TR-07-04

HumRRO

Human Resources Research Organization

Independent Evaluation of the Alignment of the California Standards Tests (CSTs) and the California Alternate Performance Assessment (CAPA)

Leslie R. Taylor
Lauress L. Wise
Arthur A. Thacker
Robin L. Moody
Lisa E. Koger
Emily R. Dickinson
D. Matthew Trippe

Prepared for: **California Department of Education**
Sacramento, CA
HumRRO Contract Number: 07-01

April 17, 2007

66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314 • 703.549.3611 • fax. 703.549.9025
www.humrro.org

Executive Summary

Independent Evaluation of the Alignment of the California Standards Tests (CSTs) and the California Alternate Performance Assessment (CAPA)

The California Department of Education (CDE) issued a request for proposal to conduct an external, independent evaluation of the California standards and assessment system. The Human Resources Research Organization (HumRRO) was awarded a contract to conduct this evaluation, and work began on October 27, 2006. The evaluation included two main tasks. The first task, detailed in this report, was an independent review of the alignment of the assessments used for school and district accountability with the California content standards. The second activity required the development of descriptions of each achievement level corresponding with the assessments. The details of the Performance Level Descriptors (PLD) task are reported separately (Wise, Taylor, Becker, Gladden, Handy, Thacker et al., 2007).

This independent evaluation of the California assessment system developed in response to results of California's peer review from the United States Department of Education (USDE). The USDE requested that California provide independent evidence of alignment between the assessments used to calculate Adequate Yearly Progress (AYP) and the State content standards on which those assessments are based. Alignment results should demonstrate that the assessments represent the full range of the content standards, and that these assessments measure student knowledge in the same manner and at the same level of complexity as expected in the content standards.

This requirement by the USDE developed from the No Child Left Behind (NCLB) Act of 2001. NCLB is a Federal education act that challenges each state to establish a coherent assessment system based on solid academic standards. All states receiving Title I funds must present evidence of establishing a fair and consistent assessment system that is based on rigorous standards, sufficient alignment between standards and assessments, and high-quality educational results. Concerning alignment, all aspects of the state assessment system must coincide, including the academic content standards, achievements standards (linked to cut scores), performance level descriptors, and each assessment.

In previous years, California conducted alignment studies to examine the match between the state assessments and the content standards for the California Standards Tests (CSTs) and the California High School Exit Exam (CAHSEE). HumRRO conducted an independent alignment review of the CAHSEE and standards in 2005, and this evidence was submitted and approved by the USDE. An alignment review of the Standardized Testing and Reporting Program compared with the content standards was conducted in 2002, but this review occurred prior to the current assessment configuration and included a norm-referenced assessment. In addition, the first administration of the new science assessments for Grades 8 and 10 occurred in 2006. The California Alternate Performance Assessment (CAPA), originally administered in

2003, is being updated substantially for the 2007 assessment. A new review is needed at this time to provide an independent evaluation of content alignment of the CSTs and of the CAPA.

In addition, the CSTs have not undergone an independent review of the alignment of the assessments to the achievement levels used in reporting[1]. The USDE also requires evidence that the assessments are aligned to the achievement levels, meaning that they provide accurate information for students scoring at different levels. The current evaluation includes an assessment alignment to the performance standards that define the achievement levels as well as alignment to the content standards.

Alignment of the Assessments to the California Content Standards

The alignment workshops conducted to evaluate the CSTs and the CAPA were held November 28 through December 3, 2007. Panelists recruited to participate in the workshops are current California educators or educational consultants. A total of 62 panelists participated in the workshops. Fifty-two panelists reviewed the CSTs and eight reviewed the CAPA. These panelists were nominated by district and test coordinators from across the State of California. Selection of nominees was made by CDE and the State Board of Education (SBE).

HumRRO applied the Webb alignment method to collect and analyze the alignment data. This alignment method has undergone substantial research (e.g., Webb, 1997; 1999; 2005), and it has been used successfully in approximately 15 other states and in two other countries. HumRRO used the standard Webb method to evaluate the alignment of the 2006 CSTs to the California Content Standards. For the CAPA performance tasks, HumRRO used a version of the Webb method specifically designed for alternate assessments to evaluate performance tasks or portfolios.

Webb uses slightly different terminology compared to California to refer to levels of the content standards documents. Specifically, Webb applies the term *standard* to mean the highest, most general level of the content expectations (in place of *domain* or *reporting category*). The results of the analyses are reported at this standard (domain) level, meaning how well the test items align with each of these broad content categories (Webb, 2005). Webb uses the term *content objective* to reference the most specific level of content expectations. In some cases, content standards documents include an additional, intermediate level of organization between the *standard* and the *objective*, which Webb refers to as a *goal*. A *goal* level (comparable to *strand* in the California content standards) usually explains the general expectations for a group of related content objectives. For example, the California Content Standards for Grade 5 science include a *standard* called Physical Science. One *goal* under this standard is the expectation that students will understand that “Elements and their combinations account for all the varied types of matter in the world”. A specific *content objective* under this

[1] Currently, there is no Federal requirement for the development of PLDs for alternate assessments, partly because alternate assessments are supposed to be aligned directly with the primary content standards as well as any alternate content standards.

goal is that “Students know that during chemical reactions the atoms in the reactants rearrange to form products with different properties.”. These terms will be used in this way throughout the report since the Webb alignment method has been applied.

The Webb method includes four major criteria to evaluate alignment. These criteria link with statistical procedures used to assess how well individual portions of the assessments and standards documents actually match. The four alignment criteria are as follows: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation.

Categorical concurrence is a basic measure of alignment between content standards and test items. This term refers to the proportion of overlap between the content stated in the standards document and that assessed by items on the test. Webb suggests that the mean number of items per standard should be at least six for acceptable content coverage.

Depth-of-Knowledge (DOK) measures the type of cognitive processing required by items and content standards. For example, is a student expected to simply identify or recall basic facts, or is the student expected to use reasoning by manipulating information or strategizing? In mathematics, a student may be asked to identify the appropriate use of a decimal among several answer choices. This task should be less complex than trying to explain the concept of a decimal and how and why it can be moved. In English-language arts, asking a student to identify Greek mythology requires less processing compared with asking a student to use knowledge of Greek mythology to understand the origin and meaning of new words.

The purpose of using depth-of-knowledge as a measure of alignment is to determine whether a test item (or performance task) and corresponding standard are both written at the same level of cognitive complexity. Reviewers make two separate judgments about cognitive complexity, one for the standard and one for the item. These two judgments are compared to determine whether the item is written at the same level as the standard to which it is linked. Results are summarized in terms of the percent of items with cognitive complexity ratings at or above (more complex than) the rating of the selected content objective. Webb’s suggests that at least 50% of the items should have complexity ratings at or above the level of the corresponding content objective. Webb refers to his comparison as *Depth-of-Knowledge Consistency*.

Another measure examines the ***range-of-knowledge correspondence*** between the assessment and content standards. The range-of-knowledge measure looks at the breadth of knowledge represented by test items in greater detail. Categorical concurrence simply notes whether a sufficient number of items on the test covers each general content topic (individual strands). However, states generally lay out more specific *content objectives*, or standards, under each strand. The range indicates the number of content objectives assessed by items. Webb’s minimum level of acceptability for range of correspondence is 50% per standard. This means that at least 50% of the objectives must be matched to one or more items.

Finally, the ***balance-of-representation*** criterion focuses on content coverage in yet more detail. In this case, the number of items matched to the content objective does matter. The balance-of-knowledge representation indicator determines whether the assessment measures the content objectives equitably within each standard. Based on Webb's method, items should be distributed evenly across the objectives per standard for good balance. The balance-of-knowledge representation is determined by calculating an index, or score, for each standard. Each standard should meet or surpass a minimum index level to demonstrate adequate balance. According to Webb, the minimum acceptable index for a single standard is 70 (on a scale of 0 to 100, with 100 representing perfect balance).

Each criterion provides different information about the degree of alignment between the assessment and content standards. However, all four of Webb's criteria must be considered for a complete picture of alignment.

The overall alignment results for both the CSTs and for the CAPA were good. However, some subject areas may require reconsideration to improve the quality of alignment, particularly for the CST assessments for mathematics and for history-social science.

Alignment of the CSTs to the Performance Standards

After analyzing the alignment of each test form to the corresponding *content* standards, HumRRO reviewed each of the 2006 assessments included in this study for alignment with the *performance* standards. The key question addressed by these analyses is whether the tests provide useful information about students at each achievement or performance level. The CSTs results are reported in terms of five performance levels ranging from Far Below Basic up to Advanced.

In developing descriptors for each performance level, we (Wise, et al. 2007) used item maps developed by Educational Testing Service (ETS). The item maps assigned each item to the lowest performance level at which most of the students could answer the item correctly. For the present analyses, we ensured that every operational item was assigned to a performance level, specifically to the lowest performance level at which 60% or more of the students answered correctly. Note that, in a few cases, fewer than 60% of students at the advanced level answered items correctly. Thus, for purposes of this mapping, we created an additional, "Beyond Advanced" category to count these items.

The analysis of the item maps and the review of error of measurement data demonstrate a reasonable coverage of the performance levels defined by California's achievement standards. Several tests could benefit from a further increase in accuracy, particularly Grade 8 science.

Summary and Recommendations

The results of these reviews provide confirmation of the content validity of the CSTs and the CAPA for California overall. Most aspects of the assessments under review aligned to the content standards. These results offer reasonable evidence to the USDE that California clearly has established a rigorous and coherent assessment system for all students. HumRRO did find that some aspects of the assessments, particularly for specific subject areas and grade levels, could benefit from additional review by CDE and the test developer to improve alignment. As a result, HumRRO offers the following recommendations to CDE for alignment improvement.

Recommendations for the CSTs and California Content Standards

1. **Review the cognitive requirements (depth-of-knowledge) of the assessment items and the content standards to establish greater consistency.** This recommendation pertains to English-language arts (ELA) Grade 6 and 8; math Grades 2 and 7; the general math test; all three integrated math tests; and, all three history-social science tests. Increasing depth-of-knowledge consistency can be accomplished by modifying existing operational items and/or by modifying content expectations of the standards. Given that the content standards underwent thorough review prior to Board approval, working with the test contractor to bring the current operational items more in line with the standards is a reasonable course of action. Furthermore, while modifying the content standards may be appropriate in some cases, California should be cautious about reducing the cognitive demands of its content expectations. If California does choose to revise the content standards at some point, it may be worthwhile to evaluate the content standards of other states whose assessment systems have been approved by the USDE to compare cognitive expectations. Alternatively, CDE and SBE could examine the structure of the content frameworks for the National Assessment of Educational Progress (NAEP). A number of states (e.g., Georgia, Kentucky, Missouri) have revised their content standards to model the NAEP content frameworks successfully.

2. **Expand the content coverage on the assessments to match the breadth of the content expectations in California Content Standards.** This recommendation pertains to the mathematics tests for Grades 2 through 5, the integrated math tests, and the history-social science tests. In evaluating the test blueprints, the narrow range of content coverage seems to stem from the limited number of items targeted for assessment in the first place. Necessarily, standardized assessments must limit the total number of items included on a single test form. Thus, HumRRO does not expect CDE and the test developer to lengthen the test to increase content coverage. Instead, several strategies working within the existing test forms may be possible: (a) redistribute items to increase content coverage on some standards; (b) consider whether some content is appropriate for standardized assessment or

could be assessed in the classroom; or (c) consider modifying or merging related content objectives to increase the number of items targeting a given content area.

Recommendations for the CAPA and the Alternate Content Standards

HumRRO recommends that CDE and SBE consider the following recommendations for the CAPA based on the outcomes of the alignment review and analyses:

1. **Review the appropriateness of the number of content objectives for the alternate standards.** One of the challenges of alternate assessments and standards is condensing and modifying the content expectations developed for the regular assessment to more appropriately evaluate special needs students. At the same time, the alternate assessment should not be reduced to the extent that the expectations are entirely different from those laid out for the regular assessment. California appears to have made good progress on achieving this goal by including a reasonable set of content expectations linked to the full content standards. However, it may be the case that further review is necessary to consider the quantity of content objectives currently in place, particularly for ELA Levels I and II and Math Levels II and III.

2. **Review the cognitive requirements (depth-of-knowledge) of the performance tasks and the alternate standards to establish greater consistency.** This recommendation applies specifically to ELA Level I (Reading and Listening/Speaking) and Math Level I (Statistics, Probability, and Data Analysis). Both the new performance tasks and the standards should be evaluated together to determine the appropriate degree of content expectations for students at this level.

Recommendations for the CSTs and Performance Levels

Coverage of the performance levels by test items was generally good for each of the CSTs, particularly for the Proficient and Basic categories. A few areas may benefit from further improvements, however. Some specific suggestions include:

1. **Review the assessments for Grade 8 science and Integrated Mathematics III for test accuracy due to larger standard errors of measurement**. To ensure that these tests measure student performance as accurately as possible, CDE should consider whether the present criteria established for the performance levels are appropriate. Two approaches may be useful in making this decision. First, the newly developed performance level descriptors (Wise et al., 2007) could be used to target item development to each performance level more distinctly. Alternatively, stricter standards might be established for test accuracy curves generated from field test information when new test forms are assembled.

2. **Review the number of items assigned to Far Below Basic and Below Basic to distinguish between these performance levels more clearly for each subject area.** Currently, many of the tests include a limited number of items not only at the Far Below Basic level but also at the Below Basic level. If these distinctions should be retained, assigning more items, at least to the Below Basic level, would be helpful to more accurately determine student performance at this level.

3. **Examine the number of items assigned to the Advanced level for ELA, math, and science.** Some grades and subject areas also include a limited number of items assigned to assess performance at the Advanced level. For ELA Grade 3 and for math Grade 4, the number of items assigned to the Advanced level is limited. For science Grades 8 and 10, Integrated Mathematics II and III, and for Algebra I, some items also appear to assess student knowledge beyond the Advanced level. Again, the new performance level descriptors might be used to improve the targeting of items to this performance level.

INDEPENDENT EVALUATION OF THE ALIGNMENT OF THE CALIFORNIA STANDARDS TESTS (CSTS) AND THE CALIFORNIA ALTERNATE PERFORMANCE ASSESSMENT (CAPA)

Table of Contents

Executive Summary i

Table of Contents viii

Chapter 1: Introduction 1
- Organization and Contents of the Report 2
- Alignment of Assessments to Standards 2
  - Webb's Terminology 3

Chapter 2: Alignment of the CSTs to the California Content Standards 7
- Alignment Study Design for the CST Review 7
  - Standard Webb Method 7
  - Supplementary Alignment Analyses 7
- Methodology for CSTs 8
  - Workshop Panelists 8
  - Workshop Review Groups 9
  - Materials 10
  - Procedures 11
- Results of Panelist-Test Developer Agreement Analyses 12
- Results of Alignment Analyses for ELA and Mathematics Grade-Level Tests 15
  - Webb Alignment Criteria 15
- Results of Alignment Analyses for Mathematics End-of-Course Tests 22
- Results of Alignment Analyses for Science Tests 25
- Results of Alignment Analyses for History-Social Science Tests 27
- Panelist Ratings of Alignment and Item Quality for CSTs 29
- Summary and Discussion of Alignment Results for the CSTs 31

Chapter 3: Alignment of the CAPA to the California Content Standards 37
- Alignment Study Design for the CAPA Review 37
- Methodology for CAPA 38
  - Workshop Panelists 38
  - Materials 38
  - Procedures 40
- Results for the CAPA 41
  - Webb Alignment Criteria 41

Panelist Ratings of Alignment CAPA Performance Tasks 47
Summary and Discussion of Alignment Results for CAPA 48

Chapter 4: Alignment of the CSTs to the Performance Standards 51

Chapter 5: Summary and Recommendations of Alignment Results 57
Alignment of the Assessments to the California Content Standards 57
California Standards Tests (CSTs) 57
California Alternate Performance Assessments 61
Alignment of the CSTs to the Performance Standards 63
HumRRO Recommendations 63
Recommendations for the CSTs and California Content Standards 63
Recommendations for the CAPA and the Alternate Content Standards 64
Recommendations for the CSTs and Performance Levels 65

References 66

Appendix A: Alignment Results for Webb Criteria by Content Area and Grade Level for the 2006 CST Test Forms A-1
English-Language Arts: Grades 2 through 8 ..... A-2
Categorical Concurrence A-2
Depth-of-Knowledge Consistency A-6
Range-of-Knowledge Correspondence A-10
Balance-of-Knowledge Representation A-14
Mathematics: Grades 2 through 7 ..... A-18
Categorical Concurrence A-18
Depth-of-Knowledge Consistency A-21
Range-of-Knowledge Correspondence A-24
Balance-of-Knowledge Representation A-27
Mathematics: Grade 8 End-of-Course Exams A-30
Categorical Concurrence A-30
Depth-of-Knowledge Consistency A-32
Range-of-Knowledge Correspondence A-34
Balance-of-Knowledge Representation A-36
Mathematics: Integrated Math I, II, and III A-39
Categorical Concurrence A-39
Depth-of-Knowledge Consistency A-41
Range-of-Knowledge Correspondence A-43
Balance-of-Knowledge Representation A-45
Science: Grades 5, 8, and 10 ..... A-47
Categorical Concurrence A-47
Depth-of-Knowledge Consistency A-49
Range-of-Knowledge Correspondence A-51
Balance-of-Knowledge Representation A-53

History-Social Science: Grades 8, 10, and 11 ..... A-55
Categorical Concurrence A-55
Depth-of-Knowledge Consistency A-57
Range-of-Knowledge Correspondence A-59
Balance-of-Knowledge Representation A-61

Appendix B: Alignment Results on Webb Criteria by Content Area and Grade Level for the CAPA B-1
CAPA – English-Language Arts B-2
Categorical Concurrence B-2
Depth-of-Knowledge Consistency B-4
Range-of-Knowledge Correspondence B-7
Balance-of-Knowledge Representation B-9
CAPA – Mathematics B-12
Categorical Concurrence B-12
Depth-of-Knowledge Consistency B-14
Range-of-Knowledge Correspondence B-17
Balance-of-Knowledge Representation B-19

Appendix C: Panelist Comments on Assessment Items C-1
CST Panelist Comments on Individual Items C-2
CAPA Panelist Comments on Individual Items C-5

Appendix D Workshop Materials for Panelists D-1
Example of Alignment Instructions for CST Review D-2
Depth-of-Knowledge (DoK) Descriptions for CSTs D-4
Example Format of Depth-of-Knowledge (DoK) Rating Sheet for CST Standards D-7
Example of Item Rating Sheet for CST Review D-11
Example of Alignment Instructions for CAPA Review D-12
Depth-of-Knowledge (DOK) Descriptions for CAPA D-14
Example Format of Depth-of-Knowledge (DoK) Rating Sheet for CAPA Standards.. D-15
Example of Item Rating Sheet for CAPA Review D-17

List of Tables

Table 1. Webb Labels Applied to California Academic Content Standards 4
Table 2. Webb Labels Applied to Grade 5 Science Content Standards 4
Table 3. Professional and Demographic Characteristics of CST Alignment Panelists 9
Table 4. Workshop Review Groups for CSTs by Content Area and Grade Span 9
Table 5. Characteristics of the CSTs Reviewed 10
Table 6. Depth-of-Knowledge Levels from Alignment Instructions Sheet 11
Table 7. Rate of Agreement between Panelists and ETS on Content Standards Assigned to Items 13
Table 8. Minimum and Maximum Rates of Agreement with ETS across Panelists by Subject and Overall 14

Table 9. Summary of Categorical Concurrence Results for Grade Level ELA and Math Test Forms 16
Table 10. Summary of Depth-of-Knowledge Results for Grade Level ELA and Math Test Forms 18
Table 11. Summary of Range-of-Knowledge Results for Grade Level ELA and Math Test Forms 20
Table 12. Summary of Balance-of-Representation Results for Grade Level ELA and Math Test Forms 21
Table 13. Alignment Results for Algebra I, Algebra II, Geometry, and General Mathematics (Math End-of-Course Tests) 23
Table 14. Alignment Results for Integrated Mathematics I, II, and III Tests (Math End-of-Course Tests) 24
Table 15. Alignment Results for Science Grades 5, 8, and 10 26
Table 16. Alignment Results for History-Social Science 28
Table 17. Summary of Alignment and Quality Ratings for the CSTs 30
Table 18. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for English-language Arts and Mathematics 32
Table 19. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for Math End-of-Course Tests 33
Table 20. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for Science 34
Table 21. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for History-Social Science 35
Table 22. Professional and Demographic Characteristics of Alignment Panelists 38
Table 23. ELA: Number of CAPA Objectives and Targeted Tasks by Level and Content Standard 39
Table 24. Math: Number of CAPA Objectives and Targeted Tasks by Level and Content Standard 39
Table 25. Depth-of-Knowledge Levels for Alternate Assessments from CAPA Alignment Instructions Sheet 40
Table 26. Summary of Categorical Concurrence Results for CAPA ELA and Mathematics by Level 42
Table 27. Summary of Depth-of-Knowledge Results for CAPA ELA and Mathematics by Level 43
Table 28. Summary of Range-of-Knowledge Results for CAPA ELA and Mathematics by Level 44
Table 29. Summary of Balance-of-Knowledge Representation Results for CAPA ELA and Mathematics by Level 46
Table 30. Summary of Alignment Ratings for the CAPA Tasks 47
Table 31. Summary Alignment Outcomes for CAPA English-language Arts and Mathematics on Webb Criteria 49
Table 32. Distribution of Test Questions by Performance Level 52
Table 33. Error of Measurement at the Minimum for Each Performance Level 54
Table 34. Summary Alignment Outcomes for the CSTs per Subject Area and Grade 58
Table 35. Summary Alignment Outcomes for the CAPA per Level 62

Table A-1. Categorical Concurrence for ELA Grade 2: Mean Items per Standard.....A-2
Table A-2. Categorical Concurrence for ELA Grade 3: Mean Items per Standard.....A-3
Table A-3. Categorical Concurrence for ELA Grade 4: Mean Items per Standard.....A-3
Table A-4. Categorical Concurrence for ELA Grade 5: Mean Items per Standard.....A-4
Table A-5. Categorical Concurrence for ELA Grade 6: Mean Items per Standard.....A-4
Table A-6. Categorical Concurrence for ELA Grade 7: Mean Items per Standard.....A-5
Table A-7. Categorical Concurrence for ELA Grade 8: Mean Items per Standard.....A-5
Table A-8. Depth-of-Knowledge Consistency for ELA Grade 2: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-6
Table A-9. Depth-of-Knowledge Consistency for ELA Grade 3: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-7
Table A-10. Depth-of-Knowledge Consistency for ELA Grade 4: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-7
Table A-11. Depth-of-Knowledge Consistency for ELA Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-8
Table A-12. Depth-of-Knowledge Consistency for ELA Grade 6: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-8
Table A-13. Depth-of-Knowledge Consistency for ELA Grade 7: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-9
Table A-14. Depth-of-Knowledge Consistency for ELA Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives.....A-9
Table A-15. Range-of-Knowledge for ELA Grade 2: Mean Percent Objectives per Standard Linked with Items..... A-10
Table A-16. Range-of-Knowledge for ELA Grade 3: Mean Percent Objectives per Standard Linked with Items..... A-10
Table A-17. Range-of-Knowledge for ELA Grade 4: Mean Percent Objectives per Standard Linked with Items..... A-11
Table A-18. Range-of-Knowledge for ELA Grade 5: Mean Percent Objectives per Standard Linked with Items..... A-11
Table A-19. Range-of-Knowledge for ELA Grade 6: Mean Percent Objectives per Standard Linked with Items..... A-12
Table A-20. Range-of-Knowledge for ELA Grade 7: Mean Percent Objectives per Standard Linked with Items..... A-12
Table A-21. Range-of-Knowledge for ELA Grade 8: Mean Percent Objectives per Standard Linked with Items..... A-13
Table A-22. Balance-of-Knowledge Representation for ELA Grade 2: Mean Balance Index per Standard A-14
Table A-23. Balance-of-Knowledge Representation for ELA Grade 3: Mean Balance Index per Standard A-15
Table A-24. Balance-of-Knowledge Representation for ELA Grade 4: Mean Balance Index per Standard A-15
Table A-25. Balance-of-Knowledge Representation for ELA Grade 5: Mean Balance Index per Standard A-16
Table A-26. Balance-of-Knowledge Representation for ELA Grade 6: Mean Balance Index per Standard A-16
Table A-27. Balance-of-Knowledge Representation for ELA Grade 7: Mean Balance Index per Standard A-17

