	76
	Multiple Group	264	7	19	15	26	34
	Autism	1349	10	10	10	28	42
	Traumatic Brain Injury	37	3	14	8	30	46
	Unknown	134	4	10	8	28	50

*Results for groups with fewer than 11 members are not reported

Table 4.16 Performance Score Distributions*: Level IV English–Language Arts

Performance Score Distributions*: Level IV English–Language Arts							
ELA Level 1V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	10353	13	13	17	21	36
Grade	6	3387	15	16	18	22	29
	7	3574	13	12	16	21	38
	8	3392	12	12	16	20	40
Gender	Male	6542	13	13	17	21	36
	Female	3760	13	14	17	20	36
	Unknown	51	16	12	10	20	43
Race Ethnicity	American Indian or Alaska Native	95	14	15	20	17	35
	Asian	651	17	15	16	21	31
	Pacific Islander	50	14	22	14	22	28
	Filipino	262	17	14	15	17	37
	Hispanic or Latino	4921	14	14	18	22	33
	African American	1198	12	12	16	22	38
	White(not Hispanic origin)	3025	12	13	17	20	39
	Unknown	151	9	13	11	22	46
Language Fluency	English Only	6423	13	13	17	21	37
	Initially-Fluent English Proficient	277	23	16	19	19	23
	English Learner	3272	14	14	17	21	33
	Reclassified-Fluent English Proficient	176	9	18	18	25	31
	Unknown	205	7	9	11	24	49
Economic Disadvantage	Yes	6197	14	13	18	21	34
	No	3783	13	13	15	20	38

Performance Score Distributions*: Level IV English–Language Arts

ELA Level 1V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
	Unknown	373	10	10	12	23	46
Primary Disability	Mental Retardation	5163	14	16	20	21	29
	Hard of Hearing	75	20	7	21	16	36
	Deafness	175	7	15	25	34	19
	Speech/Language Impairment	286	1	8	7	21	62
	Visual Impairment	88	19	11	17	26	26
	Emotional Disturbance	102	4	4	7	22	64
	Orthopedic Impairment	621	14	16	22	20	28
	Other Health Impairment	359	8	10	13	21	47
	Specific Learning Impairment	877	1	2	5	22	70
	Multiple group	401	18	17	16	21	28
	Autism	1879	19	12	14	19	36
	Traumatic Brain Injury	60	3	7	22	32	37
	Unknown	259	10	9	13	24	44

*Results for groups with fewer than 11 members are not reported

Table 4.17 Performance Score Distributions*: Level IV Mathematics

Performance Score Distributions*: Level IV Mathematics							
Mathematics Level IV	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	10337	17	14	18	25	25
Grade	6	3384	20	15	20	24	21
	7	3568	17	14	17	26	26
	8	3385	16	14	18	24	28
Gender	Male	6534	17	13	19	25	26
	Female	3753	17	16	18	25	24
	Unknown	50	16	20	14	24	26
Race Ethnicity	American Indian or Alaska Native	95	20	18	15	19	28
	Asian	651	24	20	20	21	15
	Pacific Islander	50	20	14	16	32	18
	Filipino	260	22	19	18	23	17
	Hispanic or Latino	4915	17	14	18	25	26
	African American	1195	15	13	18	25	28
	White(not Hispanic origin)	3020	17	14	19	25	26
	Unknown	151	14	12	21	25	28
Language Fluency	English Only	6411	17	14	19	24	25
	Initially-Fluent English Proficient	277	26	22	13	18	20
	English Learner	3268	17	14	18	26	25
	Reclassified-Fluent English Proficient	175	19	13	21	30	17
	Unknown	206	8	10	21	31	30

Performance Score Distributions*: Level IV Mathematics

Mathematics Level IV	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Economic Disadvantage	Yes	6186	17	13	18	25	27
	No	3778	19	16	19	25	22
	Unknown	373	12	13	22	25	28
Primary Disability	Mental Retardation	5154	19	16	20	25	20
	Hard of Hearing	74	15	19	15	27	24
	Deafness	175	7	15	23	30	25
	Speech/Language Impairment	284	2	3	17	33	45
	Visual Impairment	87	22	15	20	16	28
	Emotional Disturbance	102	5	3	5	26	61
	Orthopedic Impairment	620	20	15	17	27	21
	Other Health Impairment	359	11	12	15	28	35
	Specific Learning Impairment	877	1	2	7	29	61
	Multiple group	402	22	16	21	23	18
	Autism	1876	26	17	20	20	17
	Traumatic Brain Injury	60	10	3	17	33	37
	Unknown	259	11	10	21	29	29

*Results for groups with fewer than 11 members are not reported

Table 4.18 Performance Score Distributions*: Level V English–Language Arts

Performance Score Distributions*: Level V English–Language Arts							
ELA Level V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	9367	13	10	14	28	36
Grade	9	3206	13	10	14	29	33
	10	3250	12	9	13	28	37
	11	2911	13	9	13	26	39
Gender	Male	5737	12	9	14	28	36
	Female	3587	14	10	13	27	36
	Unknown	43	9	7	12	28	44
Race Ethnicity	American Indian or Alaska Native	68	7	4	10	29	49
	Asian	608	17	10	15	28	30
	Pacific Islander	41	17	10	15	17	41
	Filipino	238	18	7	16	29	30
	Hispanic or Latino	4129	13	11	14	29	33
	African American	1201	11	8	14	28	39
	White(not Hispanic origin)	2958	12	8	12	26	42
	Unknown	124	12	5	10	28	44
Language Fluency	English Only	5977	12	9	13	27	39
	Initially-Fluent English Proficient	321	17	12	13	26	31
	English Learner	2606	13	11	15	29	32
	Reclassified-Fluent English Proficient	230	17	10	15	27	32
	Unknown	233	11	5	10	33	40
Economic Disadvantage	Yes	5304	13	10	14	28	35
	No	3702	13	9	13	27	38

Performance Score Distributions*: Level V English–Language Arts

ELA Level V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
	Unknown	361	9	7	11	28	44
Primary Disability	Mental Retardation	5304	13	11	15	29	32
	Hard of Hearing	78	10	13	19	33	24
	Deafness	135	6	13	24	36	21
	Speech/Language Impairment	157	2	4	4	31	59
	Visual Impairment	91	19	11	14	26	30
	Emotional Disturbance	163	5	3	4	13	75
	Orthopedic Impairment	497	18	11	14	23	34
	Other Health Impairment	299	6	5	13	30	45
	Specific Learning Impairment	812	1	2	5	25	67
	Multiple group	419	23	13	15	23	26
	Autism	1144	19	10	14	28	29
	Traumatic Brain Injury	90	7	7	11	27	49
	Unknown	175	12	6	13	31	38

*Results for groups with fewer than 11 members are not reported

Table 4.19 Performance Score Distributions*: Level V Mathematics

Performance Score Distributions*: Level V Mathematics							
Mathematics Level V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
All	All	9347	23	7	19	25	27
Grade	9	3194	24	7	19	25	25
	10	3246	21	7	19	25	28
	11	2907	23	7	18	24	29
Gender	Male	5722	22	6	18	24	30
	Female	3582	24	7	21	25	23
	Unknown	43	21	0	12	26	42
Race Ethnicity	American Indian or Alaska Native	68	15	4	15	24	43
	Asian	608	27	7	20	24	23
	Pacific Islander	41	32	0	7	32	29
	Filipino	238	27	8	19	24	21
	Hispanic or Latino	4121	23	7	19	24	26
	African American	1198	20	6	21	27	25
	White(not Hispanic origin)	2950	22	6	18	24	30
	Unknown	123	20	7	17	24	33
Language Fluency	English Only	5965	23	7	19	24	27
	Initially-Fluent English Proficient	321	29	6	19	22	24
	English Learner	2603	22	7	19	25	27
	Reclassified-Fluent English Proficient	228	24	11	14	25	27
	Unknown	230	17	3	19	30	31