Table A-28. Balance-of-Knowledge Representation for ELA Grade 8: Mean Balance Index per Standard A-17
Table A-29. Categorical Concurrence for Math Grade 2: Mean Number of Items per Standard A-18
Table A-30. Categorical Concurrence for Math Grade 3: Mean Number of Items per Standard A-18
Table A-31. Categorical Concurrence for Math Grade 4: Mean Number of Items per Standard A-19
Table A-32. Categorical Concurrence for Math Grade 5: Mean Number of Items per Standard A-19
Table A-33. Categorical Concurrence for Math Grade 6: Mean Number of Items per Standard A-20
Table A-34. Categorical Concurrence for Math Grade 7: Mean Number of Items per Standard A-20
Table A-35. Depth-of-Knowledge Consistency for Math Grade 2: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-21
Table A-36. Depth-of-Knowledge Consistency for Math Grade 3: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-21
Table A-37. Depth-of-Knowledge Consistency for Math Grade 4: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-22
Table A-38. Depth-of-Knowledge Consistency for Math Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-22
Table A-39. Depth-of-Knowledge Consistency for Math Grade 6: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-23
Table A-40. Depth-of-Knowledge Consistency for Math Grade 7: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-23
Table A-41. Range-of-Knowledge for Math Grade 2: Mean Percent Objectives per Standard Linked with Items A-24
Table A-42. Range-of-Knowledge for Math Grade 3: Mean Percent Objectives per Standard Linked with Items A-24
Table A-43. Range-of-Knowledge for Math Grade 4: Mean Percent Objectives per Standard Linked with Items A-25
Table A-44. Range-of-Knowledge for Math grade 5: Mean Percent Objectives per Standard Linked with Items A-25
Table A-45. Range-of-Knowledge for Math Grade 6: Mean Percent Objectives per Standard Linked with Items A-26
Table A-46. Range-of-Knowledge for Math Grade 7: Mean Percent Objectives per Standard Linked with Items A-26
Table A-47. Balance-of-Knowledge Representation for Math Grade 2: Mean Balance Index per Standard A-27
Table A-48. Balance-of-Knowledge Representation for Math Grade 3: Mean Balance Index per Standard A-27
Table A-49. Balance-of-Knowledge Representation for Math Grade 4: Mean Balance Index per Standard A-28
Table A-50. Balance-of-Knowledge Representation for Math Grade 5: Mean Balance Index per Standard A-28

Table A-51. Balance-of-Knowledge Representation for Math Grade 6: Mean Balance Index per Standard A-29
Table A-52. Balance-of-Knowledge Representation for Math Grade 7: Mean Balance Index per Standard A-29
Table A-53. Categorical Concurrence for Algebra I: Mean Number of Items per Standard A-30
Table A-54. Categorical Concurrence for Algebra II: Mean Number of Items per Standard A-30
Table A-55. Categorical Concurrence for Geometry: Mean Number of Items per Standard A-31
Table A-56. Categorical Concurrence for General Math: Mean Number of Items per Standard A-31
Table A-57. Depth-of-Knowledge Consistency for Algebra I: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives A-32
Table A-58. Depth-of-Knowledge Consistency for Algebra II: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives A-32
Table A-59. Depth-of-Knowledge Consistency for Geometry: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives A-33
Table A-60. Depth-of-Knowledge Consistency for General Math: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives A-33
Table A-61. Range-of-Knowledge for Algebra I: Mean Percent Objectives per Standard Linked with Items A-34
Table A-62. Range-of-Knowledge for Algebra II: Mean Percent Objectives per Standard Linked with Items A-34
Table A-63. Range-of-Knowledge for Geometry: Mean Percent Objectives per Standard Linked with Items A-35
Table A-64. Range-of-Knowledge for General Math: Mean Percent Objectives per Standard Linked with Items A-35
Table A-65. Balance-of-Knowledge Representation for Algebra I: Mean Balance Index per Standard A-36
Table A-66. Balance-of-Knowledge Representation for Algebra II: Mean Balance Index per Standard A-37
Table A-67. Balance-of-Knowledge Representation for Geometry: Mean Balance Index per Standard A-37
Table A-68. Balance-of-Knowledge Representation for General Math: Mean Balance Index per Standard A-38
Table A-69. Categorical Concurrence for Integrated Math I: Mean Number of Items per Standard A-39
Table A-70. Categorical Concurrence for Integrated Math II: Mean Number of Items per Standard A-39
Table A-71. Categorical Concurrence for Integrated Math III: Mean Number of Items per Standard A-40
Table A-72. Depth-of-Knowledge Consistency for Integrated Math I: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives ... A-41
Table A-73. Depth-of-Knowledge Consistency for Integrated Math II: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives ... A-41

Table A-74. Depth-of-Knowledge Consistency for Integrated Math III: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives... A-42
Table A-75. Range-of-Knowledge for Integrated Math I: Mean Percent Objectives per Standard Linked with Items ... A-43
Table A-76. Range-of-Knowledge for Integrated Math II: Mean Percent Objectives per Standard Linked with Items ... A-43
Table A-77. Range-of-Knowledge for Integrated Math III: Mean Percent Objectives per Standard Linked with Items ... A-44
Table A-78. Balance-of-Knowledge Representation for Integrated Math I: Mean Balance Index per Standard ... A-45
Table A-79. Balance-of-Knowledge Representation for Integrated Math II: Mean Balance Index per Standard ... A-46
Table A-80. Balance-of-Knowledge Representation for Integrated Math III: Mean Balance Index per Standard ... A-46
Table A-81. Categorical Concurrence for Science Grade 5: Mean Number of Items per Standard ... A-47
Table A-82. Categorical Concurrence for Science Grade 8: Mean Number of Items per Standard ... A-48
Table A-83. Categorical Concurrence for Science Grade 10: Mean Number of Items per Standard ... A-48
Table A-84. Depth-of-Knowledge Consistency for Science Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives... A-49
Table A-85. Depth-of-Knowledge Consistency for Science Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives... A-49
Table A-86. Depth-of-Knowledge Consistency for Science Grade 10: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives... A-50
Table A-87. Range-of-Knowledge for Science Grade 5: Mean Percent Objectives per Standard Linked with Items ... A-51
Table A-88. Range-of-Knowledge for Science Grade 8: Mean Percent of Objectives per Standard Linked with Core Items ... A-51
Table A-89. Range-of-Knowledge for Science Grade 10: Mean Percent of Objectives per Standard Linked with Core Items... A-52
Table A-90. Balance-of-Knowledge Representation for Science Grade 5: Mean Balance Index per Standard ... A-53
Table A-91. Balance-of-Knowledge Representation for Science Grade 8: Mean Balance Index per Standard ... A-54
Table A-92. Balance-of-Knowledge Representation for Science Grade 10: Mean Balance Index per Standard ... A-54
Table A-93. Categorical Concurrence for History and Social Science Grade 8: Mean Number of Items per Standard ... A-55
Table A-94. Categorical Concurrence for History and Social Science Grade 10: Mean Number of Items per Standard ... A-56
Table A-95. Categorical Concurrence for History and Social Science Grade 11: Mean Number of Items per Standard ... A-56
Table A-96. Depth-of-Knowledge Consistency for History and Social Science Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives... A-57

Table A-97. Depth-of-Knowledge Consistency for History and Social Science Grade 10: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-58
Table A-98. Depth-of-Knowledge Consistency for History and Social Science Grade 11: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives A-58
Table A-99. Range-of-Knowledge for History and Social Science Grade 8: Mean Percent Objectives per Standard Linked with Items A-59
Table A-100. Range-of-Knowledge for History and Social Science Grade 10: Mean Percent of Objectives per Standard Linked with Core Items A-60
Table A-101. Range-of-Knowledge for History and Social Science Grade 11: Mean Percent of Objectives per Standard Linked with Core Items A-60
Table A-102. Balance-of-Knowledge Representation for History and Social Science Grade 8: Mean Balance Index per Standard A-61
Table A-103. Balance-of-Knowledge Representation for History and Social Science Grade 10: Mean Balance Index per Standard A-62
Table A-104. Balance-of-Knowledge Representation for History and Social Science Grade 11: Mean Balance Index per Standard A-63

Table B-1. Categorical Concurrence for ELA CAPA Level I: Mean Number of Performance Tasks per Standard B-2
Table B-2. Categorical Concurrence for ELA CAPA Level II: Mean Number of Performance Tasks per Standard B-2
Table B-3. Categorical Concurrence for ELA CAPA Level III: Mean Number of Performance Tasks per Standard B-3
Table B-4. Categorical Concurrence for ELA CAPA Level IV: Mean Number of Performance Tasks per Standard B-3
Table B-5. Categorical Concurrence for ELA CAPA Level V: Mean Number of Performance Tasks per Standard B-3
Table B-6. Depth-of-Knowledge Consistency for ELA CAPA Level I: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives B-4
Table B-7. Depth-of-Knowledge Consistency for ELA CAPA Level II: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives B-4
Table B-8. Depth-of-Knowledge Consistency for ELA CAPA Level III: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives B-5
Table B-9. Depth-of-Knowledge Consistency for ELA CAPA Level IV: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives B-5
Table B-10. Depth-of-Knowledge Consistency for ELA CAPA Level V: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives B-6
Table B-11. Range-of-Knowledge for ELA CAPA Level I: Mean Percent of Objectives per Standard Linked with Performance Tasks B-7

Table B-12. Range-of-Knowledge for ELA CAPA Level II: Mean Percent of Objectives per Standard Linked with Performance Tasks.....B-7
Table B-13. Range-of-Knowledge for ELA CAPA Level III: Mean Percent of Objectives per Standard Linked with Performance Tasks.....B-8
Table B-14. Range-of-Knowledge for ELA CAPA Level IV: Mean Percent of Objectives per Standard Linked with Performance Tasks.....B-8
Table B-15. Range-of-Knowledge for ELA CAPA Level V: Mean Percent of Objectives per Standard Linked with Performance Tasks.....B-8
Table B-16. Balance-of-Knowledge Representation for ELA CAPA Level I: Mean Balance Index per Standard B-9
Table B-17. Balance-of-Knowledge Representation for ELA CAPA Level II: Mean Balance Index per Standard B-9
Table B-18. Balance-of-Knowledge Representation for ELA CAPA Level III: Mean Balance Index per Standard B-10
Table B-19. Balance-of-Knowledge Representation for ELA CAPA Level IV: Mean Balance Index per Standard B-10
Table B-20. Balance-of-Knowledge Representation for ELA CAPA Level V: Mean Balance Index per Standard B-11
Table B-21. Categorical Concurrence for Math CAPA Level I: Mean Number of Performance Tasks per Standard..... B-12
Table B-22. Categorical Concurrence for Math CAPA Level II: Mean Number of Performance Tasks per Standard..... B-12
Table B-23. Categorical Concurrence for Math CAPA Level III: Mean Number of Performance Tasks per Standard..... B-13
Table B-24. Categorical Concurrence for Math CAPA Level IV: Mean Number of Performance Tasks per Standard..... B-13
Table B-25. Categorical Concurrence for Math CAPA Level V: Mean Number of Performance Tasks per Standard..... B-13
Table B-26. Depth-of-Knowledge Consistency for Math CAPA Level I: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives..... B-14
Table B-27. Depth-of-Knowledge Consistency for Math CAPA Level II: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives..... B-14
Table B-28. Depth-of-Knowledge Consistency for Math CAPA Level III: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives..... B-15
Table B-29. Depth-of-Knowledge Consistency for Math CAPA Level IV: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives..... B-15
Table B-30. Depth-of-Knowledge Consistency for Math CAPA Level V: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives..... B-16
Table B-31. Range-of-Knowledge for Math CAPA Level I: Mean Percent of Objectives per Standard Linked with Performance Tasks..... B-17
Table B-32. Range-of-Knowledge for Math CAPA Level II: Mean Percent of Objectives per Standard Linked with Performance Tasks..... B-17

Table B-33. Range-of-Knowledge for Math CAPA Level III: Mean Percent of Objectives per Standard Linked with Performance Tasks B-18
Table B-34. Range-of-Knowledge for Math CAPA Level IV: Mean Percent of Objectives per Standard Linked with Performance Tasks B-18
Table B-35. Range-of-Knowledge for Math CAPA Level V: Mean Percent of Objectives per Standard Linked with Performance Tasks B-18
Table B-36. Balance of Representation for Math CAPA Level I: Mean Balance Index per Standard B-19
Table B-37. Balance of Representation for Math CAPA Level II: Mean Balance Index per Standard B-19
Table B-38. Balance of Representation for Math CAPA Level III: Mean Balance Index per Standard B-20
Table B-39. Balance of Representation for Math CAPA Level IV: Mean Balance Index per Standard B-20
Table B-40. Balance of Representation for Math CAPA Level V: Mean Balance Index per Standard B-21

Table C-1. ELA Grades 2 to 8: Summary of Panelist Comments on Items C-2
Table C-2. Math Grades 2 to 7: Summary of Panelist Comments on Items C-3
Table C-3. Math End-of-Course Tests: Summary of Panelist Comments on Items C-3
Table C-4. Math Integrated Math Tests I, II, and III: Summary of Panelist Comments on Items C-4
Table C-5. Science Grades 5, 8 and 10: Summary of Panelist Comments on Items C-4
Table C-6. History-Social Science Grades 8, 10, and 11: Summary of Panelist Comments on Items C-4
Table C-7. CAPA ELA Levels I to V: Summary of Panelist Comments on Items C-5
Table C-8. CAPA Math Levels I to V: Summary of Panelist Comments on Items C-5

INDEPENDENT EVALUATION OF THE ALIGNMENT OF THE CALIFORNIA STANDARDS TESTS (CSTS) AND THE CALIFORNIA ALTERNATE PERFORMANCE ASSESSMENT (CAPA)

Chapter 1: Introduction

The California Department of Education (CDE) issued a request for proposal to conduct an external, independent evaluation of the California standards and assessment system. The Human Resources Research Organization (HumRRO) was awarded a contract to conduct this evaluation, and work began on October 27, 2006. The evaluation included two main tasks. The first task, detailed in this report, was an independent review of the alignment of the assessments used for school and district accountability with the California content standards. The second activity required the development of descriptions of each achievement level corresponding with the assessments. The details of the Performance Level Descriptors (PLD) task are reported separately (Wise, Taylor, Becker, Gladden, Handy, Thacker et al., 2007).

This independent evaluation of the California assessment system developed in response to results of California's peer review from the United States Department of Education (USDE). The USDE requested that California provide independent evidence of alignment between the assessments used to calculate Adequate Yearly Progress (AYP) and the State content standards on which those assessments are based. Alignment results should demonstrate that the assessments represent the full range of the content standards, and that these assessments measure student knowledge in the same manner and at the same level of complexity as expected in the content standards.

This requirement by the USDE developed from the No Child Left Behind (NCLB) Act of 2001. NCLB is a Federal education act that challenges each state to establish a coherent assessment system based on solid academic standards. All states receiving Title I funds must present evidence of establishing a fair and consistent assessment system that is based on rigorous standards, sufficient alignment between standards and assessments, and high-quality educational results. Concerning alignment, all aspects of the state assessment system must coincide, including the academic content standards, achievements standards (linked to cut scores), performance level descriptors, and each assessment.

In previous years, California conducted alignment studies to examine the match between the state assessments and the content standards for the California Standards Tests (CSTs) and the California High School Exit Exam (CAHSEE). HumRRO conducted an independent alignment review of the CAHSEE and standards in 2005, and this evidence was submitted and approved by the USDE. An alignment review of the Standardized Testing and Reporting Program compared with the content standards was conducted in 2002, but this review occurred prior to the current assessment configuration and included a norm-referenced assessment. In addition, the first administration of the new science assessments for Grades 8 and 10 occurred in 2006. The California Alternate Performance Assessment (CAPA), originally administered in

2003, is being updated substantially for the 2007 assessment. A new review is needed at this time to provide an independent evaluation of content alignment of the CSTs and of the CAPA.

In addition, the CSTs have not undergone an independent review of the alignment of the assessments to the achievement levels used in reporting[2]. The USDE also requires evidence that the assessments are aligned to the achievement levels, meaning that they provide accurate information for students scoring at different levels. The current evaluation includes an assessment alignment to the performance standards that define the achievement levels as well as alignment to the content standards.

Organization and Contents of the Report

This report includes five chapters. This introductory chapter explains alignment methodologies, including general methods used to evaluate alignment of regular assessments and alternate assessments. Chapter 2 focuses on the alignment approach and results for the CSTs. Chapter 3 describes the alignment approach and results used to evaluate the CAPA. Chapter 4 examines the alignment of the CSTs to the achievement levels. Chapter 5 concludes this report with a summary of the alignment findings and recommendations for changes that might be considered to further improve the alignment of the CSTs and CAPA to California's content and performance standards.

Additional information is provided in the appendices to this report. Appendix A contains tables providing more detail on the content alignment results for the CSTs. Appendix B includes detailed alignment results for the CAPA. The text of chapters 2 and 3 covers a summary of the major results for the CSTs and for the CAPA for brevity. Appendix C provides examples of the rating forms and training materials used in the alignment workshops.

Alignment of Assessments to Standards

As a preface to the discussion of the alignment tasks and results, we first describe several key concepts related to assessment and alignment research. The term *alignment* refers to "the degree to which [content] expectations and assessments are in agreement" (Webb, 2005). Alignment analyses (a) reveal the breadth, or scope, of knowledge included in the assessment, and (b) examine the Depth-of-Knowledge, or cognitive processing, required of students by the assessment compared with the state's content standards. Alignment analyses help to answer questions such as:

- How much content is covered by the assessment?
- Is this content sufficiently similar to the expectations of the standards?
- Are students asked to demonstrate this knowledge at the same level of rigor as expected in the content standards?

[2] Currently, there is no Federal requirement for the development of PLDs for alternate assessments, partly because alternate assessments are supposed to be aligned directly with the primary content standards as well as any alternate content standards.

Several alignment methods are in current use. Most methods involve rating several aspects of test items relative to the content standards. The ratings are analyzed statistically to determine the extent of alignment. For the current alignment reviews, HumRRO adapted a method developed by Norman Webb (1997; 1999; 2005) to evaluate the CSTs and CAPA.

The Webb alignment approach has several advantages over other common alignment methodologies because this method:

- provides distinct statistical criteria and outcomes for judging the quality of alignment;
- is supported by the Council of Chief State School Officers (CCSSO);
- has been used in the State Collaborative on Assessment and Student Standards (SCASS) project, a subsidiary of CCSSO, to guide the development and implementation of many state assessment systems; and
- has been widely researched (Webb, 1997; 1999; 2005).

Versions of the Webb alignment process have been used to evaluate the assessment systems for more than 16 states to meet alignment requirements of NCLB, as well as to evaluate the assessments and frameworks of the National Assessment of Educational Progress (NAEP). One version of the Webb alignment process was used by HumRRO to conduct the CAHSEE alignment review workshops of 2005.

Webb's Terminology

Before explaining more specifics about the Webb method, some explanation of terminology is required. Many states use varying, although similar, terminology to refer to their content standards documents. California uses the terms *domain*, *strand*, and *standard* to label different levels of the content standards documents. The term *domain* refers to the broad content categories within a subject area. For example, Number Sense is one domain included in the grade-level math content standards documents. S*trand* generally refers to the subcategories within the domain. Strands are written as brief statements describing a general skill that students would be expected to know within a particular domain. For instance, under the domain Number Sense, one strand for Grade 2 is 'Students understand the relationship between numbers, quantities, and place value in whole numbers up to 1,000'. Finally, the most specific level of the content expectations for California is a *standard*. This term refers to statements describing specific skills that students are expected to demonstrate. Using math Grade 2 for another example, one standard is 'Count, read, and write whole numbers to 1,000 and identify the place value for each digit'. This standard falls under the strand listed above on understanding number relationships.

Webb uses slightly different terminology compared to California to refer to levels of the content standards documents. Specifically, Webb applies the term *standard* to mean the highest, most general level of the content expectations (in place of *domain* or

reporting category). The results of the analyses are reported at this standard (domain) level, meaning how well the test items align with each of these broad content categories (Webb, 2005). Webb uses the term *content objective* to reference the most specific level of content expectations. In some cases, content standards documents include an additional, intermediate level of organization between the *standard* and the *objective*, which Webb refers to as a *goal*. A *goal* level (comparable to *strand* in the California content standards) usually explains the general expectations for a group of related content objectives. Table 1 presents general comparison between the California and Webb terminology.

Table 1. Webb Labels Applied to California Academic Content Standards

California Content Standards Terminology	Webb Terminology
• Domain • Strand • Standard	• Standard • Goal • Objective

Table 2 includes a specific example from Grade 5 Science to illustrate the comparison between the California Content Standards and Webb. The Grade 5 science content standards document includes a *standard* (domain*)* called Physical Science. One *goal* (strand*)* under this standard is the expectation that students will understand that "Elements and their combinations account for all the varied types of matter in the world." A specific *content objective* (standard*)* under this goal is that "Students know that during chemical reactions the atoms in the reactants rearrange to form products with different properties."

Table 2. Webb Labels Applied to Grade 5 Science Content Standards

Current Labels for Grade 5 California Content Standard in Science

Domain	Strand	Standard
Physical Science	1. Elements and their combinations account for all the varied types of matter in the world.	a. *Students know* that during chemical reactions the atoms in the reactants rearrange to form products with different properties.

Webb Labels Applied to Grade 5 California Content Standard in Science

Standard	Goal	Content Objective
Physical Science	1. Elements and their combinations account for all the varied types of matter in the world.	a. *Students know* that during chemical reactions the atoms in the reactants rearrange to form products with different properties.

The Webb method includes four major criteria to evaluate alignment. These criteria link with statistical procedures used to assess how well individual portions of the assessments and standards documents actually match. The four alignment criteria are as follows: categorical concurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance-of-knowledge representation.

Categorical concurrence is a basic measure of alignment between content standards and test items. This term refers to the proportion of overlap between the content stated in the standards document and that assessed by items on the test.

Depth-of-Knowledge (DOK) measures the type of cognitive processing required by items and content standards. For example, is a student expected to simply identify or recall basic facts, or is the student expected to use reasoning by manipulating information or strategizing? In mathematics, a student may be asked to identify the appropriate use of a decimal among several answer choices. This task should be less complex than trying to explain the concept of a decimal and how and why it can be moved. In English-language arts, asking a student to identify Greek mythology requires less processing compared with asking a student to use knowledge of Greek mythology to understand the origin and meaning of new words.

The purpose of using DOK as a measure of alignment is to determine whether a test item (or performance task) and corresponding standard are both written at the same level of cognitive complexity. Reviewers make two separate judgments about cognitive complexity, one for the standard and one for the item. These two judgments are compared to determine whether the item is written at the same level as the standard to which it is linked. Webb refers to his comparison as *Depth-of-Knowledge consistency*.

Another measure examines the ***range-of-knowledge correspondence*** between the assessment and content standards. The range-of-knowledge measure looks at the breadth of knowledge represented by test items in greater detail. Categorical concurrence simply notes whether a sufficient number of items on the test covers each general content topic (individual strands). However, states generally lay out more specific *content objectives*, or standards, under each strand. The range indicates the number of content objectives assessed by items.

Finally, the ***balance-of-knowledge representation*** criterion focuses on content coverage in yet more detail. In this case, the number of items matched to the content objective does matter. The balance of representation determines whether the assessment measures the content objectives equitably within each standard. Based on Webb's method, items should be distributed evenly across the objectives per standard for good balance. The balance-of-knowledge representation is determined by calculating an index, or score, for each standard. Each standard should meet or surpass a minimum index level to demonstrate adequate balance.

Each criterion provides different information about the degree of alignment between the assessment and content standards. However, all four of Webb's criteria must be considered for a complete picture of alignment.

Chapter 2: Alignment of the CSTs to the California Content Standards

This section of the report outlines the methods and overall alignment results from the review of the operational items from the 2006 CST test forms for English-language arts (ELA), mathematics, science, and history-social sciences (HSS). Grade levels reviewed differed across these subjects (as specified in the contract). After a discussion of the study design and methodology, this chapter summarizes the results of the alignment reviews for each of four criteria, developed by Webb in 2005, as explained below.

Alignment Study Design for the CST Review

The alignment approach used to conduct the CST review involved an adaptation of Webb's alignment method (2005). Specifically, we followed the standard criteria used by Webb to evaluate alignment, and we included several supplementary analyses to enhance this approach. We describe these methods below.

Standard Webb Method

The Webb method requires a set of raters to evaluate each test item on two different dimensions: (a) the standard(s) targeted by items, and (b) the depth-of-knowledge required of students to respond to items. These ratings form the basis of the four separate Webb alignment analyses described previously (see Alignment of Assessments to Standards), which compare the test items with the content standards to examine the breadth and depth of content coverage.

The statistical procedures used to evaluate these criteria allow for separate judgments for each content area about the degree of alignment between the assessment and standards. An overall alignment judgment across all the assessment and content area standards, however, is not appropriate.

Supplementary Alignment Analyses

The standard Webb alignment analyses indicate the general degree of match between the assessment and content standards. HumRRO added several other ratings, as well as analyses, to the alignment evaluation to gain a more comprehensive picture of this match. Concerning ratings, we asked item reviewers to determine just how well they considered the item to assess the selected standard. For example, a reviewer may decide that an ELA item assesses Reading Comprehension, but the reviewer may not consider the item to be the best example of this standard. By using a rating scale, the reviewer can provide more information on how well the item matches the standard (i.e., 'Not at all aligned' to 'Perfectly aligned'). We also asked for a general rating of the item quality. While evaluating the form and content of items was not a primary goal of the alignment task, this information gives further insight into how well the items assess student knowledge.

Additional analyses involved an evaluation of two types of agreement: (1) agreement among reviewers on ratings given to items, and (2) agreement between the reviewers' ratings and the test blueprint constructed by the test developer. The first type of analysis gives some indication of how closely reviewers overlapped in their assigned ratings. The second analysis tells us whether the items actually assessed the content intended by the test developer.

Methodology for CSTs

Workshop Panelists

HumRRO recruited all panelists for participation in the alignment workshops in several ways: (a) general solicitation through letters, email, and phone calls to districts across the State of California; (b) targeted solicitation of specific districts, schools, and staff; and (c) recruitment of educators who participated in previous item development, standards setting, and alignment reviews for California. In all, HumRRO contacted and received nominations for more than 400 candidates, and we accepted 82 panelists for participation. Due to cancellations and absences at the time of the workshops, the final number of panelists who participated in the CST and CAPA reviews totaled 62 across content areas. The final number of panelists reviewing the CSTs was 54.

To ensure high quality panelists, HumRRO and CDE agreed upon several minimum requirements for participation. First, panelists should be current educators (or retired within the past year) working as teachers, district/school curriculum coordinators, special education and English learner specialists, or educational consultants. Second, panelists needed strong familiarity with the grade-level content standards. In addition to these requirements, HumRRO and CDE worked to develop a diverse group of panelists for each content area to include people working in varying demographic areas and with diverse racial/ethnic groups, students with disabilities, and English learners. The resulting panel reflected the true demographic structure of California. Finally, we tried to obtain several panelists from each of the grade levels under review per content area. Table 3 indicates the experience level and demographic characteristics of the panelists.

Table 3. Professional and Demographic Characteristics of CST Alignment Panelists

Professional Position	Number of Panelists	Average Years of Experience	Region of Origin in California			Gender		Ethnicity					
			North	Central	South	M	F	Caucasian	Asian	Hispanic	African-American	Pacific Islander	Other
Teacher	38	14	13	5	20	15	23	26	3	4	2	1	2
Educational Consultant	12	17	3	2	7	2	10	8	1	3	0	0	0
District Coordinators	2	23	1	0	1	0	2	2	0	0	0	0	0
Testing Coordinator	2	23	0	1	1	1	1	2	0	0	0	0	0
Total Panelists	54		17	8	29	18	36	38	4	7	2	1	2

Workshop Review Groups

Panelists split into groups to conduct the reviews. In addition, ELA and mathematics panelists split further by grade span. This procedure was implemented to preserve a manageable number of panelists within each group. Each group included between four and eight panelists with experience teaching one or more grade levels. The breakdown of the panels by content area and grade span is presented in Table 4.

Table 4. Workshop Review Groups for CSTs by Content Area and Grade Span

Content Area	CST Groups and Subgroups	Number of Panelists per Group
ELA	*Grades 2, 3, and 4, including the Grade 4 Writing assessment	6
	**Grades 5, 6, 7, and 8, including the Grade 7 Writing assessment	6
Science	*Grades 5, 8, and 10	7
Mathematics	**Grades 2, 3, and 4	7
	**Grades 5, 6, and 7	5
	*End-of course exams: General Mathematics, Algebra I, Geometry, Algebra II	8
	*Integrated Mathematics 1, 2, and 3	8
History-social science	*Grades 8, 10, and 11	7
	Total panelists	54

* All tests were reviewed by all panelists.
* All tests were reviewed by the *majority* of panelists. One test (varied within and between groups) was not reviewed by one to two panelists due to time constraints.

Materials

Reviewers evaluated the alignment between the assessments and their corresponding standards using Webb's alignment methods and rating forms. Rating forms were in an electronic format

Test Forms. Reviewers assessed the 2006 CST test forms for English-language arts, mathematics, science, and history-social science. Table 5 lists the number of operational items reviewed for each grade-level test per subject. For ELA Grades 4 and 7, these items included one constructed response prompt.

The last column lists the number of content standards assessed using the Webb method. For ELA and the math end-of-course tests in particular, we used the reporting categories corresponding with the standards document for a more fair and appropriate evaluation of the item-to-content match.

Table 5. Characteristics of the CSTs Reviewed

Subject	Grade or Course	Number of Operational Items	Number of Standards/Reporting Categories
ELA	2	65	5
	3	65	5
	4	75	5
	5	75	5
	6	75	5
	7	75	5
	8	75	5
Mathematics	2	65	5
	3	65	5
	4	65	5
	5	65	5
	6	65	5
	7	65	5
	General Math	65	6
	Algebra I	65	4
	Geometry	65	4
	Algebra II	65	4
	Integrated Math I	65	5
	Integrated Math II	65	5
	Integrated Math III	65	5
Science	5	60	4
	8	60	9
	10 (Life Science)	60	6
History-social science	8	75	6
	10	60	6
	11	60	6

Rating Forms and Instructions. Reviewers used two rating forms to make judgments about the standards and the assessment items separately. For the California content standards, reviewers used the depth-of-knowledge (DOK) rating sheet to evaluate the level of knowledge expected by each assessed content objective. This rating form paralleled the format of California content standards with the addition of a column in which to insert the DOK rating next to each content objective (see Appendix D).

For the assessment items, reviewers used the Item Ratings handout to evaluate the items on each of four dimensions. The dimensions included:

(1) match to the specific California content standards;
(2) the Depth-of-Knowledge, or cognitive complexity, expected of students to respond to the assessment items relative to content standards;
(3) item quality using a scale range; and
(4) an overall rating of alignment for each test item per assessment using a 5-point scale.

A sample of the assessment rating form can be found in Appendix D.

To perform the alignment task, reviewers received a copy of the Item Alignment Tasks instruction sheet. This sheet explained how to use each rating form with several examples. The sheet also included definitions for each DOK level, as shown in Table 6.

Table 6. Depth-of-Knowledge Levels from Alignment Instructions Sheet

Level	Title	Description
Level 1	Recall	Item requires simple recall of information, such as facts, definitions, terms, or procedures.
Level 2	Skills/Concepts	Item calls for engagement in some mental processing and decisions beyond habitual response.
Level 3	Strategic Thinking	Item requires students to reason, plan, and use evidence.
Level 4	Extended Thinking	Item requires complex reasoning, planning, and thinking, typically over an extended period of time.

Procedures

HumRRO conducted two 2-day alignment workshops to review the assessments for ELA, math, science, and history-social science. The two workshops were organized by content groups for management purposes. The ELA and science workshops occurred on November 28–29, followed by the math and history-social science workshops on December 1–2. The general procedure and order of sessions were the same for both sets of workshops.

The workshops began with an introduction of staff and observers. Panelists then read and signed an affidavit of non-disclosure regarding any secure materials they would be reviewing over the two-day workshop. HumRRO staff gave a brief

presentation on alignment and the tasks reviewers would perform to all panelists together.

Following the general introduction, panelists split into content groups. For the ELA and mathematics reviews, panelists were split further by grade span. Two HumRRO staff members facilitated each group by discussing the rating procedures in more detail relative to the content area, training reviewers on sample standards and assessment items, and answering questions about the alignment process. Each panelist received a laptop with the rating forms already uploaded and formatted. Panelists received brief instruction on how to open and enter ratings into the electronic forms. Regarding instructions on how to rate standards and items, HumRRO staff provided general suggestions and comments when appropriate; however, they emphasized to reviewers that staff would not give explicit direction on how to rate standards or items because reviewers were valued as content experts.

After reviewing sample DOK evaluations as a group, reviewers proceeded to rate the content objectives from the California blueprint document relevant to their content area and grade span. They first made independent evaluations without discussion. Once all reviewers had completed their ratings, the HumRRO member led the group through a discussion of the objectives to achieve consensus DOK ratings. The consensus ratings were entered into the laptop spreadsheet.

Reviewers then received more specific instructions for rating the assessment items. In particular, reviewers were instructed to assign a *primary standard* to an item based on a judgment that an item clearly measured this content objective. Furthermore, reviewers could assign an *additional standard* only if the item seemed to assess another standard as clearly as the primary standard. Reviewers then evaluated and discussed sample items as a group. After completing the sample items, reviewers proceeded to rate the 2006 test forms relevant to their content area and grade span group. Again, these ratings were entered individually into electronic rating forms on their laptops. Due to time constraints, panelists did not achieve consensus on all items. However, group leaders conducted calibration checks periodically on a small set of items to evaluate the agreement between raters. Panelists reviewed each test form one at a time. For some content groups, test forms were spiraled to ensure that all grade level test forms were reviewed by a sufficient number of panelists[3].

Results of Panelist-Test Developer Agreement Analyses

Before presenting the alignment results on the Webb criteria, we review the agreement levels of our panelists' ratings, particularly compared to the intended content match established by ETS (the test developer). The agreement levels were sufficiently high as to provide further confirmation of the validity of the alignment process and outcomes reported here.

[3] Test spiraling refers to the process of giving each reviewer a different grade-level form at the same time so that no more than two persons are evaluating the same test form simultaneously.

Table 7 shows the percentage of items the panelists matched to the same standards, goals, and objectives (using Webb's terminology) targeted by ETS. Column 4 indicates the total number of ratings made by all of the panelists across the items for each grade/course test. This number was used to calculate percent agreement (as opposed to the number of individual panelists).

Table 7. Rate of Agreement between Panelists and ETS on Content Standards Assigned to Items

Subject	Grade or Course Reviewed	Number of Operational Items per Form	Total Number of Panelist Ratings across Items	Percent Agreement with ETS Codes*			
				Exact	Goal Level	Standard Level	None
ELA	2	65	390	60.8	22.6	10.3	6.4
	3	65	390	64.9	23.9	10.0	1.3
	4	75	456	57.9	24.6	5.0	12.5
	5	75	300	59.7	21.3	13.3	5.7
	6	75	375	59.7	31.2	5.9	3.2
	7	75	304	58.6	25.0	12.8	3.6
	8	75	450	54.7	29.1	8.7	7.6
Math	2	65	260	83.5	10.4	1.5	4.6
	3	65	260	79.6	8.5	2.7	9.2
	4	65	260	67.7	16.5	6.9	8.9
	5	65	260	86.2	7.7	0.0	6.2
	6	65	260	67.3	19.6	4.6	8.5
	7	65	260	67.3	11.2	10.8	10.8
	General Math	65	455	73.2	10.6	12.1	4.2
	Algebra I	65	520	81.2	0.0	16.4	2.5
	Geometry	65	455	67.0	0.2	32.5	0.2
	Algebra II	65	325	76.6	0.0	16.3	7.1
	Int. Math I	65	455	73.4	0.0	22.2	4.4
	Int. Math II	65	455	62.0	0.0	33.4	4.6
	Int. Math III	65	455	75.0	0.0	17.6	7.5
History-social science	8	75	525	79.6	9.3	6.5	4.6
	10	60	413	72.6	14.8	7.8	4.8
	11	60	420	65.0	15.7	6.9	12.4
Science	5	60	420	76.2	15.0	4.5	4.3
	8	60	420	81.9	12.4	0.5	5.2
	10	60	420	63.3	10.0	18.3	8.3
Total**		1,725	9,963	69.7	12.6	11.8	5.9

* Note: Agreement percents were computed across all panelists and operational items in each test.
** Totals agreement percents were computed across all tests as well as all items and panelists.

Under 'Percent Agreement with ETS Codes', we present the results of several analyses evaluating the degree of match between the panelists' ratings and the ETS codes since agreement levels were not exact across the board. The first column indicates the percentage of ratings by panelists that matched the ETS codes exactly across the standard, goal, and objective levels. The next column under 'Percent Agreement with ETS Codes' indicates the percentage of agreement between panelist ratings and ETS at

the goal (or strand) level per standard. If the panelist assigned a different objective, but one that was within the same content subdivision of the target standard, we counted this as a match at the 'Goal Level'. The third column shows the percent agreement between panelists' ratings and ETS at the standard level only. In some cases, panelists chose the same standard (or domain) as ETS, but they did not match the item to the same goal or objective level. Thus, if the content grouping within the standard was not the same as assigned by ETS, we still noted whether the content standard matched by panelists was the same. Finally, the last column shows the percentage of ratings by panelists that did not match the ETS coding at all on items. In this case, some items were matched to an entirely different standard, while other items were judged by panelists as not assessing any of the available California content standards.

One more point should be made regarding the way we made decisions about what counted as a match. If panelists matched an item to *two* different objectives (the most specific content expectations), we counted the one that matched the ETS target objective.

Across nearly 10,000 judgments, the panelists agreed with the ETS content experts 70% of the time overall. The agreement rate was over 94% at the standard level alone and at least 80% for the content grouping within the standard. Agreement rates were somewhat lower for the ELA tests, with less than 60% exact agreement for the tests used with several grades. For each of the tests, the panelists agreed with the ETS ratings more than half of the time. For mathematics, exact agreement rates were generally higher, above 80% for some of the tests.

We also examined the agreement rates separately for each panelist to determine whether any of the individual panelists provided significantly divergent matches. Table 8 displays the minimum and maximum exact agreement rates across the different panelists for each subject. The minimum and maximum percent of time there was no agreement at all is also shown. All of the panelists agreed with the specific objectives targeted by ETS for the majority (more than 50%) of the items reviewed. The maximum rate of no agreement was 15%, meaning that every panelist matched the item to the overall standard targeted by ETS at least 85% of the time. Based on these results, we were comfortable including all of the panelists in the alignment analyses.

Table 8. Minimum and Maximum Rates of Agreement with ETS across Panelists by Subject and Overall

Subject	Percent of items matched by panelists to targeted ETS **objectives**		Percent of items not matched to any ***standard*** by panelists or matched to a different standard than targeted by ETS	
	Minimum (across panelists)	Maximum (across panelists)	Maximum (across panelists)	Minimum (across panelists)
ELA	52	65	9	1
Mathematics	57	89	15	1
History-Social Science	69	78	9	5
Science	52	90	15	3
Overall*	52	90	15	1

* Overall agreement rates are the minima and maxima across all panelists who participated in any of the alignment workshops.

The generally high agreement rates are evidence of the accuracy of the assignments made by the panelists. They also provide an independent validation of the process used by ETS to match items to content standards and objectives.

Results of Alignment Analyses for ELA and Mathematics Grade-Level Tests

In this section of the alignment report, we present the summary of alignment outcomes on the Webb criteria and a summary of panelists' judgments on individual items.

The key alignment results on the Webb criteria are introduced by content area, starting with the elementary and middle grade tests for ELA and mathematics. These assessments are organized around a common set of strands (standards), and it is informative to see how coverage of each of these standards varies across grades. We focus on these content areas and grade levels first because ELA and math must be used for NCLB requirements currently.

Following the results for the ELA and math grade-level tests, we provide the results for the end-of-course mathematics assessments, the science assessments, and the history-social science assessments. Each of these tests is organized around a different set of content standards, or reporting categories. As a result, we used a somewhat different format to present the tabular results for these content areas by including all of the outcomes of the Webb criteria together.

For all tables, we highlight any alignment results that do not meet the minimum criteria. Those table items highlighted in yellow indicate that the assessment is partially aligned to this standard. Red highlighting in the table indicates that no assessment items matched well to this standard.

Webb Alignment Criteria

For each of the four Webb measures, we calculated indicator values for each panelist separately. The detailed result tables provided in Appendix B show the mean and standard deviation of each measure. The standard deviations are a primary measure of agreement among the panelists on the key indicators used in evaluating test alignment. Where the standard deviations are small (relative to the difference between the average index value and the minimum value needed for acceptable alignment), agreement was adequate to high. In a few cases, the standard deviations were higher, suggesting lack of complete agreement on the index values.

Categorical Concurrence. Categorical concurrence is a basic measure of alignment between standards and test items. This measure indicates how much general emphasis each standard receives on an assessment. To determine categorical concurrence, we first counted the number of items that each reviewer judged as assessing each standard. Next, we calculated the mean statistic (M) across all of the reviewers to find the average item rating per standard. For example, if Reviewer A assigned a standard six items, Reviewer B assigned seven items, and Reviewer C

assigned eight items, then the average number of items assigned to that standard is seven[4]. Webb suggests that the mean number of items per standard should be at least six for acceptable content coverage.

Table 9 summarizes the alignment results for the grade-level ELA and math tests on categorical concurrence. The number of tasks matched to each standard varied slightly from one panelist to the next. The numbers shown in Table 9 are averages across all panelists. The more detailed tables in Appendix A include standard deviations indicating how much the panelists varied in the number of tasks matched to each standard. Again, yellow highlighting indicates partial alignment of the assessment to the standard, while red highlighting indicates weak alignment of the assessment to the standard.

Table 9. Summary of Categorical Concurrence Results for Grade Level ELA and Math Test Forms

English-Language Arts – Grades 2–8

Grade	Number of Operational Items per Form	Mean Number of Items Matched to Standard: *Word Analysis	Reading Comprehension	Literary Response and Analysis	**Written and Oral Conventions	Writing Strategies	Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
2	65	19.8	16.7	**4.3**	14.8	8.5	4 of 5	Literary Response
3	65	18.7	17.5	6.7	13.2	8.7	5 of 5	None
4	75	19.2	11.0	7.7	18.2	15.0	5 of 5	None
5	75	12.5	20.0	9.0	17.3	13.3	5 of 5	None
6	75	12.2	17.0	11.6	17.0	15.8	5 of 5	None
7	75	12.0	15.8	14.3	19.5	13.3	5 of 5	None
8	75	8.3	18.5	15.8	13.8	15.8	5 of 5	None

Mathematics – Grades 2–7

Grade	Number of Operational Items per Form	Mean Number of Items Matched to Standard: Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability	Math Reasoning	Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
2	65	37.50	**5.80**	13.30	7.00	**3.00**	3 of 5	Algebra, Math Reasoning
3	65	32.3	11.0	16.3	**3.8**	**2.0**	3 of 5	Statistics, Math Reasoning
4	65	30.00	15.80	11.80	**4.50**	**5.33**	3 of 5	Statistics, Math Reasoning
5	65	29.00	17.30	15.00	**3.30**	6.00	4 of 5	Statistics
6	65	22.30	19.50	9.50	11.00	7.33	5 of 5	None
7	65	20.80	22.30	15.80	**4.80**	8.00	4 of 5	Statistics

Note: Standards were fully covered if at least six items were matched to the standard.
*The full title of reporting category is 'Word Analysis, Fluency, and Vocabulary Development.
** The full title of the reporting category is 'Written and Oral English Language Conventions'.

[4] To obtain the average of 7, one would use the following formula: (6+7+8)/3 = 7.

The results for ELA were generally positive. At least six items assessed most standards, demonstrating general concurrence between the content of assessment and expectations of the ELA standards. The one exception occurred with the standard Literary Response and Analysis. Panelists matched an average of 4.3 items to this standard, which falls below the criterion level. For ELA Grades 4 and 7 in particular, the operational items included one constructed response item each. In addition to matching the prompt to a writing standard, four of six panelists also indicated that the prompt could assess some content objectives in the standard Literary Response and Analysis for both Grades 4 and 7.

For math, two standards were assessed by a reasonable number of items across all grade levels, while the remaining three standards fell short of the minimum criterion. For Algebra and Functions, only the Grade 2 assessment was matched to less than six items on average. However, the mean number of items (M=5.80) matched to this standard clearly is just below the criterion, and the mode across these panelists is seven items. Thus, the items currently written to assess Algebra and Functions likely are sufficient.