Performance Score Distributions*: Level V Mathematics

Mathematics Level V	Subgroup	N	Percent				
			Far Below Basic	Below Basic	Basic	Proficient	Advanced
Economic Disadvantage	Yes	5295	22	7	19	25	27
	No	3693	24	7	18	25	26
	Unknown	359	17	5	17	26	35
Primary Disability	Mental Retardation	5299	24	8	22	25	21
	Hard of Hearing	78	23	4	18	29	26
	Deafness	135	13	7	20	27	33
	Speech/Language Impairment	157	4	2	9	25	59
	Visual Impairment	91	33	3	21	29	14
	Emotional Disturbance	162	6	1	9	27	58
	Orthopedic Impairment	494	30	6	21	21	21
	Other Health Impairment	299	11	4	19	31	35
	Specific Learning Impairment	809	2	1	8	25	63
	Multiple group	418	36	8	18	22	16
	Autism	1141	33	7	15	22	23
	Traumatic Brain Injury	89	16	3	13	22	45
	Unknown	172	16	2	24	26	33

*Results for groups with fewer than 11 members are not reported

Table 4.20 Scale Score Distributions*: Level I English–Language Arts

Scale Score Distributions*: Level I English–Language Arts													
ELA Level I	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	8923	43.23	13.03	15	17	35	45	54	60	60	0.90	4.12
Grade	2	1227	45.27	12.58	15	19	38	48	57	60	60	0.90	3.98
	3	974	43.75	12.64	15	19	36	45	54	60	60	0.90	4.00
	4	850	43.81	12.36	15	19	36	45	52	60	60	0.89	4.10
	5	936	42.62	12.90	15	17	35	44	54	60	60	0.90	4.08
	6	899	43.57	12.82	15	17	36	44	55	60	60	0.90	4.05
	7	831	41.63	13.61	15	15	32	43	52	60	60	0.91	4.08
	8	909	42.34	13.31	15	15	33	44	52	60	60	0.91	3.99
	9	845	43.53	13.02	15	19	35	45	55	60	60	0.91	3.91
	10	775	42.87	13.48	15	15	33	45	54	60	60	0.92	3.81
	11	677	41.62	13.48	15	15	32	43	52	60	60	0.91	4.04
Gender	Male	5466	43.72	12.94	15	17	36	45	55	60	60	0.90	4.09
	Female	3425	42.46	13.12	15	17	33	44	54	60	60	0.91	3.94
	Unknown	32	41.34	13.48	15	17	35	42	52	60	60	0.88	4.67
Race Ethnicity	American Indian or Alaska Native	54	44.00	13.70	15	15	35	45.5	55	60	60	0.92	3.87
	Asian	672	43.10	12.74	15	17	36	44	54	60	60	0.90	4.03
	Pacific Islander	35	46.23	10.51	15	27	41	46	55	60	60	0.84	4.20
	Filipino	278	42.05	12.98	15	17	33	44	52	60	60	0.89	4.30
	Hispanic or Latino	4391	43.08	13.40	15	17	35	44	55	60	60	0.91	4.02
	African American	1003	44.24	12.97	15	17	36	46	55	60	60	0.90	4.10
	White(not Hispanic origin)	2371	43.21	12.44	15	19	36	44	52	60	60	0.89	4.13
	Unknown	119	42.87	13.16	15	17	35	45	54	60	60	0.91	3.95

Scale Score Distributions*: Level I English–Language Arts

ELA Level I	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Language Fluency	English Only	5435	42.97	12.72	15	19	35	44	52	60	60	0.90	4.02
	Initially-Fluent English Proficient	332	40.83	12.98	15	15	32	42.5	51	60	60	0.89	4.30
	English Learner	2919	43.89	13.49	15	15	35	45	57	60	60	0.92	3.82
	Reclassified-Fluent English Proficient	94	43.22	15.54	15	15	29	46	59	60	60	0.94	3.81
	Unknown	143	45.29	12.39	15	19	37	46	57	60	60	0.89	4.11
Economic Disadvantage	Yes	5048	44.04	13.01	15	19	36	45	55	60	60	0.91	3.90
	No	3708	42.06	13.01	15	15	33	43	52	60	60	0.90	4.11
	Unknown	167	44.80	11.85	15	21	37	45	55	60	60	0.88	4.10
Primary Disability	Mental Retardation	3047	46.64	12.32	15	21	39	49	59	60	60	0.90	3.90
	Hard of Hearing	70	43.73	11.54	15	19	37	46	51	60	60	0.88	4.00
	Deafness	41	44.59	11.51	17	23	39	48	52	60	60	0.86	4.31
	Speech/Language Impairment	53	48.92	9.10	24	32	43	49	57	60	60	0.82	3.86
	Visual Impairment	249	39.85	12.66	15	17	31	41	49	60	60	0.89	4.20
	Emotional Disturbance	19	50.53	10.31	31	31	40	54	60	60	60	0.92	2.92
	Orthopedic Impairment	2159	38.74	13.18	15	15	29	40	49	60	60	0.91	3.95
	Other Health Impairment	193	41.00	13.68	15	15	31	42	52	60	60	0.91	4.10
	Specific Learning Impairment	86	48.33	13.14	15	17	44	50	60	60	60	0.93	3.48
	Deaf Blindness	28	35.64	12.18	15	15	29	36	41.5	59	60	0.89	4.04
	Multiple group	1185	38.31	13.26	15	15	27	39	49	60	60	0.90	4.19
	Autism	1562	46.86	10.43	19	26	41	49	55	60	60	0.85	4.04
	Traumatic Brain Injury	63	38.35	15.14	15	15	25	42	52	60	60	0.92	4.28
	Unknown	168	44.78	13.31	15	19	36.5	46	58	60	60	0.91	3.99

*Results for groups with fewer than 11 members are not reported

Table 4.21 Scale Score Distributions*: Level I Mathematics

Scale Score Distributions*: Level I Mathematics													
Mathematics Level I	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	8898	33.18	11.50	15	15	24	33	39	57	60	0.92	3.25
Grade	2	1226	34.72	11.31	15	15	27	35	41	57	60	0.91	3.39
	3	970	33.2	11.10	15	15	25	33	39	57	60	0.92	3.14
	4	846	32.84	10.55	15	15	25	33	39	51	60	0.91	3.17
	5	932	32.34	10.90	15	15	23	33	39	57	60	0.91	3.27
	6	897	33.79	11.80	15	15	24	34	40	60	60	0.93	3.12
	7	827	32.07	11.89	15	15	22	32	39	57	60	0.93	3.15
	8	907	32.89	12.07	15	15	23	33	39	60	60	0.93	3.19
	9	841	33.57	11.61	15	15	24	33	40	60	60	0.92	3.28
	10	775	33.36	11.96	15	15	23	33	40	57	60	0.94	2.93
	11	677	32.24	11.77	15	15	23	33	39	57	60	0.93	3.11
Gender	Male	5452	33.68	11.49	15	15	25	34	40	57	60	0.92	3.25
	Female	3415	32.41	11.47	15	15	23	33	39	57	60	0.92	3.24
	Unknown	31	31.29	12.93	15	15	19	30	44	51	57	0.94	3.17
Race Ethnicity	American Indian or Alaska Native	54	32.98	12.85	15	15	23	32.5	39	60	60	0.94	3.15
	Asian	670	32.96	10.78	15	15	24	33	39	51	60	0.91	3.23
	Pacific Islander	33	35.27	10.51	15	18	28	33	41	60	60	0.86	3.93
	Filipino	278	31.79	10.98	15	15	22	32	39	51	60	0.92	3.11
	Hispanic or Latino	4381	33.06	11.73	15	15	23	33	40	57	60	0.93	3.10
	African American	1002	34.27	11.45	15	15	26	35	40	57	60	0.92	3.24
	White	2363	33.15	11.30	15	15	24	33	39	57	60	0.92	3.20
	Unknown	117	33.48	12.03	15	15	23	34	40	60	60	0.93	3.18
Language Fluency	English Only	5420	32.93	11.27	15	15	24	33	39	57	60	0.92	3.19
	Initially-Fluent English Proficient	331	30.61	10.33	15	15	22	31	37	46	60	0.91	3.10
	English Learner	2913	33.87	11.99	15	15	24	34	41	60	60	0.93	3.17
	Reclassified-Fluent English Proficient	94	33.05	12.39	15	15	23	33	41	60	60	0.94	3.03
	Unknown	140	34.89	10.95	15	15.5	28	35	41	58.5	60	0.90	3.46