In comparison, panelists matched even fewer items to the standard Statistics, Data Analysis, and Probability and the standard Math Reasoning for the majority of grade test forms. These outcomes may reflect an issue in the test design. For Statistics, Data Analysis, and Probability, CDE and the SBE might review whether four or five items (listed in the test blueprints) are sufficient to support a separate reporting category. For Math Reasoning, panelists mapped only a small number of items to the Math Reasoning standard for at least three grade levels. The main cause for the low alignment stems from the blueprint organization for this standard. The test design does not call for any items to be assigned to this standard independently. All items measuring Math Reasoning skills also measure knowledge or skill in one of the other content categories. This fact by itself is not necessarily problematic if California has reason to develop the test blueprint in this way. CDE and ETS might wish to review the test items to ensure that they do in fact clearly measure Math Reasoning. Another option is to review the descriptions of the Math Reasoning objectives to see if further clarification might make it easier to match specific items to these objectives.

Depth-of-Knowledge Consistency. Depth-of-knowledge (DOK) measures the type of cognitive processing required by each item compared to the requirements implied by the content objectives. To make these judgments, reviewers first determined the DOK level for each objective of a standard using a rating scale (see Table 6 for Webb's guidelines). Next, as they reviewed items, panelists rated the level of processing needed to respond to the item using the same DOK rating scales. These two separate judgments about cognitive complexity (one for the standard, one for the item) then were compared to determine the proportion of items written at the appropriate level. Webb refers to this comparison as *depth-of-knowledge consistency*.

Table 10 summarizes the depth-of-knowledge consistency results for each subject and grade level of ELA and math. Since reviewers evaluated depth-of-knowledge at the most specific level of the standards document (content objectives), the table refers to consistency between the items and the content objectives to which they were matched. Results are summarized in terms of the percent of items with cognitive complexity ratings at or above (more complex than) the rating of the selected content objective. Webb's suggests that at least 50% of the items should have complexity ratings at or above the level of the corresponding content objective.

Table 10. Summary of Depth-of-Knowledge Results for Grade Level ELA and Math Test Forms

English-Language Arts – Grades 2–8

Grade	Percent of Items with DOK at or Above Target					Standards Covered at Adequate Depth	Standards Not Covered at Adequate Depth
	*Word Analysis	Reading Comprehension	Literary Response and Analysis	**Written and Oral Conventions	Writing Strategies		
2	76%	50%	68%	100%	84%	5 of 5	None
3	69%	83%	100%	100%	79%	5 of 5	None
4	60%	90%	73%	90%	74%	5 of 5	None
5	60%	69%	66%	**36%**	81%	4 of 5	Written/Oral Conventions
6	**39%**	**48%**	**36%**	**14%**	51%	1 of 5	Word Analysis, Reading Comp, Lit. Response, Written/Oral Conv.
7	82%	82%	69%	**26%**	60%	4 of 5	Written/Oral Conventions
8	**36%**	62%	67%	**47%**	52%	3 of 5	Word Analysis, Written/Oral Conv.

Mathematics – Grades 2–7

Grade	Percent of Items with DOK at or Above Target					Standards Covered at Adequate Depth	Standards Not Covered at Adequate Depth
	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability	Math Reasoning		
2	57%	**18%**	71%	92%	**0%**	3 of 5	Algebra, Math Reasoning
3	78%	57%	79%	100%	**44%**	4 of 5	Math Reasoning
4	80%	77%	71%	55%	79%	5 of 5	None
5	63%	73%	68%	79%	92%	5 of 5	None
6	80%	73%	72%	100%	62%	5 of 5	None
7	81%	**44%**	60%	51%	57%	4 of 5	Algebra

Note: Standards were covered at adequate depth if 50% of the tasks were at or above the complexity level for the matched content objective.

While test items assessed many standards at appropriate cognitive depth, coverage of standards for several grade assessments of ELA and math fell below the Webb criterion. For ELA, fewer than 50% of the Grade 6 assessment items demonstrated agreement with the depth of knowledge levels of four out of five standards. The Grade 8 items assessed student knowledge at a lower cognitive level than the standards for Word Analysis, Fluency, and Vocabulary Development and for Written and Oral English Language Conventions.

In math, over 50% of test items for Grades 2 and 7 covered the Algebra and Functions standard below the content expectations. The items assessing Math Reasoning were particularly problematic at Grade 2. This finding indicates that, of the few items panelists matched to Math Reasoning (M=3 from Table 7), none of these items assessed the standard at the appropriate level of cognitive processing. A number of Grade 3 items assessed Math Reasoning below the content expectations as well. Again, this outcome on Math Reasoning largely can be attributed to the intentional organization of the test blueprint. If each item is written primarily to assess another content standard in addition to Math Reasoning, matching the item on complexity to BOTH standards can be more difficult. Writing items to assess Math Reasoning only is not really feasible; however, CDE may review of the language in the current items in effort to meet the complexity of these standards more accurately.

Range-of-Knowledge. Range-of-knowledge measures how fully the test items cover each of the content objectives within each standard. The assessed objectives within a standard should be linked with at least one item. Webb's minimum level of acceptability for range of correspondence is 50% per standard. This means that at least 50% of the objectives must be matched to one or more items.

Table 11 summarizes the range-of-knowledge results for the grade-level ELA and math test forms. We computed the number of objectives covered for each standard separately for each panelist and then used averages across panelists as the summary indicator.

Results shown in Table 11 indicate that test items represented an adequate range-of-knowledge of all standards across each grade for ELA. As noted earlier under categorical concurrence, Grades 4 and 7 include a constructed response item, which some panelists matched not only to a writing standard but also to the standard Literary Response and Analysis. Specifically, panelists for Grade 4 considered the prompt to assess the content objectives *identifying main events…* and *determine causes of characters' actions…* as well as writing standards. For Grade 7, several panelists matched the prompt to similar content objectives, such as *identify events that advance the plot*… and *analyze characterization as delineated through a character's thoughts…* . This outcome is not problematic and, in fact, seems appropriate given that students must integrate these reading and writing skills in order to produce coherent text.

Table 11. Summary of Range-of-Knowledge Results for Grade Level ELA and Math Test Forms

English-Language Arts – Grades 2–8

Grade	Percent of Objectives per Standard Matched to at Least One Item					Standards with Adequate Range-of-Knowledge	Standards with Limited Range-of-Knowledge
	*Word Analysis	Reading Comprehension	Literary Response and Analysis	**Written and Oral Conventions	Writing Strategies		
2	80%	71%	71%	90%	100%	5 of 5	None
3	102%	71%	72%	91%	100%	5 of 5	None
4	97%	81%	73%	98%	81%	5 of 5	None
5	94%	95%	64%	100%	90%	5 of 5	None
6	90%	78%	78%	100%	84%	5 of 5	None
7	100%	79%	88%	96%	79%	5 of 5	None
8	89%	83%	83%	89%	100%	5 of 5	None

Mathematics – Grades 2–7

Grade	Percent of Objectives per Standard Matched to at Least One Item					Standards with Adequate Range-of-Knowledge	Standards with Limited Range-of-Knowledge
	Number Sense	Algebra and Functions	Measure-ment and Geometry	Statistics, Data Analysis, and Probability	Math Reasoning		
2	91%	75%	89%	94%	**25%**	4 of 5	Math Reasoning
3	95%	86%	88%	92%	**12%**	4 of 5	Math Reasoning
4	86%	75%	55%	70%	**24%**	4 of 5	Math Reasoning
5	93%	80%	96%	**45%**	**18%**	3 of 5	Statistics, Math Reasoning
6	88%	69%	75%	53%	**40%**	4 of 5	Statistics, Math Reasoning
7	83%	58%	71%	83%	**17%**	4 of 5	Math Reasoning

Note: Standards had an adequate Range-of-Knowledge if 50% of the objectives matched one or more of the performance tasks.

Operational items on the math test forms assessed an adequate range of content across many of the math standards as well. Two exceptions for math reflect the same problems discussed under categorical concurrence earlier. The number of target items for Statistics, Data Analysis, and Probability is small relative to the number of objectives in the content standards, particularly for Grades 5. In addition, many of the items were matched to a small number of content objectives for Math Reasoning by the panelists.

Balance-of-Knowledge Representation. The fourth measure of alignment included in the Webb method is *balance-of-knowledge representation.* This criterion tells us whether the number of test items matched to each content standard is

proportional to the number of specific objectives stated for the different standards. The representation of the content standards is balanced if the proportions are similar. The content balance is determined by calculating an index, or score, for each standard[5]. According to Webb, the minimum acceptable index for a single standard is 70 (on a scale of 0 to 100, with 100 representing perfect balance). To be clear, a standard may include more objectives than reviewers actually linked to performance tasks. Thus, only those objectives actually used by the reviewers are included in calculations of the balance index to make it independent of the range-of-knowledge indicator.

Table 12 summarizes balance-of-knowledge representation results for ELA and math per grade. For ELA, panelists determined that the assessment items for Grades 2, 4, 5 and 6 clustered around a small number of objectives for one standard each. This outcome indicates that items assessed some objectives disproportionately.

Table 12. Summary of Balance-of-Representation Results for Grade Level ELA and Math Test Forms

English-Language Arts – Grades 2–8

Grade	Balance Index per Standard					Standards with Adequate Balance	Standards with Limited Balance
	*Word Analysis	Reading Comprehension	Literary Response and Analysis	**Written and Oral Conventions	Writing Strategies		
2	78	**66**	91	90	81	4 of 5	Reading Comprehension
3	76	71	83	81	87	5 of 5	None
4	**59**	86	84	81	75	4 of 5	Word Analysis
5	80	73	84	85	**56**	4 of 5	Writing Strategies
6	77	75	75	81	**67**	4 of 5	Writing Strategies
7	86	76	79	85	80	5 of 5	Reading
8	79	71	75	86	79	5 of 5	None

Mathematics – Grades 2–7

Grade	Balance Index per Standard					Standards with Adequate Balance	Standards with Limited Balance
	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability	Math Reasoning		
2	81	73	81	92	100	5 of 5	None
3	78	72	82	87	94	5 of 5	None
4	82	77	82	89	81	5 of 5	None
5	73	80	79	85	81	5 of 5	None
6	78	**68**	78	81	84	4 of 5	Algebra/Functions
7	75	73	73	81	70	5 of 5	None

Note: Standards had an adequate balance of representation if the index value was 70 or greater.

[5] The exact formula for calculating the balance index is explained in detail in Norman Webb's (2005) alignment training manual: http://www.wcer.wisc.edu/WAT/index.aspx .

For the mathematics assessments, the distribution of items among matched objectives is quite good for most standards at each grade level. The items for the Grade 6 assessment did tend to cluster around a few objectives more than others for Algebra and Functions, but the resulting balance index (M=68) was just below the cut-off criterion (Min=70). These outcomes suggest that, while the test items were not evenly distributed across all content objectives, they were reasonably allocated.

The lower balance indices (highlighted in yellow) in Table 12 likely reflect an uneven allocation of test questions in the test blueprints. If CDE and the SBE are satisfied that the current distribution (assigning more items to some objectives than to others) reflects an important and necessary distinction in the curriculum, no further changes should be required. Otherwise, a further review of the test blueprints is warranted.

Results of Alignment Analyses for Mathematics End-of-Course Tests

In the final results sections, we present the alignment outcomes for each of the remaining subject areas in a single table format because the content standards vary per grade-level test or course. This section includes the results for the mathematics end-of-course tests, including Algebra I, Algebra II, Geometry, and General Mathematics as well as the Integrated Math I, II, and III tests. Table 13 shows the results for four math end-of-course tests on all Webb criteria together. The bottom row under each course displays the total alignment outcomes across standards.

As shown in Table 13, the end-of-course math tests aligned well to the content standards overall on most Webb criteria. Specifically, all of the course tests met the minimum criterion for categorical concurrence in terms of a sufficient number of test items per standard.

Each course test did also display a few alignment discrepancies on one or two standards per course. Depth-of-knowledge consistency was problematic for some items designed to assess one standard in Algebra II (Exponents and Logarithms), one standard in Geometry (Logic and Geometric Proofs), and two standards in General Math (Quantitative Relationships/Evaluating Expressions and Multistep Problems, Graphing, and Functions). In each case, fewer than 50% of items were rated as assessing these standards at the appropriate DOK level. The range-of-knowledge, or content objectives, assessed were limited only for the reporting category Number Properties, Operations, and Linear Equations in Algebra I. Finally, the item distribution across assessed content objectives tended to focus on certain objectives more than others for one standard each in Algebra II, Geometry, and General Math. However, the balance index for each of these three courses was close to the minimum criterion of 70, indicating that the item distribution across content objectives is not highly disproportionate.

Table 13. Alignment Results for Algebra I, Algebra II, Geometry, and General Mathematics (Math End-of-Course Tests)

Course	Standard	Number of Objectives per Standard	Target Number of Items per Standard	Webb Alignment Indicators*			
				Categorical Concurrence	DOK	ROK	Balance
Algebra I	Number Properties, Operations, and Linear Equations	11	17	16.25	68%	**42%**	**67**
	Graphing and Systems of Linear Equations	4	14	14.25	70%	97%	81
	Quadratics and Polynomials	8	21	21.63	71%	95%	82
	Functions and Rational Expressions	6	13	12.00	69%	83%	76
	Total Alignment Outcomes Across Standards	29	65	4 of 4	4 of 4	3 of 4	3 of 4
Algebra II	Polynomials and Rational Expressions	5	19	14.00	73%	84%	83
	Quadratics, Conics, and Complex Numbers	7	16	11.60	52%	63%	76
	Exponents and Logarithms	7	16	24.40	**26%**	74%	83
	Series, Combinatorics, Probability and Statistics	10	14	13.80	84%	50%	85
	Total Alignment Outcomes Across Standards	29	65	4 of 4	3 of 4	4 of 4	4 of 4
Geometry	Logic and Geometric Proofs	7	23	21.71	**44%**	92%	**65**
	Volume and Area Formulas	4	11	11.29	61%	86%	81
	Angle Relationships, Constructions, and Lines	6	16	16.43	79%	83%	74
	Trigonometry	9	15	15.43	65%	54%	82
	Total Alignment Outcomes Across Standards	26	65	4 of 4	3 of 4	4 of 4	3 of 4
General Math	Rational Numbers	6	14	14.86	86%	91%	**69**
	Exponents, Powers, and Roots	5	10	10.43	75%	97%	77
	Quantitative Relationships and Evaluating Expressions	6	11	10.14	**48%**	74%	80
	Multistep Problems, Graphing, and Functions	5	9	9.00	**46%**	83%	78
	Measurement and Geometry	9	12	12.00	54%	78%	77
	Statistics, Data Analysis, and Probability	8	9	8.57	80%	70%	79
	Total Alignment Outcomes Across Standards	39	65	6 of 6	4 of 6	6 of 6	5 of 6

*Note: Bolded entries indicate failure to meet the corresponding Webb criterion. The criteria are:
Categorical Concurrence, at least 6 items per standard
Depth-of-Knowledge (DOK), at least 50% at or above DOK for corresponding objective
Range-of-Knowledge (ROK), at least 50% of objectives matched to one or more item.
Balance-of-Knowledge (Balance), an index value of 70 or more.

Table 14 presents the results for the end-of-course tests for Integrated Mathematics I, II, and III. Many of the content categories for these courses were covered appropriately by the assessments. Several areas demonstrating insufficient alignment are highlighted.

Table 14. Alignment Results for Integrated Mathematics I, II, and III Tests (Math End-of-Course Tests)

Course	Standard	Number of Objectives per Standard	Target Number of Items per Standard	Webb Alignment Indicators*			
				Categorical Concurrence	DOK	ROK	Balance
Integrated Math I	Number Properties, Operations, and Linear Equations	7	15	13.71	70%	55%	80
	Graphing	3	9	9.00	**38%**	100%	75
	Quadratics and Polynomials	5	14	14.71	**49%**	91%	77
	Functions and Rational Expressions	2	7	6.57	**38%**	**19%**	88
	Geometry	10	20	20.00	62%	**10%**	71
	Total Alignment Outcomes Across Standards	27	65	5 of 5	2 of 5	3 of 5	5 of 5
Integrated Math II	Algebra I	13	20	18.57	**43%**	70%	73
	Logic and Geometric Proofs	6	22	20.71	**30%**	83%	70
	Angle Relationships, Constructions, and Lines	5	8	9.71	62%	60%	84
	Trigonometry	4	10	10.14	54%	93%	85
	Algebra II/Probability and Statistics	3	5	**5.00**	74%	**18%**	79
	Total Alignment Outcomes Across Standards	31	65	4 of 5	3 of 5	4 of 5	5 of 5
Integrated Math III	Geometry	1	5	**4.86**	**3%**	100%	100
	Polynomials and Rational Expressions	5	23	18.71	77%	97%	77
	Quadratics, Conics, and Complex Numbers	7	16	15.86	**39%**	**14%**	81
	Exponents and Logarithms	6	16	13.43	**40%**	**6%**	83
	Series, Combinatorics, Probability and Statistics	7	9	9.14	58%	**14%**	90
	Total Alignment Outcomes Across Standards	26	65	4 of 5	2 of 5	2 of 5	5 of 5

*Note: Bolded entries indicate failure to meet the corresponding Webb criterion. The criteria are:
Categorical Concurrence, at least 6 items per standard
Depth-of-Knowledge (DOK), at least 50% at or above DOK for corresponding objective
Range-of-Knowledge (ROK), at least 50% of objectives matched to one or more item.
Balance-of-Knowledge (Balance), an index value of 70 or more.

Tables 14 shows that each test form aligned well with the standards on categorical concurrence and balance-of-knowledge representation. The test forms contained a reasonable number of items for most reporting categories, except Algebra II/Probability and Statistics in Integrated Math II and Geometry in Integrated Math III. In addition, items were distributed well among the reporting categories (balance-of-knowledge representation) in all cases.

The items included in the three test forms did not align as well on depth-of-knowledge consistency or on range-of-knowledge correspondence. Some of the items assessed student knowledge at a lower level of cognitive complexity than expected in the standards. For each course test, two to three (out of five) reporting categories did not meet the minimum requirement on depth-of-knowledge consistency. These outcomes warrant review by CDE and the test developer because they indicate that students are not being assessed at the appropriate level of cognitive depth for the majority of the integrated math standards.

Content coverage by items *within* standards also resulted in insufficient alignment for each integrated math test. For example, as few as 6% of the content objectives for Exponents and Logarithms (Integrated Math III) were assessed by at least one test item (minimum range-of-knowledge criterion is 50%). CDE should review the assessment and the content standards for these courses to determine whether items could be reassigned across content categories (as opposed to increasing overall item numbers); or, whether some content objectives could be deleted, merged, or receive in-class assessment.

Results of Alignment Analyses for Science Tests

This section includes the alignment outcomes for the Grade 5, 8, and 10 science tests. California first administered a Grade 5 science assessment in 2003. However, the Grade 8 and 10 assessments evaluated in the current alignment review are new and were first administered in 2006.

While the general science domains covered in the California curriculum can be divided into physical, earth, and life sciences, the specific science content taught does differ per grade level, as shown by the list of content standards in Table 15. The alignment between the grade-level test forms and the content standards was good overall, although panelists found lower levels of alignment with several standards for each grade-level test.

Table 15. Alignment Results for Science Grades 5, 8, and 10

Grade	Standard	Number of Objectives per Standard	Target Number of Items per Standard	Webb Alignment Indicators*			
				Categorical Concurrence	DOK	ROK	Balance
5	Physical Science	16	18	17.1	75%	88%	86
	Life Science	14	18	17.0	51%	82%	78
	Earth Science	17	18	17.9	72%	79%	82
	Investigation and Experimentation	15	6	8.0	65%	**43%**	90
	Total Alignment Outcomes Across Standards	62	60	4 of 4	4 of 4	3 of 4	4 of 4
8	Motion	6	8	9.0	72%	93%	77
	Forces	7	8	8.1	86%	90%	85
	Structure of Matter	6	8	7.9	74%	86%	88
	Earth Science	5	7	7.0	76%	100%	82
	Reactions	5	8	7.6	71%	100%	81
	Chemistry of Living Systems	3	3	**2.9**	55%	86%	95
	Periodic Table	3	7	7.3	86%	100%	81
	Density and Buoyancy	4	5	6.4	92%	93%	77
	Investigation and Experimentation	7	6	**3.6**	60%	**39%**	91
	Total Alignment Outcomes Across Standards	46	60	7 of 9	9 of 9	8 of 9	9 of 9
10	Cell Biology	8	10	9.0	58%	88%	86
	Genetics	10	12	12.7	58%	83%	78
	Ecology	9	11	11.3	**40%**	83%	83
	Evolution	10	11	10.6	69%	74%	82
	Physiology	8	10	11.4	**40%**	89%	77
	Investigation and Experimentation	9	6	**4.7**	67%	**41%**	87
	Total Alignment Outcomes Across Standards	54	60	5 of 6	4 of 6	5 of 6	6 of 6

*Note: Bolded entries indicate failure to meet the corresponding Webb criterion. The criteria are:
Categorical Concurrence, at least 6 items per standard
Depth-of-Knowledge (DOK), at least 50% at or above DOK for corresponding objective
Range-of-Knowledge (ROK), at least 50% of objectives matched to one or more item.
Balance-of-Knowledge (Balance), an index value of 70 or more.

Items for each grade test form did not align well to the standard Investigation and Experimentation on two Webb criteria. For Grade 5, Investigation and Experimentation was the only standard that received narrow coverage of the content objectives (range-of-knowledge) on the assessments. For Grade 8, panelists matched the standards Investigation and Experimentation, as well as Chemistry of Living Systems, to a small number of items overall. Within the standard, these few items targeted a handful of objectives for Investigation and Experimentation (approximately three of seven). The same pattern resulted for Grade 10.

The small number of items matched to Chemistry and Living Systems for Grade 8 seems tied to the number of items targeted for assessment in the test blueprint (N = 3). Clearly, Table 15 indicates that panelists' ratings agreed with this number (Mean Items Matched = 2.9). Thus, the outcome in this case does not represent poor alignment between the assessment and standards so much as an insufficient number of items targeted for assessment. CDE and the test developer may wish to review this issue.

The Grade 10 test items appeared to assess students at lower depth-of-knowledge than expected in the standards on Ecology and on Physiology. These science items could undergo review to increase cognitive complexity to better match the standards.

Results of Alignment Analyses for History-Social Science Tests

Table 16 provides the results for Grades 8, 10, and 11 history-social science test forms per content standard. One point should be noted about history-social science concerning the History and Social Science Analysis Skills standard included for each grade. Items written to assess this standard are embedded, meaning that these items assess another content standard as the primary target in addition to analysis skills. Our panelists did not assign this standard as the primary standard for any item on any of the history-social science grade-level tests. However, they did assign the standard as an additional target along with the primary standard.