Scale Score Distributions*: Level I Mathematics

Mathematics Level I	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Economic Disadvantage	Yes	5038	33.95	11.60	15	15	25	34	40	57	60	0.92	3.28
	No	3696	32.11	11.29	15	15	23	33	39	57	60	0.92	3.19
	Unknown	164	33.87	11.66	15	15	25	33.5	40	60	60	0.92	3.30
Primary Disability	Mental Retardation	3040	35.96	11.31	15	16	28	36	43	60	60	0.91	3.39
	Hard of Hearing	70	33.51	9.90	15	15	29	33	38	48	60	0.87	3.57
	Deafness	41	33.05	9.17	18	19	24	35	39	44	57	0.88	3.18
	Speech/Language Impairment	53	39.64	11.94	18	20	33	37	48	60	60	0.89	3.96
	Visual Impairment	247	30.33	10.57	15	15	22	30	38	48	60	0.92	2.99
	Emotional Disturbance	19	41.84	13.30	23	23	32	39	57	60	60	0.91	3.99
	Orthopedic Impairment	2153	28.96	10.65	15	15	21	28	36	48	60	0.92	3.01
	Other Health Impairment	193	31.66	12.25	15	15	22	32	39	60	60	0.93	3.24
	Specific Learning Impairment	86	42.60	14.97	15	15	33	40	60	60	60	0.97	2.59
	Deaf Blindness	27	25.41	7.49	15	15	20	24	33	40	41	0.91	2.25
	Multiple group	1182	28.92	10.72	15	15	21	28	36	48	60	0.93	2.84
	Autism	1560	36.71	9.94	15	20	31	36	43	57	60	0.87	3.58
	Traumatic Brain Injury	63	30.46	13.61	15	15	19	29	39	60	60	0.95	3.04
	Unknown	164	34.80	12.83	15	15	23	35	43	60	60	0.94	3.14

*Results for groups with fewer than 11 members are not reported

Table 4.22 Scale Score Distributions*: Level II English–Language Arts

Scale Score Distributions*: Level II English–Language Arts													
ELA Level II	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	6206	37.78	8.13	15	26	33	37	41	60	60	0.89	2.70
Grade	2	3069	36.76	7.44	15	26	33	36	41	47	60	0.88	2.58
	3	3137	38.77	8.64	15	26	34	38	42	60	60	0.90	2.73
Gender	Male	4146	37.83	8.17	15	26	34	37	41	60	60	0.89	2.71
	Female	2020	37.66	8.01	15	26	33	37	41	60	60	0.89	2.66
	Unknown	40	38.00	10.27	15	21	32	36.5	44	60	60	0.88	3.56
Race Ethnicity	American Indian or Alaska Native	65	40.78	7.19	20	34	37	40	42	60	60	0.79	3.29
	Asian	416	36.55	7.80	15	23	32	36	41	47	60	0.89	2.59
	Pacific Islander	38	38.42	7.35	23	28	35	37	42	60	60	0.87	2.65
	Filipino	208	36.18	8.64	15	23	32	36	40	60	60	0.89	2.87
	Hispanic or Latino	3063	37.72	8.11	16	26	33	37	41	60	60	0.89	2.69
	African American	647	38.22	8.13	15	26	34	37	42	60	60	0.88	2.82
	White	1660	38.15	8.21	15	26	34	38	42	60	60	0.88	2.84
	Unknown	109	36.69	7.60	15	25	33	36	40	47	60	0.87	2.74
Language Fluency	English Only	3812	37.87	8.31	15	26	34	37	42	60	60	0.89	2.76
	Initially-Fluent English Proficient	117	37.03	8.29	18	26	32	36	41	60	60	0.91	2.49
	English Learner	2084	37.65	7.91	15	27	33.5	37	41	60	60	0.89	2.62
	Reclassified-Fluent English Proficient	43	36.56	6.69	22	27	31	36	41	44	60	0.88	2.32
	Unknown	150	37.95	6.76	22	26	34	38	42	47	60	0.86	2.53
Economic Disadvantage	Yes	3659	38.02	7.99	16	27	34	37	42	60	60	0.88	2.77
	No	2308	37.29	8.43	15	25	33	37	41	60	60	0.89	2.80
	Unknown	239	38.68	7.07	22	29	34	38	42	47	60	0.84	2.83

Scale Score Distributions*: Level II English–Language Arts

ELA Level II	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	2419	36.30	7.11	16	26	32	36	40	47	60	0.88	2.46
	Hard of Hearing	44	37.86	7.50	20	27	34	37.5	40.5	47	60	0.86	2.81
	Deafness	82	37.80	7.71	15	28	34	37.5	42	47	60	0.88	2.67
	Speech/Language Impairment	525	42.22	7.58	29	32	38	41	44	60	60	0.80	3.39
	Visual Impairment	53	35.49	7.34	15	23	32	36	39	47	60	0.88	2.54
	Emotional Disturbance	33	44.97	9.05	25	34	40	42	47	60	60	0.88	3.14
	Orthopedic Impairment	347	37.83	8.19	15	26	34	37	41	60	60	0.88	2.84
	Other Health Impairment	248	40.58	8.21	23	30	36	39	44	60	60	0.86	3.07
	Specific Learning Impairment	409	43.60	7.85	30	33	39	42	47	60	60	0.80	3.51
	Multiple group	234	36.26	7.95	15	23	33	36	40	47	60	0.89	2.64
	Autism	1609	36.58	8.48	15	22	32	37	41	47	60	0.90	2.68
	Traumatic Brain Injury	33	36.82	5.28	30	30	33	36	41	47	47	0.83	2.18
	Unknown	167	39.73	8.45	15	28	35	38	44	60	60	0.87	3.05

*Results for groups with fewer than 11 members are not reported

Table 4.23 Scale Score Distributions*: Level II Mathematics

Scale Score Distributions*: Level II Mathematics													
Mathematics Level II	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	6189	39.43	8.32	15	27	35	39	45	55	60	0.88	2.88
Grade	2	3057	38.45	7.86	15	27	34	38	42	55	60	0.87	2.83
	3	3132	40.39	8.63	15	27	36	40	45	55	60	0.88	2.99
Gender	Male	4133	39.59	8.42	15	27	35	39	45	55	60	0.88	2.92
	Female	2017	39.08	8.06	15	27	34	39	43	55	60	0.87	2.91
	Unknown	39	40.26	9.35	15	30	34	38	49	55	60	0.90	2.96
Race Ethnicity	American Indian or Alaska Native	65	42.83	7.91	27	34	37	42	46	60	60	0.83	3.26
	Asian	416	38.02	8.15	15	25	34	38	42	55	60	0.86	3.05
	Pacific Islander	38	39.61	7.65	27	27	34	39	45	55	60	0.87	2.76
	Filipino	206	37.13	8.18	15	25	34	37	41	55	55	0.87	2.95
	Hispanic or Latino	3057	39.48	8.18	15	27	35	39	45	55	60	0.88	2.83
	African American	645	39.91	8.54	15	27	36	39	45	55	60	0.88	2.96
	White	1653	39.73	8.48	15	27	35	39	45	55	60	0.87	3.06
	Unknown	109	38.37	8.03	15	25	34	38	42	55	60	0.86	3.00
Language Fluency	English Only	3800	39.52	8.58	15	27	35	39	45	55	60	0.88	2.97
	Initially-Fluent English Proficient	117	38.70	7.89	25	27	34	37	43	55	60	0.87	2.84
	English Learner	2080	39.34	7.92	15	27	35	39	43	55	60	0.87	2.86
	Reclassified-Fluent English Proficient	43	38.16	6.88	23	27	33	37	43	49	55	0.89	2.28
	Unknown	149	39.33	7.39	15	30	36	39	43	55	60	0.85	2.86
													0.00
Economic Disadvantage	Yes	3649	39.88	8.11	15	27	35	39	45	55	60	0.87	2.92
	No	2302	38.63	8.66	15	25	34	38	43	55	60	0.88	3.00
	Unknown	238	40.17	7.36	20	30	36	39	43	55	60	0.83	3.03

Scale Score Distributions*: Level II Mathematics

Mathematics Level II	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	2413	38.00	7.30	15	27	34	37	42	49	60	0.87	2.63
	Hard of Hearing	44	42.30	9.33	17	29	37.5	42	46	60	60	0.90	2.95
	Deafness	82	43.27	9.76	15	29	38	43	46	60	60	0.89	3.24
	Speech/Language Impairment	525	43.51	7.35	27	34	39	42	46	60	60	0.80	3.29
	Visual Impairment	53	36.83	8.29	15	15	33	37	42	46	55	0.90	2.62
	Emotional Disturbance	33	48.00	8.99	30	35	41	46	55	60	60	0.87	3.24
	Orthopedic Impairment	345	39.29	8.44	15	27	35	39	43	55	60	0.88	2.92
	Other Health Impairment	248	42.00	8.35	15	30	37	41	46	55	60	0.88	2.89
	Specific Learning Impairment	408	45.17	7.14	31	34	40	45	49	60	60	0.78	3.35
	Multiple group	233	38.23	8.47	15	25	34	37	43	55	60	0.89	2.81
	Autism	1606	38.07	8.83	15	20	34	38	43	55	60	0.88	3.06
	Traumatic Brain Injury	33	38.18	6.98	27	29	33	38	42	49	60	0.89	2.32
	Unknown	163	41.23	8.05	25	32	36	40	45	60	60	0.84	3.22

*Results for groups with fewer than 11 members are not reported

Table 4.24 Scale Score Distributions*: Level III English–Language Arts

Scale Score Distributions*: Level III English–Language Arts													
ELA Level III	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	6269	37.81	10.24	16	22	31	37	44	60	60	0.90	3.24
Grade	4	3113	37.25	9.91	16	22	30	37	43	56	60	0.89	3.29
	5	3156	38.35	10.52	15	22	31	37	44	60	60	0.91	3.16
Gender	Male	4120	37.95	10.34	15	22	31	37	44	60	60	0.90	3.27
	Female	2120	37.50	10.04	16	22	30	37	44	58	60	0.90	3.17
	Unknown	29	39.24	9.94	15	25	33	40	46	56	60	0.89	3.30
Race Ethnicity	American Indian or Alaska Native	58	39.48	10.11	15	22	33	39	44	60	60	0.88	3.50
	Asian	426	35.45	9.95	15	20	28	36	41	56	60	0.90	3.15
	Pacific Islander	25	37.12	9.41	20	22	31	37	40	56	56	0.88	3.26
	Filipino	176	35.40	10.31	16	18	29	36	41	56	60	0.91	3.09
	Hispanic or Latino	3052	37.35	10.10	16	22	30	37	43	60	60	0.90	3.19
	African American	678	39.41	10.61	15	23	32	38	46	60	60	0.90	3.36
	White	1770	38.70	10.28	16	22	32	38	44	60	60	0.90	3.25
	Unknown	84	38.60	8.96	16	23	33	38	44	56	60	0.85	3.47
Language Fluency	English Only	3881	38.13	10.42	15	22	31	37	44	60	60	0.90	3.30
	Initially-Fluent English Proficient	134	34.49	10.09	16	21	28	33	38	60	60	0.90	3.19
	English Learner	2073	37.36	9.87	16	22	31	37	43	56	60	0.90	3.12
	Reclassified-Fluent English Proficient	75	38.11	9.64	22	25	30	37	44	56	60	0.87	3.48
	Unknown	106	38.56	10.54	15	19	34	38	46	56	60	0.90	3.33
Economic Disadvantage	Yes	3798	37.90	10.12	16	22	31	37	44	60	60	0.90	3.20
	No	2301	37.56	10.45	15	21	30	37	44	60	60	0.90	3.30
	Unknown	170	39.08	9.85	15	21	34	38.5	44	56	60	0.89	3.27

Scale Score Distributions*: Level III English–Language Arts

ELA Level III	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	2779	36.12	9.45	16	22	30	35	41	56	60	0.90	2.99
	Hard of Hearing	46	37.46	11.01	18	23	28	36.5	44	60	60	0.92	3.11
	Deafness	60	35.03	6.13	21	25	31	35.5	38	47.5	49	0.80	2.74
	Speech/Language Impairment	323	42.96	9.06	21	29	37	43	49	60	60	0.86	3.39
	Visual Impairment	44	35.14	7.71	15	22	30.5	37	40	44	56	0.84	3.08
	Emotional Disturbance	56	45.98	10.48	25	30	38.5	44	56	60	60	0.89	3.48
	Orthopedic Impairment	392	37.44	9.78	15	22	31	37	43	60	60	0.89	3.24
	Other Health Impairment	260	40.27	9.77	19	22.5	33	40	46	60	60	0.89	3.24
	Specific Learning Impairment	518	46.01	8.90	25	32	40	44	56	60	60	0.83	3.67
	Multiple group	266	36.49	10.53	15	21	30	35	43	60	60	0.91	3.16
	Autism	1352	36.54	10.75	15	20	29	37	43	56	60	0.91	3.23
	Traumatic Brain Injury	37	38.11	7.96	22	22	34	39	43	49	60	0.84	3.18
	Unknown	134	39.18	10.50	15	22	32	38.5	46	60	60	0.90	3.32