Table 16. Alignment Results for History-Social Science

Grade	Standard	Number of Objectives per Standard	Target Number of Items per Standard	Webb Alignment Indicators*			
				Categorical Concurrence	DOK	ROK	Balance
8	Ancient Civilizations	48	16	16.71	**39%**	**32%**	92
	Late Antiquity and Middle Ages	40	14	13.57	**42%**	**29%**	88
	Renaissance and Reformation	21	10	9.57	56%	**41%**	92
	U.S. Constitution and Early Republic	42	22	21.86	53%	**42%**	85
	Civil War and its Aftermath	27	13	13.14	51%	**38%**	85
	History and Social Skills Analysis	Embedded		10.00	**46%**	**33%**	70
	Total Objectives/Items Acceptable Standards	178	75	6 of 6	3 of 6	0 of 6	6 of 6
10	Modern Political Thought	8	13	12.86	79%	75%	76
	Industrial Expansion	11	10	10.00	51%	68%	84
	First World War	9	14	13.86	**35%**	79%	78
	Second World War	8	13	13.00	**36%**	86%	78
	Post World War II	11	10	9.14	**22%**	52%	80
	History and Social Skills Analysis	Embedded		3.00	**18%**	**26%**	89
	Total Objectives/Items Acceptable Standards	48	60	6 of 6	2 of 6	5 of 6	6 of 6
11	American Political and Social Thought	9	10	6.00	56%	54%	85
	U.S. Role as a World Power	15	13	15.86	63%	64%	76
	U.S. Between the World Wars	12	12	13.00	**32%**	66%	73
	World War II and Foreign Affairs	15	12	12.86	**33%**	55%	74
	Post World War II Domestic Issues	22	13	12.29	**49%**	**44%**	84
	History and Social Skills Analysis	Embedded		3.14	**9%**	**29%**	85
	Total Objectives/Items Acceptable Standards	73	60	6 of 6	2 of 6	4 of 6	6 of 6

*Note: Bolded entries indicate failure to meet the corresponding Webb criterion. The criteria are:
Categorical Concurrence, at least 6 items per standard
Depth-of-Knowledge (DOK), at least 50% at or above DOK for corresponding objective
Range-of-Knowledge (ROK), at least 50% of objectives matched to one or more item.
Balance-of-Knowledge (Balance), an index value of 70 or more.

As demonstrated in Table 16, the three assessments definitely targeted and assessed an appropriate number of items per standard (strong categorical concurrence), and these items were distributed rather equally across content objectives (balance of representation). However, two areas of weakness in alignment are evident. First, across grade tests, panelists rated items as assessing student knowledge at a lower cognitive level than expected by the history-social science content standards. Thus, the consistency in depth-of-knowledge between the items and many of the standards is inadequate based on the ratings of our panelists.

The second area requiring attention is the range of content covered by the assessments. The Grade 8 test in particular demonstrated a narrow span of content assessment to the extent that none of the standards met the minimum criterion (50% of objectives linked to at least one item). This outcome is not entirely surprising because most of the standards include at least twice as many content objectives as test items, making it difficult to assess the standards adequately with the current test design. For this reason, CDE and the SBE should review the structure of the content specifications for Grade 8 history-social science especially.

One final point pertains to the History and Social Science Analysis Skills standard. Items assessing this standard were designed to target another history-social science standard as the primary standard. However, as noted in the test blueprint, 25% of the items should assess History and Social Science Analysis Skills in addition to another standard. From Table 14 (and also noted in more specific data tables in Appendix A), less than 15% of items were matched to this standard by our panelists for each grade. A review of the items intended to assess this standard may be warranted.

Panelist Ratings of Alignment and Item Quality for CSTs

This final section focuses on more qualitative outcomes from the alignment review. First, we present results on the ratings panelists provided on overall item alignment and item quality. Table 17 includes the percentage of items rated as a 3 or 4. For item alignment, a 3 indicates that an item was judged ‘Highly Aligned to Content Standard’, while a 4 indicates ‘Fully Aligned to Content Standard’. For item quality, a rating of 3 means that panelists found the item to be of ‘Good Quality’ or typical of an item assessing this standard, while a 4 indicates that panelists determined that an item was ‘Excellent Quality’ or almost textbook in example of assessment item for the standard.

Table 17. Summary of Alignment and Quality Ratings for the CSTs

Test	Percentage of Items with Good Alignment Ratings (Rating of 3 or 4)	Percentage of Items with Good Quality Ratings (Rating of 3 or 4)
	English-Language Arts	
Grade 2	92	87
Grade 3	89	88
Grade 4	87	93
Grade 5	88	94
Grade 6	83	96
Grade 7	87	95
Grade 8	90	97
	Mathematics	
Grade 2	91	95
Grade 3	80	92
Grade 4	85	95
Grade 5	94	93
Grade 6	96	93
Grade 7	95	99
General Math	95	91
Algebra I	93	93
Geometry	96	92
Algebra II	89	82
Int. Math I	93	99
Int. Math II	95	99
Int. Math III	93	99
	History-Social Science	
Grade 8	92	94
Grade 10	87	95
Grade 11	88	96
	Science	
Grade 5	92	90
Grade 8	95	96
Grade 10	95	97

Both the item alignment and item quality ratings were quite high for the CSTs. The lowest item quality rating was 82% for Algebra II. CDE and the test developer may wish to examine the quality ratings for the individual items in the Algebra II test.

In addition to providing more standardized ratings, panelists were given the opportunity to make notations about items during the item rating period. Specific comments referenced secure test items, which precludes us from including the details in this report. These comments have been shared with CDE and the test developers for appropriate action.

Overall, the comments were typical and minor in nature. Comments often pointed to problems in item wording or clarity. Some panelists did point to items they considered inappropriate for the grade level (often lower than grade-level expectations). However, this type of comment pertained only to 1% to 3% of items at most per grade-level test and subject. Many subjects and grades did not receive this comment at all.

Summary and Discussion of Alignment Results for the CSTs

The alignment review of the CSTs for English-language arts, mathematics, science, and history-social science involved an evaluation of the operational items from the 2007 test forms compared to the California Content Standards. HumRRO applied the Webb alignment method to conduct the review. The overall alignment results for the CSTs were good. However, some subject areas may require reconsideration to improve the quality of alignment. We present a summary of the alignment outcomes for the CSTs in this section.

Summary alignment judgments are based on Webb (1999). These summary judgments focus on the percentage of content standards represented well by the assessment. Webb outlined a scale with a range of potential alignment outcomes applied to each criterion:

- Fully aligned – assessments align to all content standards (100%);
- Highly aligned – assessments align to the majority of standards (70–90%)
- Partially aligned – assessments align well to some standards (50–69%); and
- Weakly aligned – assessments align to less than half the standards (below 50%).

Webb's alignment method does not allow for a *single* judgment of overall alignment across the four criteria. However, one can get a sense of overall alignment between the assessments and standards by looking at all of the criteria together.

The summary tables in this section are linked to the column labeled 'Standards Covered at Adequate Depth' in each of Tables 7 through 14 of the Results section. Thus, these summary judgments reflect a final evaluation of each grade-level or course assessment per Webb criteria *across* the standards for a given grade.

Table 18 presents the summary alignment outcomes for the ELA and math elementary and middle grade tests based on the above scale. The table includes the alignment judgment, along with the percentage of standards covered well by the assessment for each Webb criterion. As shown in the table, a number of the alignment outcomes for ELA and math are good. However, some aspects of the ELA and Math tests demonstrated low levels of alignment to the content standards on one or more of the Webb criteria. Those subject and grade-level tests with partial to weak alignment are highlighted in each table. As in the Results section, yellow highlighting indicates partial alignment to the standards, while red highlighting indicates weak alignment to the standards.

Table 18. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for English-language Arts and Mathematics

	Percentage of Standards that Met Webb Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
Grade	English-Language Arts			
2	Highly aligned (80%)	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (80%)
3	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)
4	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (80%)
5	Fully aligned (100%)	Highly aligned (80%)	Fully aligned (100%)	Highly aligned (80%)
6	Fully aligned (100%)	Weakly aligned (20%)	Fully aligned (100%)	Highly aligned (80%)
7	Fully aligned (100%)	Highly aligned (80%)	Fully aligned (100%)	Fully aligned (100%)
8	Fully aligned (100%)	Partially aligned (60%)	Fully aligned (100%)	Fully aligned (100%)
Grade	Mathematics			
2	Partially aligned (60%)	Partially aligned (60%)	Highly aligned (80%)	Fully aligned (100%)
3	Partially aligned (60%)	Highly aligned (80%)	Highly aligned (80%)	Fully aligned (100%)
4	Partially aligned (60%)	Fully aligned (100%)	Highly aligned (80%)	Fully aligned (100%)
5	Highly aligned (80%)	Fully aligned (100%)	Partially aligned (60%)	Fully aligned (100%)
6	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (80%)	Highly aligned (80%)
7	Highly aligned (80%)	Highly aligned (80%)	Highly aligned (80%)	Fully aligned (100%)

The majority of the 2007 ELA test forms aligned well with the content standards across the Webb criteria. The two exceptions were Grades 6 and 8 on the depth-of-knowledge criterion. For Grade 6, panelists determined that less than half of the items developed for four of five standards (all but Writing Strategies) assessed student knowledge at the same cognitive level expected in the content expectations. The items for ELA Grade 8 also produced a mismatch in cognitive demand with the Word Analysis, Fluency, and Vocabulary Development standard as well as the Written and Oral English Language Conventions standard. These outcomes suggest that a review of the item pool for these two grade levels could be useful to evaluate the appropriateness of the complexity level.

The lower levels of alignment for math at Grades 2 through 7 on some criteria mostly occurred with the same two standards: Statistics, Data Analysis, and Probability and Math Reasoning. Panelists' ratings suggest that the mathematics tests for Grades 3, 4, 5, and 7 do not include a sufficient number of items to assess the Statistics standard adequately, producing low categorical concurrence. Reduced alignment between test items and the Math Reasoning standard was more comprehensive. However, as noted in earlier sections of this report, many of these outcomes are explainable as the result of the test blueprint because Math Reasoning is *intended* to be an additional standard targeted for assessment by some items. One suggestion is to re-evaluate the items selected for the assessment of Math Reasoning since our panelists found it difficult to clearly match many items to this standard and they considered these items to be less complex overall in comparison to the standard.

The assessment of the Algebra and Functions Standard also resulted in lower alignment with a couple of standards for Grades 2 and 7 on the categorical concurrence and depth-of-knowledge consistency criteria. While some review of the items assessing this standard may be appropriate, the alignment issues between the test forms and standards for Grades 2 and 7 is not as extensive.

Table 19 includes the summary alignment outcomes for the end-of-course mathematics tests given in Grades 8 through 11.

Table 19. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for Math End-of-Course Tests

	Percentage of Standards that Met Webb Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
Course	Math End-of-Course Tests			
Algebra I	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (75%)	Highly aligned (75%)
Algebra II	Fully aligned (100%)	Highly aligned (75%)	Fully aligned (100%)	Fully aligned (100%)
Geometry	Fully aligned (100%)	Highly aligned (75%)	Fully aligned (100%)	Highly aligned (75%)
General Math	Fully aligned (100%)	Partially aligned (67%)	Fully aligned (100%)	Highly aligned (83%)
Course	Integrated Mathematics			
Integrated Math I	Fully aligned (100%)	Weakly aligned (40%)	Partially aligned (60%)	Fully aligned (100%)
Integrated Math II	Highly aligned (80%)	Partially aligned (60%)	Highly aligned (80%)	Fully aligned (100%)
Integrated Math III	Highly aligned (80%)	Weakly aligned (40%)	Weakly aligned (40%)	Fully aligned (100%)

From Table 19, three of the tests aligned very well to the standards. For General Math, panelists did rate the items as low in depth-of-knowledge compared to the content objectives for two standards. The integrated math test forms did not align well to many of the content standards on depth-of-knowledge consistency or on range-of-knowledge representation. For Integrated Math II and III, the test forms exhibited weak content coverage of the majority of standards, as determined by our panelists. It appears that a small number of content objectives was covered within these standards, and that these items assess student knowledge at a low cognitive level.

Table 20 includes the summary alignment outcomes for the science tests given in Grades 5, 8, and 10. Despite deficits in alignment for some individual standards shown in Table 15, the overall alignment picture for science was fairly positive per grade-level test. As shown by the highlighting in the table, the cognitive complexity of the Grade 10 items may require review, however.

Table 20. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for Science

	Percentage of Standards that Met Webb Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
Grade	Science			
5	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (83%)	Fully aligned (100%)
8	Highly aligned (78%)	Fully aligned (100%)	Highly aligned (89%)	Fully aligned (100%)
10	Highly aligned (83%)	Partially aligned (67%)	Highly aligned (83%)	Fully aligned (100%)

Table 21 includes the summary alignment outcomes for the history-social science tests given in Grades 8, 10, and 11. The most prominent alignment issue warranting attention by CDE and the test developer is the inconsistency found between the test items and content standards on depth-of-knowledge across all three grade-level tests. In addition, the tests for Grades 8 and 11 assessed a narrow range of content within the standards.

Table 21. Summary Alignment Outcomes on Each Webb Criterion per Grade Level for History-Social Science

	Percentage of Standards that Met Webb Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
Grade	History-Social Science			
8	Fully aligned (100%)	Partially aligned (67%)	Weakly aligned (0%)	Fully aligned (100%)
10	Fully aligned (100%)	Weakly aligned (40%)	Highly aligned (83%)	Fully aligned (100%)
11	Fully aligned (100%)	Weakly aligned (40%)	Partially aligned (67%)	Fully aligned (100%)

We provide a more detailed discussion of the implications of the alignment outcomes in chapter 5. This chapter also includes recommendations to CDE for increasing alignment for those subject areas and grades in need of improvement.

Chapter 3: Alignment of the CAPA to the California Content Standards

The CAPA is an alternate assessment designed for those students with significant cognitive disabilities. This assessment was developed in 2003 to meet the requirements of the Individuals with Disabilities Education Act (IDEA) and the No Child Left Behind Act (NCLB). The assessment is composed of a set of performance tasks, as opposed to the selected-response and constructed-response items included in the California Standards Tests (CSTs). These performance tasks allow students to demonstrate their knowledge through hands-on activities. The CAPA performance tasks are organized into five levels each for ELA and math.

A new version of the CAPA will be implemented with the 2008 assessment. We evaluated alignment of the new CAPA by reviewing field tryout tasks administered with the operational 2007 assessment. Each version of the 2007 forms included four field tryout tasks along with eight operational tasks. The number of test form versions for 2007 with distinct sets of field tryout tasks varied from four to seven, yielding 16 to 28 new tasks. Since each new operational form will include eight scored tasks, the number of tasks reviewed was two to three times the number of tasks in an operational form. Results of the evaluation of the alignment of each of the five levels of the CAPA ELA and mathematics assessments to the corresponding content standards are reported in this section.

Alignment Study Design for the CAPA Review

Since the CAPA is based on performance tasks and alternate standards, the alignment methods used to evaluate the regular on-grade assessments are not entirely appropriate. Instead, HumRRO used a modified version of the Webb alignment method developed for alternate assessments (Almond, Filbin, Hall, & Tindal, 2005; Browder, Flowers, Ahlgrim-Delzell, Karvonen, Spooner, & Algozzine, 2004; Tindal, 2005). This approach has been used successfully to evaluate the alignment of alternate assessments in seven other states (Roach, Elliott, & Webb, 2005).

This alignment approach differs from Webb's traditional method in three major ways. First, as noted earlier, panelists evaluate performance tasks, or portfolios, instead of discrete items. Second, the cognitive complexity present in alternate assessments necessarily differs from regular assessments in type and degree. As a result, the depth-of-knowledge descriptions used at each level to evaluate the performance tasks differ from those used to rate test items. Third, while each of the four Webb alignment measures are retained, Webb adjusted the minimum conditions for demonstrating alignment between the alternate assessment and the content standards. For example, an alternate assessment must include at least one performance task per standard to demonstrate adequate categorical concurrence, compared to the requirement of a minimum of six items per content standard expected for a regular assessment. This change is appropriate because the CAPA tasks take longer to complete and contain more information in responses than do individual test items.

Methodology for CAPA

Workshop Panelists

HumRRO recruited panelists to participate in the CAPA workshops in the same manner and during the same timeframe as occurred for the CSTs. The same expectations for quality panelists and for diversity in background and demographics were applied in the recruiting process. We were successful in recruiting teachers and other content experts with extensive experience with special education students. However, achieving a diverse panel proved more difficult for the CAPA compared to the CST panels due to the smaller pool of teachers/educators who work with this group and to somewhat lower ethnic diversity among current Special Education (SpED) teachers.

The resulting panels included four reviewers each for ELA and for math (total N=8). The experience and characteristics of these panelists are presented in Table 22.

Table 22. Professional and Demographic Characteristics of Alignment Panelists

Professional Position	Number of Panelists per Type of Position	Mean Years of Experience	Region of Origin in California			Gender		Ethnicity			
			North	Central	South	M	F	Caucasian	Asian	Hispanic	Other
SpED Teacher	3	17.67	0	2	1	1	2	2	1	0	0
Educational Consultant	2	7.5	1	0	1	0	2	2	0	0	0
Curriculum Specialist	2	25	1	0	1	0	2	2	0	0	0
SpED Specialist	1	30	1	0	0	0	1	1	0	0	0
Total Panelists	8										

Materials

Reviewers evaluated the alignment between the assessments and their corresponding standards using Webb's alignment methods and rating forms. Rating forms were in an electronic format

Rating Forms and Instructions. Reviewers used two separate rating forms to make judgments about the CAPA content standards and the performance tasks. For the CAPA standards, reviewers used the rating sheet for Depth-of-Knowledge Levels for CAPA to evaluate the level of knowledge expected by each assessed content objective. This rating form paralleled the format of CAPA test blueprints with the addition of a

column in which to insert the DOK rating next to each content objective (see Appendix C). The test blueprints indicate the number of performance tasks needed to assess each content standard (strand). The format of these blueprints includes the major standards from the regular California Content Standards, along with one or more specific CAPA objectives linked to each content standard. Tables 23 and 24 show the number of objectives and targeted tasks for each level of the CAPA.

Table 23. ELA: Number of CAPA Objectives and Targeted Tasks by Level and Content Standard

English-Language Arts				
Level	Grade Range	Reading	Writing	Listening and Speaking
		Number of Objectives		
I	Ungraded	6	4	8
II	2 – 3	8	6	4
III	4 – 5	10	6	0
IV	6 – 8	11	4	5
V	9 – 12	7	3	3
		Number of Tasks Targeted		
I	Ungraded	4	1	3
II	2 – 3	4	1	3
III	4 – 5	5	3	0
IV	6 – 8	4	2	2
V	9 – 12	4	2	2

Table 24. Math: Number of CAPA Objectives and Targeted Tasks by Level and Content Standard

Mathematics					
Level	Grade Range	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis and Probability,
			Number of CAPA Objectives		
I	Ungraded	7	5	7	1
II	2 – 3	14	4	6	2
III	4 – 5	11	2	6	6
IV	6 – 8	6	2	1	0
V	9 – 12	8	0	1	0
			Number of Tasks Targeted		
I	Ungraded	3	1	3	1
II	2 – 3	4	1	2	1
III	4 – 5	3	1	2	2
IV	6 – 8	5	2	1	0
V	9 – 12	7	0	1	0

For the performance tasks, reviewers used the CAPA Performance Task Ratings Sheet to record ratings of each task on four dimensions. These included:

- how well the task matched to the specific California content standards;
- the Depth-of-Knowledge, or cognitive complexity, expected of students to respond to the assessment items relative to content standards;
- item quality using a scale range; and
- an overall rating of alignment for each test item per assessment using a 5-point scale.

A sample of the assessment rating form can be found in Appendix D.

To perform the alignment task, reviewers received a copy of the CAPA Item Alignment Tasks instruction sheet. This sheet explained how to use each rating form, giving several examples. The sheet also included definitions for each DOK level for alternate assessments, as shown in Table 25.

Table 25. Depth-of-Knowledge Levels for Alternate Assessments from CAPA Alignment Instructions Sheet

Level	Title	Description
Level 1	Recall	Requires students to recall or observe facts, definitions, and terms. Involves simple one-step procedures. Involves computing simple algorithms (e.g., sum, quotient).
Level 2	Skills/Concepts	This level includes the engagement of some mental processing beyond a habitual response. The item requires students to make some decisions as to how to approach a problem or activity.
Level 3	Strategic Thinking	A multiple-step 'behavioral event' is executed in more than one context. Requires reasoning, planning, or use of evidence to solve problem or algorithm. May involve activity with more than one possible answer. Requires conjecture or restructuring of problems. Involves drawing conclusions from observations, citing evidence and developing logical arguments for concepts. Uses concepts to solve nonroutine problems.
Level 4	Extended Thinking	The 'behavioral event' reflects an approach (of many) to completing the task. May require complex reasoning, planning, developing and thinking. Typically requires extended time to complete problem, but time spent not on repetitive tasks. Requires students to make several connections and apply one approach among many to solve the problem. Involves complex restructuring of data, establishing and evaluating criteria to solve problems.

Procedures

HumRRO conducted separate 2-day alignment workshops for ELA and mathematics. The ELA workshop occurred on November 28–29, and the math

workshop occurred on December 1–2. The general procedures and order of sessions were the same as for the CSTs and are not repeated here.

Results for the CAPA

In this section of the alignment report, we present the results on the CAPA. As with the CSTs, we analyzed the panelist ratings of the performance tasks with the alternate standards using the four Webb criteria. Finally, we evaluated reviewers' overall ratings of the items. One point to note relevant to each of the subsequent tables is that some tables include the statement 'Not Assessed'. This statement indicates that these standards are not targeted in the standardized CAPA assessment, although this content is still taught and may be subject to in-class assessment.

Webb Alignment Criteria

We present a summary of the results for each of the four Webb criteria in this section. Tables providing more detailed results are included in Appendix B.

Categorical Concurrence. Categorical concurrence describes the extent to which the CAPA tasks cover all of the targeted content standards. For a regular assessment, Webb recommends a minimum of six test questions assessing standard (content strand), but for an alternate assessment, the criterion is one performance task per standard. We reviewed more tasks than would normally be included in an operational form, so we set the minimum coverage to one task for each eight tasks reviewed (the equivalent of an operational form).

Table 26 summarizes the CAPA alignment results for categorical concurrence. The number of tasks matched to each standard varied slightly from one panelist to the next. The numbers shown in Table 26 are averages across all four panelists. The more detailed tables in Appendix B include standard deviations indicating how much the panelists varied in the number of tasks matched to each standard.

As Table 26 indicates, each of the standards was covered adequately for all levels of the CAPA. One caveat should be pointed out for the Statistics, Data Analysis, and Probability standard for Level II of the CAPA mathematics assessment. For this particular standard, that the test developer assigned only one task for this category. While at least two panelists assigned more than one task to the standard, the test developers may have difficulty constructing more than one new form that meets the blueprint requirements if only one item exists. However, it should be possible to pilot additional tasks for this standard before more forms are needed.

Table 26. Summary of Categorical Concurrence Results for CAPA ELA and Mathematics by Level

English-Language Arts					
	Balance Index per Standard				
Level	Reading	Writing	Listening and Speaking	Standards with Adequate Balance	Standards with Limited Balance
I	71	94	100	3 of 3	None
II	79	100	85	3 of 3	None
III	84	79	Not assessed	2 of 2	None
IV	80	90	84	3 of 3	None
V	83	87	87	3 of 3	None

Mathematics						
	Balance Index per Standard					
Level	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability	Standards with Adequate Balance	Standards with Limited Balance
I	83	97	77	100	4 of 4	None
II	83	92	83	100	4 of 4	None
III	100	90	100	90	4 of 4	None
IV	81	92	100	Not assessed	3 of 3	None
V	84	Not assessed	100	Not assessed	2 of 2	None

Note: Standards were fully covered if the number of tasks matched was at least 1 per form reviewed.

Depth-of-Knowledge Consistency. Depth-of-knowledge (DOK) measures the type of cognitive processing required by each performance task compared to the requirements implied by the content objectives. To make these judgments, reviewers first determined the DOK level for each objective of a standard using a rating scale (see Table 25 for Webb's guidelines). Next, as they reviewed items, panelists rated the level of processing needed to perform the task using the same DOK rating scales. We then compared these two separate judgments about cognitive complexity (one for the standard, one for the task) to determine the proportion of tasks written at the appropriate level. Webb refers to this comparison as *depth-of-knowledge consistency*.

Table 27 summarizes the depth-of-knowledge consistency results for each subject and level of the CAPA. Since reviewers evaluated depth-of-knowledge at the most specific level of the standards document (content objectives), the table refers to consistency between the tasks and the content objectives to which they were matched. Results are summarized in terms of the percent of tasks with cognitive complexity ratings at or above (more complex than) the rating for the corresponding content

objective. Webb's suggested criterion for this alignment indicator is the same as for a regular assessment – at least 50% of the tasks should have complexity ratings at or above the level of the corresponding content objective.

Table 27. Summary of Depth-of-Knowledge Results for CAPA ELA and Mathematics by Level

English-Language Arts					
	Percent of Tasks with DOK At or Above the Level of the Objectives per Standard			Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
Level	Reading	Writing	Listening and Speaking		
I	**30%**	93%	50%	2 of 3	Reading
II	90%	100%	95%	3 of 3	None
III	59%	94%	Not assessed	2 of 2	None
IV	73%	91%	68%	3 of 3	None
V	89%	93%	73%	3 of 3	None

Mathematics						
	Percent of Tasks with DOK At or Above the Level of the Objectives per Standard				Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
Level	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability		
I	68%	70%	81%	**0%**	3 of 4	Statistics
II	100%	71%	69%	79%	4 of 4	None
III	94%	87%	100%	87%	4 of 4	None
IV	93%	80%	100%	Not assessed	3 of 3	None
V	86%	Not assessed	87%	Not assessed	2 of 2	None

Note: Standards were covered at adequate depth if 50% of the tasks were at or above the complexity level for the matched content objective.

The results show that a sufficient number of tasks assess student knowledge at or above the level of the standards for CAPA levels with two exceptions. For Level I Reading of the CAPA ELA assessment, only 31% of tasks (or approximately 2.5 out of 8) were rated as the same or above the standards in cognitive complexity. For Level I Statistics, Data Analysis, and Probability of the CAPA, panelists decided that none of the tasks assessed students at or above the level of the CAPA standards.

Level I of the CAPA is administered to the most cognitively challenged students. Thus, students at this level have difficulty performing tasks requiring higher levels of cognitive complexity. California may wish to review the appropriateness of the cognitive

objectives expected of students at this level for Reading and for Statistics, Data Analysis, and Probability in particular. If they are appropriate, the performance tasks developed for this level should be modified to match these expectations.

Range-of-Knowledge. Range-of-knowledge measures how fully the tasks cover each of the content objectives within each standard. The assessed objectives within a standard should be linked with at least one performance task. Webb's minimum level of acceptability for range-of-knowledge correspondence is 50% per standard. This means that at least 50% of the objectives must be matched to one or more tasks.

Table 28 summarizes the range-of-knowledge results for each level of the CAPA. We computed the number of objectives covered for each standard separately for each panelist and then averaged across panelists to obtain the summary alignment indicator.

Table 28. Summary of Range-of-Knowledge Results for CAPA ELA and Mathematics by Level

English-Language Arts					
Level	Percent of Objectives per Standard Matched to at Least One Task			Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
	Reading	Writing	Listening and Speaking		
I	79%	50%	**16%**	2 of 3	Listening
II	65%	67%	**25%**	2 of 3	Listening
III	75%	63%	Not assessed	2 of 2	None
IV	73%	88%	70%	3 of 3	None
V	75%	100%	92%	3 of 3	None

Mathematics						
Level	Percent of Objectives per Standard Matched to at Least One Task				Number of Standards Assessed Adequately	Specific Standards Assessed Inadequately
	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability		
I	86%	65%	57%	100%	4 of 4	None
II	**36%**	63%	50%	50%	3 of 4	Number Sense
III	**45%**	88%	67%	63%	3 of 4	Number Sense
IV	96%	100%	100%	Not assessed	3 of 3	None
V	84%	Not assessed	100%	Not assessed	2 of 2	None

Note: Standards had an adequate range-of-knowledge if 50% of the objectives were matched by one or more of the performance tasks.

The range-of-knowledge results shown in Table 28 were based on ratings of the pool of available performance tasks. In an operational form, the test blueprint includes more than twice as many individual objectives as target tasks for several levels and standards. For this reason, it is difficult to assess half of the content objectives with a small number of tasks, unless a single performance task assesses more than one content objective simultaneously. In many cases, CAPA tasks do intentionally assess more than one objective.

Content coverage by the pool of available tasks is adequate with the exception of two levels for each of two standards. For Listening and Speaking at Levels I and II, and for Number Sense at Levels II and III, the test blueprints include many more content objectives than targeted tasks. The apparent shortcoming in these areas could be resolved by summarizing the specific content objectives into a smaller number of discrete categories.

Balance-of-Knowledge Representation. The fourth measure of alignment included in the Webb method is *balance-of-knowledge representation.* This measure indicates the number of tasks linked to each objective per standard. The number of tasks should be distributed rather evenly between the objectives for each standard to achieve good balance.

The content balance is determined by calculating an index, or score, for each standard[6]. According to Webb, the minimum acceptable index for a single standard is 70 (on a scale of 0 to 100, with 100 representing perfect balance). To be clear, a standard may include more objectives than reviewers actually linked to performance tasks. Thus, only those objectives actually used by the reviewers are included in calculations of the balance index.

Table 29 summarizes the results on balance of content representation for ELA and for mathematics. The table presents the balance index for each standard separately by level. As the table demonstrates, each standard surpassed the minimum criterion (index of 70 or higher) for each level of the CAPA. These findings indicate that the CAPA includes a comparable number of performance tasks corresponding to each of the content objectives across test versions.

[6] The exact formula for calculating the balance index is explained in detail in Norman Webb's (2005) alignment training manual: http://www.wcer.wisc.edu/WAT/index.aspx .

Table 29. Summary of Balance-of-Knowledge Representation Results for CAPA ELA and Mathematics by Level

English-Language Arts					
	Balance Index per Standard				
Level	Reading	Writing	Listening and Speaking	Standards with Adequate Balance	Standards with Limited Balance
I	71	94	100	3of 3	None
II	79	100	85	3 of 3	None
III	84	79	Not assessed	2 of 2	None
IV	80	90	84	3 of 3	None
V	83	87	87	3 of 3	None

Mathematics						
	Balance Index per Standard					
Level	Number Sense	Algebra and Functions	Measurement and Geometry	Statistics, Data Analysis, and Probability	Standards with Adequate Balance	Standards with Limited Balance
I	83	97	77	100	4 of 4	None
II	83	92	83	100	4 of 4	None
III	100	90	100	90	4 of 4	None
IV	81	92	100	Not assessed	3 of 3	None
V	84	Not assessed	100	Not assessed	2 of 2	None

Note: Standards had an acceptable balance if the index score was at least 70.

Panelist Ratings of Alignment CAPA Performance Tasks

This final section focuses on more qualitative outcomes from the alignment review. We present results on the ratings panelists provided on overall item alignment. Table 30 includes the percentage of items rated as a 3 (Highly Aligned) or 4 (Fully Aligned).

Table 30. Summary of Alignment Ratings for the CAPA Tasks

	Percentage of Tasks with Good Alignment Ratings (Rating of 3 or 4)	
Level	ELA	Math
I	90	88
II	81	89
III	73	84
IV	69	84
V	76	74

These outcomes on the overall alignment ratings of the performance tasks indicate that panelists considered many of the new field-test items to b matched appropriately to a CAPA standard. The ratings of tasks for several levels, however, do suggest that some tasks should be reviewed to increase alignment. Those CAPA levels with tasks requiring review are highlighted in red in Table 30.

As with the CSTs, panelists were given the opportunity to provide comments on individual tasks. Again, security issues prevent us from detailing the comments in this report; however, a few summary statements can be made. Overall, the comments given point to minor changes that could be implemented to improve the quality of a task, such as clarity in the steps a student must demonstrate or wording in the administration manual that may be confusing to test administrators. Several more substantial comments were made by most or all reviewers regarding the adaptability of some items to students with specific disabilities. For example, panelists noted that several tasks developed for ELA Levels III through V would be difficult for students with sensory impairments (i.e., visual, auditory). For both ELA and math at several levels, panelists noted the amount of experience required of students to respond to an item correctly, which may be an inappropriate assumption for that level.

Summary and Discussion of Alignment Results for CAPA

The alignment review of the CAPA for ELA and mathematics involved an evaluation of the new performance tasks field-tested in the 2007 administration compared to the CAPA standards. HumRRO applied the Webb method designed for alternate assessments to conduct the review. Overall, the alignment results for the CAPA were very positive. We present a summary of the alignment outcomes for the CAPA in this section.

Summary alignment judgments are based on Webb (1999). These summary judgments focus on the percentage of content standards represented well by the assessment. Webb outlined a scale with a range of potential alignment outcomes applied to each criterion. Under that scale:

- Fully aligned – assessments align to all content standards (100%);
- Highly aligned – assessments align to the majority of standards (70–90%)
- Partially aligned – assessments align well to some standards (50–69%); and
- Weakly aligned – assessments align to less than half the standards (below 50%).

Webb's alignment method does not allow for a *single* judgment of overall alignment across the four criteria. However, one can get a sense of overall alignment between the assessments and standards by looking at all of the criteria together.

Table 31 presents the summary alignment outcomes for ELA and math based on the above scale. The table includes the alignment judgment, along with the percentage of standards covered well by the assessment.

These results indicate that the new performance tasks assess the majority of CAPA standards well across levels for both ELA and math. Thus, the alignment between the CAPA performance tasks and standards is sufficient overall.

The exceptions to this statement pertain to a few standards written for ELA Levels I and II. Specifically, panelists judged the performance tasks to assess student knowledge of the Reading standard at a lower cognitive level (DOK) than expected. In addition, the tasks developed to assess Listening and Speaking did not cover the full range of content objectives (ROK) for this standard. As noted earlier, this outcome could be a result of the number of content expectations for Listening and Speaking compared with the small pool of items. Due to the lower degree of alignment for these tasks and standards, CDE should consider reevaluating content expectations for Reading and for Listening and Speaking at Levels I and II of the CAPA for content coverage and cognitive complexity.

Table 31. Summary Alignment Outcomes for CAPA English-language Arts and Mathematics on Webb Criteria

	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
	English-Language Arts			
ELA Level I	Fully aligned (100%)	Partially aligned (67%)	Partially aligned (67%)	Fully aligned (100%)
ELA Level II	Fully aligned (100%)	Fully aligned (100%)	Partially aligned (67%)	Fully aligned (100%)
ELA Level III	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)
ELA Level IV	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)
ELA Level V	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)
	Mathematics			
Math Level I	Fully aligned (100%)	Highly aligned (75%)	Fully aligned (100%)	Fully aligned (100%)
Math Level II	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (75%)	Fully aligned (100%)
Math Level III	Fully aligned (100%)	Fully aligned (100%)	Highly aligned (75%)	Fully aligned (100%)
Math Level IV	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)
Math Level V	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)	Fully aligned (100%)

Chapter 4: Alignment of the CSTs to the Performance Standards

After analyzing the alignment of each test form to the corresponding *content* standards, HumRRO reviewed each of the 2006 assessments included in this study for alignment with the *performance* standards. The key question addressed by these analyses is whether the tests provide useful information about students at each achievement or performance level. The CSTs results are reported in terms of five performance levels ranging from Far Below Basic up to Advanced.

In developing descriptors for each performance level, we (Wise, et al. 2007) used item maps developed by ETS. The item maps assigned each item to the lowest performance level at which most of the students could answer the item correctly. Some of the test items did not map cleanly onto the performance levels because the percent answering correctly was similar for two or more adjacent performance levels. For the present analyses, we revised the item mapping criteria so that every operational item was assigned a performance level, specifically to the lowest performance level at which 60% or more of the students answered correctly. Note that, in a few cases, fewer than 60% of students at the advanced level answered the item correctly. Thus, for purposes of this mapping, we created an additional, “Beyond Advanced” category to count these items.

Table 32 shows the number of items in each test mapped to each of the performance levels. There is not a universally accepted standard for the minimum number of items needed to provide information about students at a given achievement level. Webb suggests a minimum of at least six items per standard in evaluating content coverage. We propose a slightly higher minimum, 10 items, to demonstrate good coverage of a performance level. Under this criterion, the Basic and Proficient levels are well covered by each of the assessments studied.

Table 32. Distribution of Test Questions by Performance Level

Grade or Course	Far Below Basic	Below Basic	Basic	Proficient	Advanced	Beyond Advanced	Total
	Performance Level						
			English-Language Arts				
2		7	24	25	9		65
3	1	15	20	23	6		65
4*		4	25	26	18	2	75
5	1	9	21	28	14	2	75
6		5	28	29	13		75
7*		5	24	31	15		75
8		9	25	26	13	2	75
			Mathematics				
2	1	17	25	12	10		65
3	1	17	14	22	11		65
4	1	8	29	20	6	1	65
5		10	19	21	14	1	65
6		3	27	21	14		65
7		1	20	30	14		65
General Math		5	17	27	13	3	65
Algebra I		2	19	19	21	4	65
Geometry	1	5	24	22	13		65
Algebra II	1	8	21	20	15		65
Int. Math 1		4	22	15	24		65
Int. Math 2		3	12	23	20	7	65
Int. Math 3		7	14	21	19	4	65
			History-Social Science				
8		4	15	31	24	1	75
10		4	11	25	20		60
11		2	14	23	20	1	60
			Science				
5		3	11	31	13	2	60
8		2	18	12	16	12	60
10		1	17	25	12	5	60

* Essay questions in the grades 4 and 7 ELA tests are not included because they have multiple score levels that do not map cleanly onto the performance levels.

It is not surprising that there are very few items at the Far Below Basic level. Items that students at the lowest level can answer correctly provide little useful information for distinguishing performance at this level from performance at higher levels. For many of the tests, only a limited number of items mapped to the Below Basic level as well. *If the distinction between Far Below Basic and Below Basic levels is important, the test design might need to be modified slightly to include more items that Below Basic students can answer correctly*. As shown in Table 32, this concern applies particularly to the ELA tests at Grades 4, 6, and 7; to the mathematics tests at Grades 6 and 7 and most of the end-of-course tests; and, to each of the history and science tests

studied. In each case there are five or fewer items that most students at the Below Basic level can answer.

In a few cases, such as the Grade 3 ELA test and the Grade 4 mathematics test, the number of items at the advanced level is also somewhat limited. For the science test at Grades 8 and 10, the Integrated Mathematics II and III tests, and for the Algebra I test, four or more items map beyond the advanced level. CDE may wish to investigate whether these items assess content that even advanced students are not expected to master, or whether other factors make these items inappropriately difficult. The earlier finding that some of the science items had cognitive complexity (depth-of-knowledge) ratings beyond the ratings for the corresponding standard may also suggest opportunities to improve test alignment to both the content and performance standards.

For each of the tests reviewed, however, the numbers of Basic and Proficient items are well above the suggested minimum. This is entirely appropriate since the assessment of whether students have reached the proficient level is given prime importance under NCLB accountability.

From a measurement perspective, the number of items mapped to each performance level is less critical than the accuracy of the performance levels reported for each student. Measurement error, resulting from day-to-day fluctuation in student performance as well as from sampling test questions from a broader domain, is present in any assessment. Students very near the minimum score for a performance category will sometimes be “misclassified” because they are truly in between the levels described by adjacent categories. ETS reports a standard error of measurement performance, in scale score units, at the minimum scale score for each performance level. These error estimates, and technical details on their estimation, are described in technical documentation developed by ETS. (See http://www.startest.org/doclibrary.html under Appendix C: STAR CST Conditional Standard Errors of Measurement (CSEM)).

Table 33 shows the estimated error of measurement reported by ETS at the lower bound for each performance level (except, of course, for Far Below Basic) for the 2006 test forms. The size of the standard errors should be compared to the width of the performance level categories. In all cases, the Basic category is 50 points wide, ranging from 300 to 349. The width of the Below Basic level varies from about 30 to 65 points and the width of the Proficient category varies from 42 points up to 78 points. The yellow highlighting indicates those grades or courses for which the standard error of measurement may be larger than is desirable for sufficient accuracy.

Table 33. Error of Measurement at the Minimum for Each Performance Level

Grade / Course	Below Basic	Basic	Proficient	Advanced
	Performance Level			
	English-Language Arts			
2	14	13	14	18
3	15	14	16	20
4	13	13	13	15
5	13	13	13	15
6	13	13	13	16
7	14	13	13	16
8	15	14	14	17
	Mathematics			
2	18	17	19	24
3	17	17	18	22
4	15	14	15	20
5	19	19	19	23
6	17	15	16	20
7	17	16	16	20
General Math	17	16	17	20
Algebra I	20	18	18	22
Geometry	17	15	16	21
Algebra II	18	18	18	22
Int. Math I	20	19	19	23
Int. Math II	20	18	17	20
Int. Math III	22	20	20	24
	History-Social Science			
8	16	15	15	18
10	18	18	17	20
11	19	18	17	20
	Science			
5	17	16	16	19
8	25	23	23	24
10	16	15	16	18

One way of deciding whether a test has sufficient accuracy is to consider the probability that a student who is actually at the bottom of a given level will have an observed score for a given administration that is more than one performance level (50 points) above or below their true score. Assuming a normal distribution of errors, we will be 99% confident that measurement error for a given student will be less than 50 points if the standard error of measurement is less than 19.4. For this reason, we suggest that a standard error of less than 20 points is a reasonable target for the CSTs.[7] According to this criterion, the CSTs cover the performance levels accurately in almost all cases.

[7] It is possible, of course, to set different targets for classification accuracy or to consider the different widths of performance levels other than Basic.

Interestingly, the areas of concern are not at the Below Basic level, as suggested by the item map data, but at the Advanced level. As shown in Table 33, this point is evident particularly for mathematics.

The Grade 8 science test, with standard errors above 20 for each of the performance levels, is one where further investigation of test accuracy is warranted. The only other test with standard errors at or above 20 for the basic and proficient performance levels is the end-of-course test for Integrated Mathematics III.

In summary, both the analysis of the item maps and the review of error of measurement data demonstrate a reasonable coverage of the performance levels defined by California's achievement standards. Several tests could benefit from a further increase in accuracy, particularly Grade 8 science. In addition, instances where items were mapped beyond the advanced category should be reviewed. Nonetheless, coverage of the content standards should take precedence and items should not be dropped if they are good measures of content that even advanced students have not yet mastered well.

Chapter 5: Summary and Recommendations of Alignment Results

The purpose of the independent review of the California assessment system was to develop and evaluate evidence of alignment between the assessments, the academic content standards, and the achievement levels. The assessment-to-standards review evaluated the agreement between the 2007 CST and CAPA forms and the content standards. The assessment-to-performance level review examined the match between the CSTs and the performance levels adopted by the SBE for reporting assessment results.

The results of these reviews provide confirmation of the content validity of the CSTs and the CAPA for California overall. These results offer evidence to the USDE that California has established a rigorous and coherent assessment system for all students.

In this section, we present a cumulative synopsis of the alignment outcomes for the assessment-to-standards review and the assessment-to-performance level review. Finally, HumRRO provides brief recommendations to California on these outcomes.

Alignment of the Assessments to the California Content Standards

California Standards Tests (CSTs)

Table 34 summarizes alignment judgments for each of the CSTs for each of the Webb criteria. Alignment results are classified into four levels of acceptability:

- Fully aligned – assessments align to all content standards (100%);
- Highly aligned – assessments align to the majority of standards (70–90%)
- Partially aligned – assessments align well to some standards (50–69%); and
- Weakly aligned – assessments align to less than half the standards (below 50%).

The highlighted portions of the table reflect those subjects and grades with lower degrees of alignment between the assessments and content standards. The implications of the alignment outcomes will be discussed subsequently for each subject area separately.

Table 34. Summary Alignment Outcomes for the CSTs per Subject Area and Grade

	Summary Alignment Outcomes per Webb Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Knowledge Representation
Grade	English-language Arts Grade-Level Tests			
2	Highly aligned	Fully aligned	Fully aligned	Highly aligned
3	Fully aligned	Fully aligned	Fully aligned	Fully aligned
4	Fully aligned	Fully aligned	Fully aligned	Highly aligned
5	Fully aligned	Highly aligned	Fully aligned	Highly aligned
6	Fully aligned	Weakly aligned	Fully aligned	Highly aligned
7	Fully aligned	Highly aligned	Fully aligned	Fully aligned
8	Fully aligned	Partially aligned	Fully aligned	Fully aligned
Grade	Mathematics Grade-Level Tests			
2	Partially aligned	Partially aligned	Highly aligned	Fully aligned
3	Partially aligned	Highly aligned	Highly aligned	Fully aligned
4	Partially aligned	Fully aligned	Highly aligned	Fully aligned
5	Highly aligned	Highly aligned	Partially aligned	Fully aligned
6	Highly aligned	Highly aligned	Highly aligned	Highly aligned
7	Highly aligned	Partially aligned	Highly aligned	Fully aligned
Course	Math End-of-Course Tests			
Algebra I	Fully aligned	Fully aligned	Highly aligned	Highly aligned
Algebra II	Fully aligned	Highly aligned	Fully aligned	Fully aligned
Geometry	Fully aligned	Highly aligned	Fully aligned	Highly aligned
General Math	Fully aligned	Partially aligned	Fully aligned	Highly aligned
Course	Integrated Mathematics			
Int. Math I	Fully aligned	Weakly aligned	Partially aligned	Fully aligned
Int. Math II	Highly aligned	Partially aligned	Highly aligned	Fully aligned
Int. Math III	Highly aligned	Weakly aligned	Weakly aligned	Fully aligned
Grade	Science			
5	Fully aligned	Fully aligned	Highly aligned	Fully aligned
8	Highly aligned	Fully aligned	Highly aligned	Fully aligned
10	Highly aligned	Partially aligned	Highly aligned	Highly aligned
Grade	History-Social Science			
8	Fully aligned	Partially aligned	Weakly aligned	Fully aligned
10	Fully aligned	Weakly aligned	Highly aligned	Fully aligned
11	Fully aligned	Weakly aligned	Partially aligned	Fully aligned

English-Language Arts Grade Level Tests. The assessments for ELA demonstrated sufficient breadth of content coverage overall, as shown by the alignment outcomes on the categorical concurrence, range-of-knowledge, and balance-of-knowledge criteria. The cognitive complexity level of many test items does not correspond well with the cognitive expectations in the standards for Grades 6 and 8. For Grade 6 in particular, the majority of standards were assessed at a lower cognitive level by most test items, as determined by our panelists. The reason for this discrepancy could be attributed to the specific sample of items selected for inclusion in the 2007 test

forms, in which case a review of these items by the test developer may be warranted for modification. Alternatively, CDE may wish to review the expectations within the standards document. The majority of content objectives were rated as DOK Level 3 by panelists, while many of the items for Grades 6 and 8 were judged as DOK Level 1 or 2. Thie emphasis on higher cognitive processing in the content standards is not necessarily problematic especially at higher grade levels, nor is it inconsistent with the cognitive expectations found in many other states. However, the assessment and content standards should be consistent.