*Results for groups with fewer than 11 members are not reported

Table 4.25 Scale Score Distributions*: Level III Mathematics

Scale Score Distributions*: Level III Mathematics													
Mathematics Level III	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	6259	41.40	10.92	17	24	33	41	49	60	60	0.91	3.28
Grade	4	3108	40.84	10.58	17	24	33	41	49	60	60	0.90	3.35
	5	3151	41.96	11.22	17	24	33	42	49	60	60	0.91	3.37
Gender	Male	4112	41.78	11.06	17	24	34	42	49	60	60	0.91	3.32
	Female	2118	40.64	10.56	17	25	33	41	49	60	60	0.91	3.17
	Unknown	29	42.93	12.83	16	21	33	42	56	60	60	0.93	3.39
Race Ethnicity	American Indian or Alaska Native	58	44.48	10.12	25	28	36	44	56	60	60	0.87	3.65
	Asian	426	39.24	11.05	15	21	31	40	46	60	60	0.91	3.32
	Pacific Islander	25	41.88	9.86	21	29	35	44	46	60	60	0.87	3.56
	Filipino	176	39.47	11.69	15	22	32	37	47.5	60	60	0.91	3.51
	Hispanic or Latino	3046	41.12	10.65	19	25	33	41	49	60	60	0.90	3.37
	African American	678	43.00	11.17	19	25	36	42	49	60	60	0.91	3.35
	White	1766	41.77	11.10	17	24	34	42	49	60	60	0.91	3.33
	Unknown	84	43.56	10.35	19	28	37	41.5	56	60	60	0.87	3.73
Language Fluency	English Only	3874	41.58	11.15	17	24	33	41	49	60	60	0.91	3.35
	Initially-Fluent English Proficient	134	38.60	10.39	21	25	30	37	46	60	60	0.89	3.45
	English Learner	2070	41.20	10.45	19	25	34	41	49	60	60	0.90	3.30
	Reclassified-Fluent English Proficient	75	41.88	9.94	26	27	33	42	49	60	60	0.88	3.44
	Unknown	106	42.07	12.17	15	20	35	42	49	60	60	0.94	2.98
Economic Disadvantage	Yes	3793	41.66	10.72	19	25	34	42	49	60	60	0.90	3.39
	No	2296	40.81	11.16	16	23	33	41	49	60	60	0.91	3.35
	Unknown	170	43.59	11.45	15	24	37	44	56	60	60	0.92	3.24

Scale Score Distributions*: Level III Mathematics

Mathematics Level III	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	2775	39.66	10.25	19	25	32	40	46	60	60	0.9	3.24
	Hard of Hearing	46	42.80	13.11	22	22	31	42	56	60	60	0.94	3.21
	Deafness	60	44.70	9.29	26	28.5	40	44	49	60	60	0.83	3.83
	Speech/Language Impairment	323	47.45	9.12	24	32	42	46	56	60	60	0.84	3.65
	Visual Impairment	44	37.43	9.97	15	23	31.5	37.5	43	56	60	0.89	3.31
	Emotional Disturbance	56	48.84	9.74	23	33	41	49	60	60	60	0.86	3.64
	Orthopedic Impairment	392	39.41	10.63	15	24	31.5	38	46	60	60	0.91	3.19
	Other Health Impairment	259	43.73	9.91	22	26	37	44	49	60	60	0.89	3.29
	Specific Learning Impairment	518	49.10	8.43	28	36	44	49	56	60	60	0.76	4.13
	Multiple group	264	38.82	11.61	15	22	29	37	46	60	60	0.93	3.07
	Autism	1349	40.63	11.57	16	22	32	41	49	60	60	0.92	3.27
	Traumatic Brain Injury	37	42.08	10.68	24	25	35	42	49	60	60	0.89	3.54
	Unknown	134	43.39	11.37	15	25	37	43	56	60	60	0.91	3.41

*Results for groups with fewer than 11 members are not reported

Table 4.26 Scale Score Distributions*: Level IV English–Language Arts

Scale Score Distributions*: Level IV English–Language Arts													
ELA Level IV	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	10353	36.70	10.66	15	20	29	36	43	60	60	0.91	3.20
Grade	6	3387	35.16	10.05	15	20	28	35	42	54	60	0.90	3.18
	7	3574	37.13	10.74	15	20	29	37	43	60	60	0.91	3.22
	8	3392	37.77	11.00	15	20	30	37	45	60	60	0.91	3.30
Gender	Male	6542	36.77	10.73	15	20	29	36	43	60	60	0.91	3.22
	Female	3760	36.56	10.53	16	20	29	36	43	60	60	0.91	3.16
	Unknown	51	37.39	11.69	15	18	28	38	45	60	60	0.92	3.31
Race Ethnicity	American Indian or Alaska Native	95	36.18	10.64	16	20	29	35	43	54	60	0.92	3.01
	Asian	651	35.24	10.72	15	17	28	35	42	54	60	0.91	3.22
	Pacific Islander	50	35.06	11.00	15	21	27	34.5	41	54	60	0.93	2.91
	Filipino	262	36.08	11.29	15	18	28	36	43	60	60	0.92	3.19
	Hispanic or Latino	4921	36.19	10.42	15	20	29	36	42	60	60	0.91	3.13
	African American	1198	37.18	10.52	15	20	30	37	43	60	60	0.91	3.16
	White	3025	37.66	10.96	16	21	30	37	43	60	60	0.91	3.29
	Unknown	151	38.19	10.35	15	19	30	39	45	60	60	0.91	3.11
Language Fluency	English Only	6423	37.08	10.81	15	20	29	37	43	60	60	0.91	3.24
	Initially-Fluent English Proficient	277	33.36	10.37	15	18	26	33	39	54	60	0.91	3.11
	English Learner	3272	36.04	10.37	15	20	29	36	42	54	60	0.91	3.11
	Reclassified-Fluent English Proficient	176	36.35	9.49	19	22	29	36.5	42	54	60	0.88	3.29
	Unknown	205	40.01	10.41	16	22	33	39	45	60	60	0.89	3.45
Economic Disadvantage	Yes	6197	36.30	10.40	15	20	29	36	43	60	60	0.91	3.12
	No	3783	37.11	10.99	15	20	29	37	43	60	60	0.91	3.30
	Unknown	373	39.09	11.19	15	20	32	39	45	60	60	0.91	3.36

Scale Score Distributions*: Level IV English–Language Arts

ELA Level IV	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	5163	35.24	9.99	16	20	28	34	41	54	60	0.90	3.16
	Hard of Hearing	75	36.09	10.81	15	20	28	35	42	60	60	0.91	3.24
	Deafness	175	34.54	6.28	15	24	30	35	39	43	48	0.84	2.51
	Speech/Language Impairment	286	43.15	9.48	24	28	38	43	48	60	60	0.86	3.55
	Visual Impairment	88	34.03	10.01	15	16	27	35	41	48	60	0.90	3.17
	Emotional Disturbance	102	43.26	10.24	20	26	36	42	48	60	60	0.88	3.55
	Orthopedic Impairment	621	35.04	9.81	15	20	28	34	41	54	60	0.90	3.10
	Other Health Impairment	359	40.04	10.88	19	22	32	39	48	60	60	0.90	3.44
	Specific Learning Impairment	877	45.05	8.98	24	31	39	43	54	60	60	0.82	3.81
	Multiple group	401	34.56	10.27	15	19	27	34	41	54	60	0.91	3.08
	Autism	1879	35.78	11.51	15	17	27	36	43	60	60	0.93	3.05
	Traumatic Brain Injury	60	39.37	10.31	15	26	32	37	45	60	60	0.87	3.72
	Unknown	259	39.17	10.70	18	22	32	38	45	60	60	0.88	3.71

*Results for groups with fewer than 11 members are not reported

Table 4.27 Scale Score Distributions*: Level IV Mathematics

Scale Score Distributions*: Level IV Mathematics													
Mathematics Level IV	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	10337	35.41	11.01	15	19	28	34	43	60	60	0.90	3.48
Grade	6	3384	34.15	10.42	15	19	27	33	40	52	60	0.89	3.46
	7	3568	35.68	11.02	15	19	28	35	43	60	60	0.90	3.48
	8	3385	36.40	11.45	15	19	28	35	43	60	60	0.90	3.62
Gender	Male	6534	35.66	11.24	15	19	28	35	43	60	60	0.90	3.55
	Female	3753	34.99	10.60	15	19	28	34	40	60	60	0.89	3.52
	Unknown	50	35.62	11.37	15	19	28	34	43	52	60	0.90	3.60
Race Ethnicity	American Indian or Alaska Native	95	35.14	11.58	16	19	27	34	43	60	60	0.91	3.47
	Asian	651	32.37	10.26	15	18	26	32	37	52	60	0.89	3.40
	Pacific Islander	50	33.74	9.72	15	18	27	34.5	40	46	60	0.89	3.22
	Filipino	260	32.95	10.16	15	19	27	32	39	52	60	0.88	3.52
	Hispanic or Latino	4915	35.58	10.92	15	19	28	35	43	60	60	0.89	3.62
	African American	1195	36.52	11.38	15	19	28	35	43	60	60	0.90	3.60
	White	3020	35.56	11.08	15	19	28	35	43	60	60	0.90	3.50
	Unknown	151	36.52	11.66	15	17	29	35	43	60	60	0.90	3.69
Language Fluency	English Only	6411	35.46	11.12	15	19	28	34	43	60	60	0.90	3.52
	Initially-Fluent English Proficient	277	32.59	10.80	15	17	24	30	39	52	60	0.91	3.24
	English Learner	3268	35.49	10.87	15	19	28	35	43	60	60	0.89	3.61
	Reclassified-Fluent English Proficient	175	33.85	9.48	16	20	28	34	39	52	60	0.88	3.28
	Unknown	206	37.91	10.40	17	22	31	36.5	43	60	60	0.85	4.03
Economic Disadvantage	Yes	6186	35.90	11.12	15	19	28	35	43	60	60	0.90	3.52
	No	3778	34.49	10.78	15	19	28	34	40	60	60	0.89	3.58
	Unknown	373	36.72	11.01	15	19	30	35	43	60	60	0.88	3.81
Primary Disability	Mental Retardation	5154	34.17	10.21	16	19	28	33	39	52	60	0.89	3.39
	Hard of Hearing	74	34.92	10.36	15	19	28	35	40	60	60	0.89	3.44
	Deafness	175	36.59	9.37	17	23	30	36	40	60	60	0.82	3.98

Scale Score Distributions*: Level IV Mathematics													
Mathematics Level IV	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
	Speech/Language Impairment	284	42.04	9.70	20	29	35	40	49	60	60	0.76	4.75
	Visual Impairment	87	33.92	11.11	15	17	27	33	43	52	60	0.90	3.51
	Emotional Disturbance	102	45.25	10.89	16	26	39	43	52	60	60	0.88	3.77
	Orthopedic Impairment	620	34.48	11.00	15	19	27	34	40	60	60	0.90	3.48
	Other Health Impairment	359	38.51	11.20	17	20	31	37	46	60	60	0.89	3.71
	Specific Learning Impairment	877	45.24	9.48	23	32	39	43	52	60	60	0.76	4.64
	Multiple group	402	33.05	10.37	15	19	26	32	39	52	60	0.90	3.28
	Autism	1876	32.46	10.89	15	17	24	31	37	60	60	0.90	3.44
	Traumatic Brain Injury	60	40.00	11.11	15	22	34	39	49	60	60	0.88	3.85
	Unknown	259	37.75	10.95	15	22	31	36	43	60	60	0.87	3.95

*Results for groups with fewer than 11 members are not reported

Table 4.28 Scale Score Distributions*: Level V English–Language Arts

Scale Score Distributions*: Level V English–Language Arts													
ELA Level V	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	9367	38.39	10.57	16	22	31	38	45	57	60	0.91	3.17
Grade	9	3206	37.67	10.27	16	22	30	38	45	57	60	0.91	3.08
	10	3250	38.74	10.52	16	23	31	38	45	57	60	0.91	3.16
	11	2911	38.80	10.91	16	22	31	38	47	60	60	0.92	3.09
Gender	Male	5737	38.42	10.45	16	22	31	38	45	57	60	0.91	3.14
	Female	3587	38.31	10.74	16	22	30	38	45	57	60	0.92	3.04
	Unknown	43	41.81	11.83	20	22	34	41	50	60	60	0.90	3.74
Race Ethnicity	American Indian or Alaska Native	68	42.18	10.48	21	23	35.5	41	48.5	60	60	0.88	3.63
	Asian	608	36.30	10.15	15	21	29	36.5	43	57	60	0.92	2.87
	Pacific Islander	41	38.37	11.94	15	21	28	37	47	57	60	0.94	2.92
	Filipino	238	36.72	10.35	15	21	29	37	43	57	60	0.91	3.11
	Hispanic or Latino	4129	37.67	10.36	18	22	30	37	45	57	60	0.91	3.11
	African American	1201	39.05	10.12	18	23	32	40	45	57	60	0.90	3.20
	White	2958	39.52	10.93	15	22	32	40	47	60	60	0.92	3.09
	Unknown	124	40.43	11.07	18	21	34	41	47	60	60	0.92	3.13
Language Fluency	English Only	5977	38.88	10.70	16	22	31	40	45	57	60	0.92	3.03
	Initially-Fluent English Proficient	321	36.79	11.14	16	21	28	37	43	60	60	0.92	3.15
	English Learner	2606	37.43	10.08	18	22	30	37	43	57	60	0.91	3.02
	Reclassified-Fluent English Proficient	230	37.03	10.61	15	22	29	37	45	57	60	0.91	3.18
	Unknown	233	40.16	10.64	15	22	34	41	47	60	60	0.92	3.01
Economic Disadvantage	Yes	5304	38.05	10.41	16	22	30	38	45	57	60	0.91	3.12
	No	3702	38.66	10.74	15	22	31	38	45	57	60	0.92	3.04
	Unknown	361	40.65	10.89	15	23	34	41	47	57	60	0.92	3.08

Scale Score Distributions*: Level V English–Language Arts													
ELA Level V	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	5304	37.54	10.19	16	22	30	37	45	57	60	0.91	3.06
	Hard of Hearing	78	36.44	8.57	15	22	30	38	41	50	57	0.90	2.71
	Deafness	135	36.62	7.92	23	24	31	36	41	50	60	0.83	3.27
	Speech/Language Impairment	157	44.08	8.33	23	28	40	43	50	60	60	0.81	3.63
	Visual Impairment	91	35.75	10.26	15	22	28	36	43	57	60	0.91	3.08
	Emotional Disturbance	163	46.01	10.14	16	26	41	47	57	60	60	0.90	3.21
	Orthopedic Impairment	497	36.97	11.14	15	21	28	36	45	57	60	0.93	2.95
	Other Health Impairment	299	41.23	9.77	18	23	35	41	47	57	60	0.88	3.38
	Specific Learning Impairment	812	45.83	8.68	23	32	40	45	50	60	60	0.81	3.78
	Multiple group	419	34.95	10.93	15	20	26	34	43	57	60	0.93	2.89
	Autism	1144	36.41	10.88	15	20	28	37	43	57	60	0.92	3.08
	Traumatic Brain Injury	90	41.69	10.36	23	24	35	41	47	60	60	0.90	3.28
	Unknown	175	39.62	10.54	16	22	33	40	47	57	60	0.91	3.16