Mathematics Grade-Level Tests. For the Grades 2 through 7 math tests, one issue seems to be the number of items assigned to assess each content objective in the test blueprint, which produced low alignment outcomes on the categorical concurrence and range-of-knowledge criteria. Several factors could have contributed to these results. For grades 2 through 5 in particular, the test blueprint indicates that the majority of test items should assess Number Sense (49% to 58% of items), followed by Measurement and Geometry (20% to 22% of items). In comparison, the standard Statistics, Data Analysis, and Probability should be assessed by 6% to 11% of items across these grades. For Math Reasoning, the blueprint does not designate any items to assess this standard individually. Instead, the blueprint specifies that items should assess one of the other four standards in addition to Math Reasoning since this form of knowledge represents an implicit cognitive task across math domains, and, hence, is designated as “embedded” on the test blueprint.

These organizational features of the standards document clearly exhibits intentional design, meaning that California considered the emphasis given to some standards over others important. For this reason, the structure of the California Content Standards should be taken into account when considering the lower alignment numbers on categorical concurrence and range-of-knowledge representation for some grades. In other words, the content emphasis found in most of the math elementary and middle-grade math test forms reflect the content emphasis established in the standards/blueprints. Furthermore, the emphasis given to Number Sense and to Measurement and Geometry in early grades does correspond with common curriculum decisions across states, and it is supported by developmental evidence of student knowledge acquisition of mathematics (e.g., Jensen, 1993; Flavell, 2002; Sternberg & Ben-Zeev, 1996). Despite this evidence, it still is the case that the standard Statistics, Data Analysis, and Probability standard was assessed by a very small number of items. While the minimum criterion established for adequate categorical concurrence is essentially a guideline, Webb (1999) suggested that six items can “produce a reasonably reliable scale for estimating students’ mastery of content on that scale” (p. 7). Hoffman, Diaz, and Dickinson (2005) demonstrated that even seven items may not produce a reliable estimate of student-level scores. Ideally, a standard and its corresponding content objectives would be linked with an even larger number of items. However, practical constraints, such as time allotted for testing, limit test length.

The outcomes on Math Reasoning for the CSTs parallel those found for the CAHSEE in 2005. Math panelists in both alignment reviews had difficulty matching

items to this standard. The test blueprints for the CSTs and CAHSEE clearly specify that Math Reasoning should be assessed in addition to at least one other content standard. However, the requirements of NCLB indicate that that the “assessment should align fully to the content standards” (USDE, 2004). If CDE finds the current blueprint layout desirable and appropriate for California students, sufficient justification should be given.

The results for depth-of-knowledge consistency do warrant review for math Grades 2 and 7. In this case, the composition of the test forms does not reflect the intention of the standards in that the majority of items fell below the expected cognitive ability level defined in several of the grade-level standards.

Mathematics End-of-Course Tests. The alignment results for the math end-of-course tests for Algebra I, Algebra II, Geometry, and General Math was strong overall. The one exception occurred for General Math on depth-of-knowledge consistency with the finding that items assessed two content areas (Quantitative Relationships and Multistep Problems) below the level of the cognitive expectations laid out in the standards document. The DOK levels found in these standards seems appropriately distributed among the four DOK levels with the majority of content objectives split between Levels 2 and 3. In comparison, panelists rated test items matched to these standards as predominantly assessing student knowledge at Levels 1 and 2. Thus, some items may require modification to match the cognitive expectations of the standards.

The results for the Integrated Math end-of-course tests were not as solid. Specifically, all three course test forms showed low alignment with at least three standards each on depth-of-knowledge consistency and on range-of-knowledge correspondence. Thus, the level of cognitive complexity assessed is lower than expected by those standards, and items target a small number of content objectives within these standards. Concerning the range of content covered, CDE and the test developer should examine the test items and the corresponding content standards in detail to determine whether items should be redistributed among standards, or whether some content objectives are in fact necessary for assessment.

Science Tests. As Table 34 shows, the science assessments for Grades 5, 8, and 10 align well to the content standards on many dimensions. Nevertheless, some review of test items developed for individual standards within each grade should occur to demonstrate appropriate alignment to all aspects of the content standards. As noted in the results section, the science blueprints target a small number (five or fewer) of items to assess some standards. In addition, some content objectives for several content standards, such as Earth Science (Grade 5) or Investigation and Experimentation (all grades), correspond with a limited number of items. The consequence is reduced alignment outcomes on categorical concurrence and range-of-knowledge. With a small item pool, any variation in item quantity and distribution will affect alignment results. CDE may consider revisiting the number of distinct content

objectives in this category or working with the test developer to adjust the targeted number of items.

History-Social Science Tests. The results for the history-social science tests for Grades 8, 10, and 11 indicate that the breadth (categorical concurrence) of content covered by the assessments is appropriate and comparable to the content standards overall. However, panelists could match the items to a limited range of content within the standards for couple of standards each in Grades 10 and 11. Panelists found narrow content coverage by the assessment items for each of the Grade 8 content standards.

In addition, many of the history-social science items assessed student knowledge at a lower cognitive level than expected in the content standards. This inconsistency between item DOK and standard DOK emerged in all three of the grade-level tests, although items for Grades 10 and 11 appeared to exhibit more serious discrepancies in alignment. All of the operational items for Grades 8, 10, and 11 used in the 2007 test forms should undergo review to ensure better alignment on depth-of-knowledge consistency for future administrations.

Finally, the History and Social Science Analysis Skills standard received less coverage on the assessment for each grade-level test than intended in the test blueprints. Items assessing this standard were designed to target another history-social science standard as the primary standard. However, as noted in the test blueprint, 25% of the items should assess History and Social Science Analysis Skills in addition to another standard. From Table 16 (and also noted in more specific data tables in Appendix A), less than 15% of items were matched to this standard by our panelists. A review of the items intended to assess this standard may be warranted.

California Alternate Performance Assessments

Table 35 includes the summary alignment judgments for the CAPA on the Webb criteria. The highlighted portions of the table reflect the subject and CAPA levels with lower degrees of alignment. The implications of the alignment outcomes will be discussed subsequently for ELA and math separately.

Table 35. Summary Alignment Outcomes for the CAPA per Level

	Alignment Criteria			
	Categorical Concurrence	Depth-of-Knowledge Consistency	Range-of-Knowledge Correspondence	Balance-of-Representation
	English-Language Arts			
ELA Level I	Fully aligned	Partially aligned	Partially aligned	Fully aligned
ELA Level II	Fully aligned	Fully aligned	Partially aligned	Fully aligned
ELA Level III	Fully aligned	Fully aligned	Fully aligned	Fully aligned
ELA Level IV	Fully aligned	Fully aligned	Fully aligned	Fully aligned
ELA Level V	Fully aligned	Fully aligned	Fully aligned	Fully aligned
	Mathematics			
Math Level I	Fully aligned	Highly aligned	Fully aligned	Fully aligned
Math Level II	Fully aligned	Fully aligned	Highly aligned	Fully aligned
Math Level III	Fully aligned	Fully aligned	Highly aligned	Fully aligned
Math Level IV	Fully aligned	Fully aligned	Fully aligned	Fully aligned
Math Level V	Fully aligned	Fully aligned	Fully aligned	Fully aligned

As shown in the table, the new 2007 field-test performance tasks aligned quite well to the CAPA content standards for both ELA and math. These results indicate that the performance tasks sufficiently assess the breadth and depth of the alternate content standards as a whole.

Two areas that may require modification to further improve alignment pertain to ELA Levels I and II. Specifically, the ratings by the CAPA panelists suggest that the performance tasks for Level I may be inadequate on depth-of-knowledge and on range-of-knowledge. In evaluating DOK, panelists considered some performance tasks to assess students at a lower level of cognitive complexity for Reading than the content objectives included for that standard. This outcome may reflect the difficulty of developing appropriate assessment items for students at this level with the most significant cognitive challenges. CDE and the test developer should review the content standards to determine whether the expectations of these students are too high, or whether an adjustment to the performance tasks is required.

Regarding the range of the content covered by the tasks, panelists found that the performance tasks assessed a small number of content objectives on Listening and Speaking at both Levels I and II. As noted in the results section earlier, this standard includes far more content objectives than targeted tasks. While some tasks may assess more than one content objective, the disproportionate number of content expectations to tasks still remains. It is likely not appropriate to increase the number of performance tasks for these students to correct this discrepancy. Instead, CDE should consider reviewing the content objectives for possible merger or for in-class assessment.

The CAPA math results do not indicate any serious alignment problems across standards for any given level. However, a review of performance tasks developed for some individual standards may be worthwhile. For example, the performance tasks matched to the Statistics, Data Analysis, and Probability standard were rated as less cognitively challenging than the content expectations for Level I. For the Number Sense

standard, panelists matched less than half of the objectives to a performance task. Given that the current performance tasks under review are field-test items, CDE could take this opportunity to evaluate these content expectations and corresponding tasks in greater detail.

Alignment of the CSTs to the Performance Standards

There are no widely accepted criteria currently for assessing the alignment of assessments to the performance levels used in reporting results. HumRRO examined two ways of determining the extent to which the performance levels were covered in each assessment. These included counting the number of items mapped to each performance level and assessing the standard error of measurement at the cut-scores dividing the different performance levels.

The results of both approaches indicated good coverage of the performance levels for the CSTs. However, the outcomes also suggested marginal coverage of the performance levels for several specific assessments. Standard errors for the Grade 8 Science Assessment were at or above 20 at each of the performance levels. Standard error of measurement was also relatively large for the end-of-course test for Integrated Mathematics III. In addition, while coverage of the basic and proficient levels was good, the ELA tests for the earlier grades varied in their coverage of the below basic and advanced levels from one grade to the next.

HumRRO Recommendations

While the overall alignment picture for the California assessment system is positive, several areas require review to improve alignment between the assessments, content standards, and performance descriptors. In this section, HumRRO outlines recommendations to CDE and SBE on how to make these improvements.

Recommendations for the CSTs and California Content Standards

1. **Review the cognitive requirements (depth-of-knowledge) of the assessment items and the content standards to establish greater consistency.** This recommendation pertains to English-language arts (ELA) Grade 6 and 8; math Grades 2 and 7; the general math test; all three integrated math tests; and, all three history-social science tests. Increasing depth-of-knowledge consistency can be accomplished by modifying existing operational items and/or by modifying content expectations of the standards. Given that the content standards underwent thorough review prior to Board approval, working with the test contractor to bring the current operational items more in line with the standards is a reasonable course of action. Furthermore, while modifying the content standards may be appropriate in some cases, California should be cautious about reducing the cognitive demands of its content expectations. If California does choose to revise the content standards at some point, it may be worthwhile to evaluate the content

standards of other states whose assessment systems have been approved by the USDE to compare cognitive expectations. Alternatively, CDE and SBE could examine the structure of the content frameworks for the National Assessment of Educational Progress (NAEP). A number of states (e.g., Georgia, Kentucky, Missouri) have revised their content standards to model the NAEP content frameworks successfully.

2. **Expand the content coverage on the assessments to match the breadth of the content expectations in California Content Standards.** This recommendation pertains to the mathematics tests for Grades 2 through 5, the integrated math tests, and the history-social science tests. In evaluating the test blueprints, the narrow range of content coverage seems to stem from the limited number of items targeted for assessment in the first place. Necessarily, standardized assessments must limit the total number of items included on a single test form. Thus, HumRRO does not expect CDE and the test developer to lengthen the test to increase content coverage. Instead, several strategies working within the existing test forms may be possible: (a) redistribute items to increase content coverage on some standards; (b) consider whether some content is appropriate for standardized assessment or could be assessed in the classroom; or (c) consider modifying or merging related content objectives to increase the number of items targeting a given content area.

Recommendations for the CAPA and the Alternate Content Standards

HumRRO recommends that CDE and SBE consider the following recommendations for the CAPA based on the outcomes of the alignment review and analyses:

1. **Review the appropriateness of the number of content objectives for the alternate standards.** One of the challenges of alternate assessments and standards is condensing and modifying the content expectations developed for the regular assessment to more appropriately evaluate special needs students. At the same time, the alternate assessment should not be reduced to the extent that the expectations are entirely different from those laid out for the regular assessment. California appears to have made good progress on achieving this goal by including a reasonable set of content expectations linked to the full content standards. However, it may be the case that further review is necessary to consider the quantity of content objectives currently in place, particularly for ELA Levels I and II and Math Levels II and III.

2. **Review the cognitive requirements (depth-of-knowledge) of the performance tasks and the alternate standards to establish greater consistency.** This recommendation applies specifically to ELA Level I (Reading and Listening/Speaking) and Math Level I (Statistics, Probability, and Data Analysis). Both the new performance tasks and the standards

should be evaluated together to determine the appropriate degree of content expectations for students at this level.

Recommendations for the CSTs and Performance Levels

Coverage of the performance levels by test items was generally good for each of the CSTs, particularly for the Proficient and Basic categories. A few areas may benefit from further improvements, however. Some specific suggestions include:

1. **Review the assessments for Grade 8 science and Integrated Mathematics III for test accuracy due to larger standard errors of measurement**. To ensure that these tests measure student performance as accurately as possible, CDE should consider whether the present criteria established for the performance levels are appropriate. Two approaches may be useful in making this decision. First, the newly developed performance level descriptors (Wise et al., 2007) could be used to target item development to each performance level more distinctly. Alternatively, stricter standards might be established for test accuracy curves generated from field test information when new test forms are assembled.

2. **Review the number of items assigned to Far Below Basic and Below Basic to distinguish between these performance levels more clearly for each subject area.** Currently, many of the tests include a limited number of items not only at the Far Below Basic level but also at the Below Basic level. If these distinctions should be retained, assigning more items, at least to the Below Basic level, would be helpful to more accurately determine student performance at this level.

3. **Examine the number of items assigned to the Advanced level for ELA, math, and science.** Some grades and subject areas also include a limited number of items assigned to assess performance at the Advanced level. For ELA Grade 3 and for math Grade 4, the number of items assigned to the Advanced level is limited. For science Grades 8 and 10, Integrated Mathematics II and III, and for Algebra I, some items also appear to assess student knowledge beyond the Advanced level. Again, the new performance level descriptors might be used to improve the targeting of items to this performance level.

References

Almond, P., Filbin, J., Hall, J. and Tindal, J. (2005, June). *Alignment to Standards: Alternate Assessments, Expanded Benchmarks, Universal Design for Learning and Individualized Education Programs.* Paper presented at the meeting of the Council of Chief State School Officers, San Antonio, TX.

Browder, D. M., Flowers, C., Ahlgrim-Delzell, L., Karvonen, M., Spooner, F., & Algozzine, R. (2004). The alignment of alternate assessment content to academic and functional curricula. *Journal of Special Education, 37,* 211-223.

California Department of Education. (2002). *California standards tests: English-language arts blueprints.* http://www.cde.ca.gov/ta/tg/sr/blueprints.asp

California Department of Education. (2002). *California standards tests: History-social science blueprints.* http://www.cde.ca.gov/ta/tg/sr/blueprints.asp

California Department of Education. (2002*). California standards tests: Mathematics blueprints.* http://www.cde.ca.gov/ta/tg/sr/blueprints.asp

California Department of Education. (2002). *California standards tests: Science blueprints.* http://www.cde.ca.gov/ta/tg/sr/blueprints.asp

California Department of Education. (2003). *CAHSEE Language Arts Blueprint*. Retrieved from http://www.cde.ca.gov/ta/tg/hs/admin.asp.

California Department of Education. (2003). *CAHSEE Mathematics Blueprint*. Retrieved from http://www.cde.ca.gov/ta/tg/hs/admin.asp.

California Department of Education. (2006). *Request for proposals: California standards and assessment system independent evaluation study.*

Flavell, J. H.. (2002). Representation and concepts. In J. H. Flavell, Miller, P. H., and Miller S. A. (Eds), *Cognitive development* (pp. 99-138). Upper Saddle River, NJ: Prentice Hall.

Hoffman, R. G., Diaz, T. E., & Dickinson, E. R. (2005). *Idaho Standards Achievement Test: Independent calculations of reliability estimates, standard errors of measurement, classification accuracy, and classification consistency. Task 3.2.* (HumRRO Report No. FR-04-87). Alexandria, VA: Human Resources Research Organization.

Jensen, R. J. (1993). *Research ideas for the classroom: Early childhood mathematics.* New York: Macmillan.

No Child Left Behind Act of 2001. Public Law 107-110.

Roach, A. T., Elliott, S. N., & Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards: Evidence content validity of the Wisconsin Alternate Assessment. *Journal of Special Education*, *38,* 218–231.

Sternber, R. J., & BenZeev, T. (1996). *The nature of mathematical thinking.* Mahwah, NJ: Erlbaum.

Tindal, G. (2005). *Alignment of alternate assessments using the Webb system: An abbreviated version.* Paper presented at the meeting of the Council of Chief State School Officers, San Antonio, TX.

U.S. Department of Education. (2004, April). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001.* Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education. Retrieved from http://www.ed.gov/policy/elsec/guid/saaprguidance.doc .

U.S. Department of Education. (2005, May). *A user's guide to preparing submissions for the NCLB standards and assessment peer review.* Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education. Retrieved from http://www.ed.gov/policy/elsec/guid/saaprguidance.doc .

U.S. Department of Education. (2004, April). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001.* Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education. Retrieved from http://www.ed.gov/policy/elsec/guid/saaprguidance.doc .

Webb, N. L. (1997). *Research Monograph No. 6: Criteria for alignment of expectations and assessments in mathematics and science education.* Washington, D.C.: Council of Chief State Schools Officers.

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states.* (Research Monograph 18). Madison, WI: National Institute for Science Education and Council of Chief State School Officers. (ERIC Document Reproduction Service No. ED440852)

Webb, N. L. (2005). *Webb alignment tool: Training manua*l. Madison, WI: Wisconsin Center for Education Research. Available: http://www.wcer.wisc.edu/WAT/index.aspx.

Wise, L. L., Becker, D. E., Harris, C. D., Taylor, L. R., Johnstone, C. J., Miller, N. A. et al. (2005). *Independent evaluation of the California High School Exit Exam (CAHSEE): 2005 evaluation report* (FR-05-43). Alexandria, VA: Human Resources Research Organization. [Online]. Available: http://www.cde.ca.gov/ta/tg/hs/year6indepeval.asp .

Wise, L. L., Taylor, L. R., Becker, D. E., Gladden, F., Handy, K., Thacker, A., et al. (2007). Development of performance level descriptors for the California standards tests (CSTs) and high school exit exam (CAHSEE) (DTR-07-01). Alexandria, VA: Human Resources Research Organization.

Appendix A:
Alignment Results for Webb Criteria by Content Area and Grade Level for the 2006 CST Test Forms

English-Language Arts: Grades 2 through 8

Categorical Concurrence

The tables below present the results for ELA on categorical concurrence for each standard separated by grade level. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion.

Table A- 1. Categorical Concurrence for ELA Grade 2: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 2			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	22	19.83	3.43	Y
2	Reading Comprehension	15	16.67	1.21	Y
3	Literary Response and Analysis	6	4.33	1.75	N
4	Written and Oral English Language Conventions	14	14.83	1.33	Y
5	Writing Strategies	8	8.50	1.22	Y
	Total	65	64.16		
	Percent of standards with at least six items				80%

Table A- 2. Categorical Concurrence for ELA Grade 3: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 3			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	22	18.67	1.37	Y
2	Reading Comprehension	15	17.50	1.64	Y
3	Literary Response and Analysis	6	6.67	1.86	Y
4	Written and Oral English Language Conventions	13	13.17	0.41	Y
5	Writing Strategies	9	8.67	0.52	Y
	Total	65	64.68		
	Percent of standards with at least six items				100%

Table A- 3. Categorical Concurrence for ELA Grade 4: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 4			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	18	19.17	1.17	Y
2	Reading Comprehension	15	11.00	1.41	Y
3	Literary Response and Analysis	9	7.67	1.97	Y
4	Written and Oral English Language Conventions	18	18.17	0.75	Y
5	Writing Strategies	15	15.00	1.79	Y
	Total	75	71.01		
	Percent of standards with at least six items				100%

Table A- 4. Categorical Concurrence for ELA Grade 5: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 5			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	14	12.50	3.00	Y
2	Reading Comprehension	16	20.00	5.72	Y
3	Literary Response and Analysis	12	9.00	4.00	Y
4	Written and Oral English Language Conventions	17	17.25	1.26	Y
5	Writing Strategies	16	13.25	2.06	Y
	Total	75	72.00		
	Percent of standards with at least six items				100%

Table A- 5. Categorical Concurrence for ELA Grade 6: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 6			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	13	12.20	1.64	Y
2	Reading Comprehension	17	17.00	1.22	Y
3	Literary Response and Analysis	12	11.60	1.34	Y
4	Written and Oral English Language Conventions	16	17.00	2.35	Y
5	Writing Strategies	17	15.80	2.05	Y
	Total	75	73.60		
	Percent of standards with at least six items				100%

Table A- 6. Categorical Concurrence for ELA Grade 7: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 7			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	11	12.00	0.82	Y
2	Reading Comprehension	18	15.75	3.40	Y
3	Literary Response and Analysis	13	14.25	3.40	Y
4	Written and Oral English Language Conventions	16	19.50	1.73	Y
5	Writing Strategies	17	13.25	2.63	Y
	Total	75	74.75		
	Percent of standards with at least six items				100%

Table A- 7. Categorical Concurrence for ELA Grade 8: Mean Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		ELA Grade 8			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	9	8.33	1.63	Y
2	Reading Comprehension	18	18.50	0.84	Y
3	Literary Response and Analysis	15	15.83	2.71	Y
4	Written and Oral English Language Conventions	16	13.83	1.94	Y
5	Writing Strategies	17	12.83	2.56	Y
	Total	75	69.32		
	Percent of standards with at least six items				100%