*Results for groups with fewer than 11 members are not reported

Table 4.29 Scale Score Distributions*: Level V Mathematics

Scale Score Distributions*: Level V Mathematics													
Mathematics Level V	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
All	All	9347	35.32	10.09	15	21	27	35	41	57	60	0.91	3.03
Grade	9	3194	34.86	9.97	15	20	27	35	41	57	60	0.91	2.99
	10	3246	35.5	9.87	15	22	29	35	41	57	60	0.91	2.96
	11	2907	35.64	10.44	15	21	27	35	42	57	60	0.91	3.13
Gender	Male	5722	35.91	10.42	15	21	27	36	42	57	60	0.91	3.13
	Female	3582	34.34	9.43	15	21	27	34	40	49	60	0.90	2.98
	Unknown	43	38.65	11.62	19	21	32	39	46	57	60	0.94	2.85
Race Ethnicity	American Indian or Alaska Native	68	39.25	10.98	17	23	30.5	38	46	60	60	0.91	3.29
	Asian	608	34.07	10.11	15	19	26	33	40	57	60	0.91	3.03
	Pacific Islander	41	34.71	9.87	15	19	25	37	42	46	60	0.93	2.61
	Filipino	238	33.61	9.98	15	18	26	33	40	57	60	0.91	2.99
	Hispanic or Latino	4121	35.19	10.00	15	22	27	35	41	57	60	0.91	3.00
	African American	1198	35.27	9.30	17	22	29	35	40	57	60	0.89	3.08
	White	2950	35.81	10.46	15	21	27	36	42	57	60	0.91	3.14
	Unknown	123	36.14	10.08	15	21	29	37	42	57	60	0.91	3.02
Language Fluency	English Only	5965	35.32	10.14	15	21	27	35	41	57	60	0.91	3.04
	Initially-Fluent English Proficient	321	34.28	10.34	15	22	25	34	40	57	60	0.91	3.10
	English Learner	2603	35.31	9.83	15	22	27	35	41	57	60	0.91	2.95
	Reclassified-Fluent English Proficient	228	35.21	10.84	17	21	27	35	41	57	60	0.92	3.07
	Unknown	230	37.16	10.53	15	22	31	37	42	60	60	0.91	3.16
Economic Disadvantage	Yes	5295	35.34	9.94	15	22	27	35	41	57	60	0.91	2.98
	No	3693	35.13	10.30	15	20	27	35	41	57	60	0.91	3.09
	Unknown	359	37.02	9.95	15	22	30	38	42	57	60	0.91	2.99

Scale Score Distributions*: Level V Mathematics

Mathematics Level V	Subgroup	N	Mean	SD	Mean at Percentiles							Alpha	SEM
					1	5	25	50	75	95	99		
Primary Disability	Mental Retardation	5299	34.10	9.19	15	22	27	34	40	49	60	0.90	2.91
	Hard of Hearing	78	35.26	9.67	15	21	29	35	41	57	60	0.90	3.06
	Deafness	135	36.96	8.62	21	23	31	37	42	49	60	0.86	3.23
	Speech/Language Impairment	157	42.96	9.36	23	27	37	42	46	60	60	0.85	3.63
	Visual Impairment	91	32.38	8.64	15	20	25	33	38	49	57	0.89	2.87
	Emotional Disturbance	162	43.14	9.94	15	26	38	42	49	60	60	0.88	3.44
	Orthopedic Impairment	494	33.24	9.90	15	18	25	33	40	49	60	0.91	2.97
	Other Health Impairment	299	38.06	9.60	15	22	33	38	42	57	60	0.88	3.33
	Specific Learning Impairment	809	44.27	9.18	23	31	38	42	49	60	60	0.83	3.79
	Multiple group	418	31.68	9.67	15	17	24	31	38	49	60	0.92	2.74
	Autism	1141	33.51	10.80	15	18	24	33	40	57	60	0.92	3.05
	Traumatic Brain Injury	89	39.72	10.51	23	24	33	39	46	57	60	0.90	3.32
	Unknown	172	36.76	9.72	15	22	31	36	42	57	60	0.90	3.07

*Results for groups with fewer than 11 members are not reported

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational And Psychological Testing*. Washington, DC: American Educational Research Association, 1999.

Crocker, L. and Algina, J. *Introduction to Classical and Modern Test Theory*. New York: Holt, 1986.

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

Dorans, N. J., and Holland, P. W. "DIF Detection and Description: Mantel-Haenszel and Standardization," in *Differential Item Functioning*. Edited by P. W. Holland and H. Wainer. Hillsdale, NJ: Lawrence Erlbaum, 1993, pp. 35–66.

Dorans N.J., and Schmitt A.P. "Constructed Response and Differential Item Functioning: A Pragmatic Approach," in *Construction Versus Choice in Cognitive Measurement*. Edited by R.E. Bennett and W.C. Ward. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993, pp.135–165.

Drasgow F. "Polychoric and Polyserial Correlations," in *Encyclopedia of Statistical Sciences*, Vol. 7. Edited by L. Kotz, N. L. Johnson. New York: Wiley, 1988, pp. 69–74.

Educational Testing Service. *ETS Standards for Quality and Fairness*. Princeton, NJ: Author, 2002.

Feldt, L. S. and Brennen, R. L. "Reliability," in *Educational Measurement*. Edited by R.L. Linn. New York: Macmillan, 1989.

Holland, P.W. and Thayer, D.T. "An Alternative Definition of the ETS Delta Scale of Item Difficulty." RR-85–43, 1985.

Livingston, S.A., and Lewis, C. "Estimating the Consistency and Accuracy of Classification Based on Test Scores." *Journal on Educational Measurement* Vol. 32 (1995), pp. 179–197.

Lord, F.M. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 1980.

Mantel, N. "Chi-square Tests with One Degree of Freedom: Extensions of the Mantel-Haenszel Procedure." *Journal of the American Statistical Association*, Vol. 58 (1963), pp. 690–700.

Mantel, N, and Haenszel, W. "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease." *Journal of the National Cancer Institute*, Vol. 22 (1959), pp. 719–748.

Muraki, E. "A Generalized Partial Credit Model: Application of an EM Algorithm." *Applied Psychological Measurement* Vol. 16 (1992), pp. 159–176.

Muraki, E. and Bock, R.D. “PARSCALE” [Computer software]. Chicago, IL: Scientific Software, Inc., 1999.

No Child Left Behind Act of 2001, 20 U.S.C. 6301 et seq (PL 107–110), 2001.

Stocking, M.L., and Lord, F.M. “Developing a Common Metric in Item Response Theory.” *Applied Psychological Measurement* (1983), pp. 201–210, 1983.

Appendix A. Individual Item Statistics

Table A.1 2006 CAPA Item Statistics: Level I

2006 CAPA Item Statistics: Level I					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level I Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
English–Language Arts	01	1	3.18	.82	
	01	2	3.20	.79	
	01	3	2.57	.83	
	01	4	2.30	.84	
	01	5	3.55	.71	
	01	6	2.86	.83	
	01	7	2.22	.81	
	01	8	3.28	.82	
	01/06	9	3.28	.80	
	01/02/03/ 04/05/06*	10	3.86	.75	
	02/05*	9	3.03	.74	
	03	9	2.73	.76	
	04	9	2.97	.77	
Mathematics	01	11	3.02	.85	
	01	12	2.75	.83	
	01	13	2.90	.83	
	01	14	2.28	.81	
	01	15	2.48	.87	
	01	16	2.87	.80	
	01	17	2.68	.86	
	01	18	2.75	.83	
	01/03/05*	19	2.59	.80	
	01/03/05*	20	2.64	.73	
	02/04/06*	19	2.56	.78	
	02/04/06*	20	2.47	.80	
Science	01	21	2.86	.82	
	01	22	3.45	.77	
	01	23	2.73	.85	
	01	24	2.59	.89	
	01	25	2.64	.88	
	01	26	3.24	.84	
	01	27	2.56	.88	
	01	28	2.40	.85	
	01	29	2.76	.79	
	01	30	3.07	.82	

2006 CAPA Item Statistics: Level I					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level I Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
	02	29	2.80	.71	
	02	30	2.25	.77	
	03	29	2.62	.81	
	03	30	2.63	.68	
	04	29	2.74	.87	
	04	30	3.06	.76	
	05	29	2.69	.82	
	05	30	2.59	.81	
	06	29	2.55	.85	
	06	30	2.47	.79	

*This item appeared on more than one field test form.

Table A.2 2006 CAPA Item Statistics: Level II

2006 CAPA Item Statistics: Level II					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level II Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
English–Language Arts	01	1	2.70	.78	
	01	2	2.48	.77	
	01	3	3.17	.75	
	01	4	2.43	.78	
	01	5	2.99	.78	
	01	6	3.23	.80	H
	01	7	2.59	.77	
	01	8	3.06	.79	
	01	9	1.95	.69	
	01	10	3.60	.73	H
	02	9	3.10	.75	
	02	10	2.51	.76	
	03	9	2.97	.75	
	03/05/06*	10/9/10*	3.21	.76	H
	04	9	3.58	.71	H
	04/05*	10	3.11	.72	
	06	9	3.28	.74	H
Mathematics	01	11	3.04	.84	
	01	12	2.06	.79	
	01	13	3.01	.81	
	01	14	2.70	.79	
	01	15	2.59	.79	
	01	16	2.04	.75	
	01	17	2.91	.67	
	01	18	2.71	.78	
	01/03*	19	3.20	.75	
	01/05*	20/19*	1.74	.70	
	02	19	3.40	.72	H
	02	20	3.00	.74	
	03	20	3.14	.64	
	04/06*	19/20*	3.56	.67	H
	04/06*	20/19*	3.09	.70	
	05	20	3.39	.68	H

*This item appeared on more than one field test form.

Table A.3 2006 CAPA Item Statistics: Level III

2006 CAPA Item Statistics: Level III					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level III Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
English–Language Arts	01	1	2.62	.79	
	01	2	2.87	.85	
	01	3	2.17	.74	
	01	4	2.87	.72	
	01	5	2.85	.84	
	01	6	2.66	.76	
	01	7	2.38	.85	
	01	8	2.43	.83	
	01	9	3.47	.65	H
	01	10	2.69	.78	
	02	9	2.84	.70	
	02	10	3.38	.71	H
	03	9	2.75	.76	
	03	10	3.57	.78	H
	04	9	1.80	.65	
	04	10	3.11	.68	
	05	9	2.99	.66	
	05	10	3.37	.78	H
	06	9	2.30	.51	R
	06	10	3.00	.67	
Mathematics	01	11	3.51	.79	H
	01	12	2.50	.84	
	01	13	2.43	.76	
	01	14	2.87	.86	
	01	15	3.14	.73	
	01	16	3.12	.83	
	01	17	3.06	.86	
	01	18	3.04	.76	
	01	19	3.59	.78	H
	01	20	3.05	.67	
	02	19	3.13	.80	
	02	20	2.29	.72	
	03	19	2.33	.62	
	03	20	3.14	.77	
	04	19	1.96	.68	
	04	20	2.33	.78	
	05	19	2.77	.72	
	05	20	3.61	.75	H
	06	19	3.12	.70	
	06	20	2.54	.74	

2006 CAPA Item Statistics: Level III					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level III Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
Science	01	21	2.82	.76	
	01	22	2.73	.66	
	01	23	3.03	.73	
	01	24	3.08	.82	
	01	25	2.87	.81	
	01	26	2.71	.80	
	01	27	2.85	.82	
	01	28	3.32	.82	H
	01	29	3.24	.76	H
	01	30	3.57	.78	H
	02	29	3.25	.72	H
	02	30	3.03	.71	
	03	29	3.45	.70	H
	03	30	3.39	.71	H
	04	29	3.17	.74	
	04	30	2.36	.64	
	05	29	2.82	.77	
	05	30	2.76	.71	
	06	29	3.08	.71	
	06	30	2.79	.65	

Table A.4 2006 CAPA Item Statistics: Level IV

2006 CAPA Item Statistics: Level IV					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level IV Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
English–Language Arts	01	1	3.13	.76	
	01	2	2.23	.87	
	01	3	2.35	.75	
	01	4	2.38	.87	
	01	5	2.36	.86	
	01	6	2.85	.82	
	01	7	2.82	.74	
	01	8	3.03	.88	
	01	9	3.12	.69	
	01	10	3.18	.84	
	02	9	3.50	.63	H
	02	10	2.97	.68	
	03	9	3.37	.73	H
	03/05*	10	2.80	.79	
	04/06*	9	2.99	.74	
	04	10	2.67	.77	
	05/06*	9/10*	2.38	.65	
Mathematics	01	11	2.49	.82	
	01	12	3.19	.72	
	01	13	2.69	.77	
	01	14	2.65	.87	
	01	15	2.61	.83	
	01	16	2.65	.78	
	01	17	3.05	.80	
	01	18	2.63	.74	
	01	19	2.97	.80	
	01	20	3.35	.79	H
	02	19	3.11	.61	
	02	20	2.19	.78	
	03/04*	19	2.92	.77	
	03/04*	20	3.22	.75	H
	05	19	2.76	.66	
	05	20	3.40	.77	H
	06	19	2.51	.80	
	06	20	2.11	.73	

2006 CAPA Item Statistics: Level IV					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level IV Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
Science	01	21	2.74	.82	
	01	22	2.77	.71	
	01	23	2.58	.78	
	01	24	2.43	.73	
	01	25	2.83	.79	
	01	26	2.65	.80	
	01	27	2.47	.77	
	01	28	2.88	.83	
	01	29	3.45	.70	H
	01	30	3.46	.76	H
	02	29	2.84	.69	
	02	30	3.26	.71	H
	03	29	3.05	.78	
	03	30	2.65	.73	
	04	29	2.98	.63	
	04	30	3.20	.72	
	05	29	3.23	.71	H
	05	30	2.14	.68	
	06	29	3.24	.77	H
	06	30	2.85	.79	

*This item appeared on more than one field-test form.

Table A.5 2006 CAPA Item Statistics: Level V

2006 CAPA Item Statistics: Level V					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level V Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
English–Language Arts	01	1	2.02	.81	
	01	2	2.97	.84	
	01	3	2.64	.80	
	01	4	2.13	.82	
	01	5	2.86	.84	
	01	6	3.00	.75	
	01	7	3.22	.87	H
	01	8	3.19	.86	
	01/05*	9	3.37	.82	H
	01/04/06*	10/10/9*	2.22	.80	
	02/04*	9	3.16	.78	
	02	10	2.69	.76	
	03	9	2.83	.81	
	03	10	3.46	.69	H
	05/06*	10	2.72	.79	
Mathematics	01	11	2.77	.82	
	01	12	2.22	.86	
	01	13	2.96	.83	
	01	14	2.51	.76	
	01	15	2.64	.82	
	01	16	1.77	.87	
	01	17	2.50	.83	
	01	18	2.11	.83	
	01/04*	19	2.96	.82	
	01	20	1.98	.83	
	02/05*	19	2.55	.81	
	02/05*	20	2.79	.82	
	03/06*	19	2.28	.81	
	03/06*	20	2.93	.78	
	04	20	2.91	.83	

2006 CAPA Item Statistics: Level V					
Flag Values: A = Low Average Item Score • **R** = Low Correlation with Criterion • **O** = High Percent of Omits/Not Responding • **H** = High Average Item Score					
Level V Content	**Version**	**Item Number**	**AIS**	**Polyserial**	**Flag**
Science	01	21	3.08	.80	
	01	22	2.89	.77	
	01	23	3.16	.81	
	01	24	2.94	.78	
	01	25	2.15	.80	
	01	26	3.02	.86	
	01	27	2.66	.80	
	01	28	3.03	.82	
	01/02/05*	29	2.83	.66	
	01/02/03*	30	2.52	.63	
	03/04/06*	29	2.81	.66	
	04/05/06*	30	2.19	.78	

*This item appeared on more than one field-test form.