Depth-of-Knowledge Consistency

The tables below present the results from the comparison between the depth-of-knowledge expected in the standards and the depth-of-knowledge assessed by items. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards along with the corresponding standard deviations. Results are separated by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A- 8. Depth-of-Knowledge Consistency for ELA Grade 2: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA Grade 2						
1	Word Analysis, Fluency, and Systematic Vocabulary Development	19.83	24	7.79	64	4.50	12	6.62	Y
2	Reading Comprehension	16.67	50	18.91	45	15.98	5	4.95	Y
3	Literary Response and Analysis	4.33	32	23.81	58	27.39	10	15.29	Y
4	Written and Oral English Language Conventions	14.83	0	0.00	88	25.82	12	25.82	Y
5	Writing Strategies	8.50	16	19.02	61	23.94	23	24.13	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 9. Depth-of-Knowledge Consistency for ELA Grade 3: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					ELA Grade 3				
1	Word Analysis, Fluency, and Systematic Vocabulary Development	18.67	31	18.09	63	13.91	6	4.87	Y
2	Reading Comprehension	17.50	17	14.56	54	14.11	29	13.36	Y
3	Literary Response and Analysis	6.67	0	0.00	44	18.02	56	18.02	Y
4	Written and Oral English Language Conventions	13.17	0	0.00	88	21.07	12	21.07	Y
5	Writing Strategies	8.67	16	23.64	56	20.75	23	16.19	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 10. Depth-of-Knowledge Consistency for ELA Grade 4: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					ELA Grade 4				
1	Word Analysis, Fluency, and Systematic Vocabulary Development	19.17	40	23.22	47	14.21	14	19.88	Y
2	Reading Comprehension	11.00	10	13.26	65	17.32	24	15.72	Y
3	Literary Response and Analysis	7.67	27	16.74	51	14.38	22	14.06	Y
4	Written and Oral English Language Conventions	18.17	10	8.08	68	28.47	22	33.33	Y
5	Writing Strategies	15.00	26	22.34	59	15.69	15	18.65	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 11. Depth-of-Knowledge Consistency for ELA Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA Grade 5						
1	Word Analysis, Fluency, and Systematic Vocabulary Development	12.50	40	43.04	40	36.46	51	8.38	Y
2	Reading Comprehension	20.00	31	22.48	31	25.30	62	15.22	Y
3	Literary Response and Analysis	9.00	34	19.98	34	14.51	54	20.56	Y
4	Written and Oral English Language Conventions	17.25	64	44.87	64	45.02	35	2.94	N
5	Writing Strategies	13.25	19	38.46	19	40.46	65	30.77	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								80%

Table A-12. Depth-of-Knowledge Consistency for ELA Grade 6: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA Grade 6						
1	Word Analysis, Fluency, and Systematic Vocabulary Development	12.20	61	22.65	35	18.76	4	8.13	N
2	Reading Comprehension	17.00	52	15.23	48	15.23	0	0.00	N
3	Literary Response and Analysis	11.60	64	15.53	36	15.53	0	0.00	N
4	Written and Oral English Language Conventions	17.00	86	25.64	14	25.64	0	0.00	N
5	Writing Strategies	15.80	49	28.24	46	27.83	6	5.64	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								20%

Table A-13. Depth-of-Knowledge Consistency for ELA Grade 7: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					ELA Grade 7				
1	Word Analysis, Fluency, and Systematic Vocabulary Development	12.00	18	27.57	67	20.87	14	11.88	Y
2	Reading Comprehension	15.75	18	14.78	76	8.37	6	7.20	Y
3	Literary Response and Analysis	14.25	31	14.36	66	16.51	3	3.40	Y
4	Written and Oral English Language Conventions	19.50	74	20.25	26	20.25	0	0.00	N
5	Writing Strategies	13.25	40	27.77	57	25.72	3	3.73	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								80%

Table A-14. Depth-of-Knowledge Consistency for ELA Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					ELA Grade 8				
1	Word Analysis, Fluency, and Systematic Vocabulary Development	8.33	64	15.01	35	17.55	2	4.08	N
2	Reading Comprehension	18.50	38	12.28	55	14.46	7	7.53	Y
3	Literary Response and Analysis	15.83	33	16.40	56	21.54	11	17.39	Y
4	Written and Oral English Language Conventions	13.83	53	20.43	44	18.45	3	4.52	N
5	Writing Strategies	12.83	48	17.02	52	17.02	0	0.00	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								60%

Range-of-Knowledge Correspondence

The tables below present the results on the range of content covered by the test items for ELA. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 15. Range-of-Knowledge for ELA Grade 2: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 2			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	9	19.83	7.17	0.75	80	Y
2	Reading Comprehension	7	16.67	5.00	0.89	71	Y
3	Literary Response and Analysis	4	4.33	2.83	0.75	71	Y
4	Written and Oral English Language Conventions	8	14.83	7.17	0.41	90	Y
5	Writing Strategies	3	8.50	3.00	0.00	100	Y
	Total	31	64.15	25.17			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 16. Range-of-Knowledge for ELA Grade 3: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 3			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	7	18.67	7.17	0.41	102	Y
2	Reading Comprehension	7	17.50	5.00	0.89	71	Y
3	Literary Response and Analysis	6	6.67	4.33	0.82	72	Y
4	Written and Oral English Language Conventions	9	13.17	8.17	0.41	91	Y
5	Writing Strategies	3	8.67	3.00	0.00	100	Y
	Total	32	64.68	27.67			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 17. Range-of-Knowledge for ELA Grade 4: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 4			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	5	19.17	4.83	0.98	97	Y
2	Reading Comprehension	6	11.00	4.83	0.75	81	Y
3	Literary Response and Analysis	5	7.67	3.67	0.52	73	Y
4	Written and Oral English Language Conventions	7	18.17	6.83	0.41	98	Y
5	Writing Strategies	8	15.00	6.50	0.55	81	Y
	Total	31	71.01	26.66			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 18. Range-of-Knowledge for ELA Grade 5: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 5			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	4	12.50	3.75	0.50	94	Y
2	Reading Comprehension	5	20.00	4.75	0.50	95	Y
3	Literary Response and Analysis	7	9.00	4.50	1.00	64	Y
4	Written and Oral English Language Conventions	5	17.25	5.00	0.82	100	Y
5	Writing Strategies	5	13.25	4.50	0.58	90	Y
	Total	26	72.00	22.50			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 19. Range-of-Knowledge for ELA Grade 6: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 6			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	4	12.20	3.60	0.55	90	Y
2	Reading Comprehension	8	17.00	6.20	1.10	78	Y
3	Literary Response and Analysis	8	11.60	6.20	1.30	78	Y
4	Written and Oral English Language Conventions	5	17.00	5.00	0.00	100	Y
5	Writing Strategies	5	15.80	4.20	0.84	84	Y
	Total	30	73.60	25.20			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 20. Range-of-Knowledge for ELA Grade 7: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 7			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	3	12.00	3.00	0.00	100	Y
2	Reading Comprehension	6	15.75	4.75	1.26	79	Y
3	Literary Response and Analysis	6	14.25	5.25	0.96	88	Y
4	Written and Oral English Language Conventions	7	19.50	6.75	0.50	96	Y
5	Writing Strategies	6	13.25	4.75	0.96	79	Y
	Total	28	74.75	24.50			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 21. Range-of-Knowledge for ELA Grade 8: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				ELA Grade 8			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	3	8.33	2.67	0.52	89	Y
2	Reading Comprehension	7	18.50	5.83	1.60	83	Y
3	Literary Response and Analysis	7	15.83	5.83	1.17	83	Y
4	Written and Oral English Language Conventions	6	13.83	5.33	0.52	89	Y
5	Writing Strategies	4	15.83	4.00	0.00	100	Y
	Total	27	72.32	23.66			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Balance-of-Knowledge Representation

The tables below present the mean balance index calculated per standard for ELA per grade level. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 22. Balance-of-Knowledge Representation for ELA Grade 2: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA Grade 2			
1 Word Analysis, Fluency, and Systematic Vocabulary Development	9	7.17	19.83	31	78	3.42	Y
2 Reading Comprehension	7	5.00	16.67	26	66	8.57	N
3 Literary Response and Analysis	4	2.83	4.33	7	91	10.15	Y
4 Written and Oral English Language Conventions	8	7.17	14.83	23	90	9.45	Y
5 Writing Strategies	3	3.00	8.50	13	81	9.43	Y
Total	31	25.17	64.15				
Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 23. Balance-of-Knowledge Representation for ELA Grade 3: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA Grade 3			
1 Word Analysis, Fluency, and Systematic Vocabulary Development	7	7.17	18.67	29	76	5.00	Y
2 Reading Comprehension	7	5.00	17.50	27	71	7.54	Y
3 Literary Response and Analysis	6	4.33	6.67	10	83	2.37	Y
4 Written and Oral English Language Conventions	9	8.17	13.17	20	81	1.37	Y
5 Writing Strategies	3	3.00	8.67	13	87	2.87	Y
Total	32	27.67	64.68				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 24. Balance-of-Knowledge Representation for ELA Grade 4: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA Grade 4			
1 Word Analysis, Fluency, and Systematic Vocabulary Development	5	4.83	19.17	27	59	9.37	N
2 Reading Comprehension	6	4.83	11.00	16	86	5.23	Y
3 Literary Response and Analysis	5	3.67	7.67	11	84	6.22	Y
4 Written and Oral English Language Conventions	7	6.83	18.17	26	81	3.01	Y
5 Writing Strategies	8	6.50	15.00	21	75	6.96	Y
Total	31	26.66	71.01				
Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 25. Balance-of-Knowledge Representation for ELA Grade 5: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					ELA Grade 5			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	4	3.75	12.50	17	80	4.56	Y
2	Reading Comprehension	5	4.75	20.00	28	73	5.19	Y
3	Literary Response and Analysis	7	4.50	9.00	12	84	4.49	Y
4	Written and Oral English Language Conventions	5	5.00	17.25	24	85	4.74	Y
5	Writing Strategies	5	4.50	13.25	18	56	9.04	N
	Total	26	22.50	72.00				
	Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 26. Balance-of-Knowledge Representation for ELA Grade 6: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					ELA Grade 6			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	4	3.60	12.20	17	77	3.42	Y
2	Reading Comprehension	8	6.20	17.00	23	75	6.62	Y
3	Literary Response and Analysis	8	6.20	11.60	16	75	3.30	Y
4	Written and Oral English Language Conventions	5	5.00	17.00	23	81	7.48	Y
5	Writing Strategies	5	4.20	15.80	21	67	7.36	N
	Total	30	25.20	73.60				
	Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 27. Balance-of-Knowledge Representation for ELA Grade 7: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					ELA Grade 7			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	3	3.00	12.00	16	86	3.25	Y
2	Reading Comprehension	6	4.75	15.75	21	76	5.16	Y
3	Literary Response and Analysis	6	5.25	14.25	19	79	6.75	Y
4	Written and Oral English Language Conventions	7	6.75	19.50	26	85	6.37	Y
5	Writing Strategies	6	4.75	13.25	18	80	0.81	Y
	Total	28	24.50	74.75				
	Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 28. Balance-of-Knowledge Representation for ELA Grade 8: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					ELA Grade 8			
1	Word Analysis, Fluency, and Systematic Vocabulary Development	3	2.67	8.33	12	79	6.81	Y
2	Reading Comprehension	7	5.83	18.50	26	71	3.05	Y
3	Literary Response and Analysis	7	5.83	15.83	22	75	4.52	Y
4	Written and Oral English Language Conventions	6	5.33	13.83	19	86	4.49	Y
5	Writing Strategies	4	4.00	15.83	22	79	5.55	Y
	Total	27	23.66	72.32				
	Percentage of standards with a balance of representation index of 70 or greater							100%

Mathematics: Grades 2 through 7

Categorical Concurrence

The tables below present the results for math 2 through 7 on categorical concurrence for each standard. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion.

Table A- 29. Categorical Concurrence for Math Grade 2: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 2		
1	Number Sense	38	37.50	0.58	Y
2	Algebra and Functions	6	5.75	1.50	N
3	Measurement and Geometry	14	13.25	0.96	Y
4	Statistics, Data Analysis, And Probability	7	7.00	0.82	Y
5	Mathematical Reasoning	Embedded	3.00	0.00	N
	Total	65	66.50		
	Percent of standards with at least six items				60%

Table A- 30. Categorical Concurrence for Math Grade 3: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 3		
1	Number Sense	32	32.25	1.50	Y
2	Algebra and Functions	12	11.00	1.41	Y
3	Measurement and Geometry	16	16.25	0.96	Y
4	Statistics, Data Analysis, And Probability	5	3.75	0.96	N
5	Mathematical Reasoning	Embedded	2.00	1.00	N
	Total	65	65.25		
	Percent of standards with at least six items				60%

Table A- 31. Categorical Concurrence for Math Grade 4: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 4		
1	Number Sense	31	30.25	2.06	Y
2	Algebra and Functions	18	15.75	2.36	Y
3	Measurement and Geometry	12	11.75	1.89	Y
4	Statistics, Data Analysis, And Probability	4	4.50	0.58	N
5	Mathematical Reasoning	Embedded	5.33	3.06	N
	Total	65	67.58		
	Percent of standards with at least six items				60%

Table A- 32. Categorical Concurrence for Math Grade 5: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 5		
1	Number Sense	29	29.00	0.00	Y
2	Algebra and Functions	17	17.25	1.26	Y
3	Measurement and Geometry	15	15.00	0.00	Y
4	Statistics, Data Analysis, And Probability	4	3.25	1.26	N
5	Mathematical Reasoning	Embedded	6.00	2.16	Y
	Total	65	70.50		
	Percent of standards with at least six items				80%

Table A- 33. Categorical Concurrence for Math Grade 6: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 6		
1	Number Sense	25	23.00	3.16	Y
2	Algebra and Functions	19	19.75	2.06	Y
3	Measurement and Geometry	10	9.75	0.50	Y
4	Statistics, Data Analysis, And Probability	11	11.00	0.00	Y
5	Mathematical Reasoning	Embedded	7.33	3.79	Y
	Total	65	70.83		
	Percent of standards with at least six items				100%

Table A- 34. Categorical Concurrence for Math Grade 7: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Math Grade 7		
1	Number Sense	22	20.75	1.26	Y
2	Algebra and Functions	25	22.25	4.57	Y
3	Measurement and Geometry	13	15.75	3.50	Y
4	Statistics, Data Analysis, And Probability	5	4.75	0.50	N
5	Mathematical Reasoning	Embedded	8.00	6.08	Y
	Total	65	71.50		
	Percent of standards with at least six items				80%

Depth-of-Knowledge Consistency

The tables below present the results for depth-of-knowledge consistency. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A-35. Depth-of-Knowledge Consistency for Math Grade 2: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math Grade 2						
1	Number Sense	37.50	43	15.36	46	12.21	11	5.90	Y
2	Algebra and Functions	5.75	82	5.16	18	5.16	0	0.00	N
3	Measurement and Geometry	13.25	29	22.15	42	18.61	29	20.23	Y
4	Statistics, Data Analysis, And Probability	7.00	8	16.67	36	24.05	56	34.26	Y
5	Mathematical Reasoning	3.00	100	0.00	0	0.00	0	0.00	N
	Percent of standards with 50% of item DOK at or above objective DOK:								60%

Table A-36. Depth-of-Knowledge Consistency for Math Grade 3: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math Grade 3						
1	Number Sense	32.25	22	3.86	38	12.61	40	8.78	Y
2	Algebra and Functions	11.00	43	18.00	50	16.69	7	4.74	Y
3	Measurement and Geometry	16.25	21	9.80	56	13.29	23	11.19	Y
4	Statistics, Data Analysis, And Probability	3.75	0	0.00	58	50.00	42	50.00	Y
5	Mathematical Reasoning	2.00	56	50.92	44	50.92	0	0.00	N
	Percent of standards with 50% of item DOK at or above objective DOK:								80%

Table A- 37. Depth-of-Knowledge Consistency for Math Grade 4: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Math Grade 4				
1	Number Sense	30.25	20	10.61	53	12.47	27	7.60	Y
2	Algebra and Functions	15.75	23	19.39	52	28.54	24	33.70	Y
3	Measurement and Geometry	11.75	29	11.95	63	9.54	8	11.79	Y
4	Statistics, Data Analysis, And Probability	4.50	45	5.77	31	25.29	24	20.56	Y
5	Mathematical Reasoning	5.33	21	26.02	58	52.04	21	26.02	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 38. Depth-of-Knowledge Consistency for Math Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Math Grade 5				
1	Number Sense	29.00	37	18.76	30	14.18	33	11.44	Y
2	Algebra and Functions	17.25	27	14.01	36	12.00	37	7.74	Y
3	Measurement and Geometry	15.00	32	6.38	48	20.64	20	25.53	Y
4	Statistics, Data Analysis, And Probability	3.25	21	25.00	51	35.00	28	37.86	Y
5	Mathematical Reasoning	6.00	8	9.69	58	29.10	34	24.19	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 39. Depth-of-Knowledge Consistency for Math Grade 6: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Math Grade 6				
1	Number Sense	23.00	20	16.59	58	3.64	22	15.82	Y
2	Algebra and Functions	19.75	27	7.23	53	18.37	19	14.86	Y
3	Measurement and Geometry	9.75	28	10.00	31	24.22	41	14.32	Y
4	Statistics, Data Analysis, And Probability	11.00	0	0.00	52	4.55	48	4.55	Y
5	Mathematical Reasoning	7.33	38	14.42	54	3.75	7	12.83	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 40. Depth-of-Knowledge Consistency for Math Grade 7: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Math Grade 7				
1	Number Sense	20.75	19	11.55	57	14.60	24	17.72	Y
2	Algebra and Functions	22.25	56	32.44	31	12.18	13	20.55	N
3	Measurement and Geometry	15.75	40	26.80	43	14.57	18	18.31	Y
4	Statistics, Data Analysis, And Probability	4.75	49	33.26	36	33.51	15	19.15	Y
5	Mathematical Reasoning	8.00	43	40.41	26	22.19	31	53.89	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								80%

Range-of-Knowledge Correspondence

The tables below present the results on the range of content covered by the test items for math 2 through 7. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 41. Range-of-Knowledge for Math Grade 2: Mean Percent Objectives per Standard Linked with Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math Grade 2			
1 Number Sense	15	37.50	12.75	0.50	91	Y
2 Algebra and Functions	3	5.75	2.25	0.96	75	Y
3 Measurement and Geometry	7	13.25	6.25	0.96	89	Y
4 Statistics, Data Analysis, And Probability	6	7.00	3.75	0.50	94	Y
5 Mathematical Reasoning	4	3.00	1.00	0.00	25	N
Total	35	66.50	26			
Percentage of standards with 50% of objectives linked to at least one item						80%

Table A- 42. Range-of-Knowledge for Math Grade 3: Mean Percent Objectives per Standard Linked with Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math Grade 3			
1 Number Sense	17	32.25	15.25	0.96	95	Y
2 Algebra and Functions	7	11.00	6.00	0.00	86	Y
3 Measurement and Geometry	10	16.25	8.75	0.50	88	Y
4 Statistics, Data Analysis, And Probability	4	3.75	2.75	0.50	92	Y
5 Mathematical Reasoning	11	2.00	1.33	0.58	12	N
Total	47	65.25	34.08			
Percentage of standards with 50% of objectives linked to at least one item						80%

Table A- 43. Range-of-Knowledge for Math Grade 4: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Math Grade 4			
1	Number Sense	17	30.25	13.75	0.50	86	Y
2	Algebra and Functions	7	15.75	5.25	1.71	75	Y
3	Measurement and Geometry	15	11.75	8.25	0.96	55	Y
4	Statistics, Data Analysis, And Probability	5	4.50	3.50	0.58	70	Y
5	Mathematical Reasoning	11	5.33	2.67	1.15	24	N
	Total	55	67.58	33.42			
	Percentage of standards with 50% of objectives linked to at least one item						80%

Table A- 44. Range-of-Knowledge for Math grade 5: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Math Grade 5			
1	Number Sense	10	29.00	9.25	0.50	93	Y
2	Algebra and Functions	5	17.25	4.00	0.00	80	Y
3	Measurement and Geometry	7	15.00	6.75	0.50	96	Y
4	Statistics, Data Analysis, And Probability	7	3.25	2.25	0.50	45	N
5	Mathematical Reasoning	11	6.00	2.00	0.82	18	N
	Total	38	70.50	24.25			
	Percentage of standards with 50% of objectives linked to at least one item						60%

Table A- 45. Range-of-Knowledge for Math Grade 6: Mean Percent Objectives per Standard Linked with Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math Grade 6			
1 Number Sense	8	23.00	7.00	0.00	88	Y
2 Algebra and Functions	9	19.75	6.25	0.96	69	Y
3 Measurement and Geometry	6	9.75	4.50	0.58	75	Y
4 Statistics, Data Analysis, And Probability	14	11.00	4.75	0.50	53	Y
5 Mathematical Reasoning	13	7.33	4.00	2.65	40	N
Total	44	70.83	26.50			
Percentage of standards with 50% of objectives linked to at least one item						60%

Table A- 46. Range-of-Knowledge for Math Grade 7: Mean Percent Objectives per Standard Linked with Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math Grade 7			
1 Number Sense	12	20.75	10.00	0.82	83	Y
2 Algebra and Functions	13	22.25	7.50	0.58	58	Y
3 Measurement and Geometry	13	15.75	8.50	1.00	71	Y
4 Statistics, Data Analysis, And Probability	3	4.75	2.50	0.58	83	Y
5 Mathematical Reasoning	14	8.00	2.33	0.58	17	N
Total	54	71.50	30.83			
Percentage of standards with 50% of objectives linked to at least one item						80%

Balance-of-Knowledge Representation

The tables below present the mean balance index calculated per standard for math per grade level. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 47. Balance-of-Knowledge Representation for Math Grade 2: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 2			
1 Number Sense	15	12.75	37.50	57	81	2.34	Y
2 Algebra and Functions	3	2.25	5.75	9	73	18.11	Y
3 Measurement and Geometry	7	6.25	13.25	20	81	3.03	Y
4 Statistics, Data Analysis, And Probability	6	3.75	7.00	11	92	10.26	Y
5 Mathematical Reasoning	4	1.00	3.00	4	100	0.00	Y
Total	35	26	66.50				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 48. Balance-of-Knowledge Representation for Math Grade 3: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 3			
1 Number Sense	17	15.25	32.25	49	78	2.75	Y
2 Algebra and Functions	7	6.00	11.00	17	72	1.72	Y
3 Measurement and Geometry	10	8.75	16.25	25	82	4.10	Y
4 Statistics, Data Analysis, And Probability	4	2.75	3.75	6	87	14.93	Y
5 Mathematical Reasoning	11	1.33	2.00	3	94	9.62	Y
Total	47	34.08	65.25				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 49. Balance-of-Knowledge Representation for Math Grade 4: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 4			
1 Number Sense	17	13.75	30.25	45	82	2.59	Y
2 Algebra and Functions	7	5.25	15.75	24	77	8.51	Y
3 Measurement and Geometry	15	8.25	11.75	18	82	1.32	Y
4 Statistics, Data Analysis, And Probability	5	3.50	4.50	7	89	7.62	Y
5 Mathematical Reasoning	11	2.67	5.33	8	81	17.35	Y
Total	55	33.42	67.58				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 50. Balance-of-Knowledge Representation for Math Grade 5: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 5			
1 Number Sense	10	9.25	29.00	41	73	3.35	Y
2 Algebra and Functions	5	4.00	17.25	24	80	4.00	Y
3 Measurement and Geometry	7	6.75	15.00	21	79	3.17	Y
4 Statistics, Data Analysis, And Probability	5	2.25	3.25	5	85	11.06	Y
5 Mathematical Reasoning	11	2.00	6.00	8	81	16.58	Y
Total	38	24.25	70.50				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 51. Balance-of-Knowledge Representation for Math Grade 6: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 6			
1 Number Sense	8	7.00	23.00	33	78	2.15	Y
2 Algebra and Functions	9	6.25	19.75	28	68	4.38	N
3 Measurement and Geometry	6	4.50	9.75	14	78	3.33	Y
4 Statistics, Data Analysis, And Probability	14	4.75	11.00	16	81	3.85	Y
5 Mathematical Reasoning	13	4.00	7.33	10	84	14.56	Y
Total	50	26.50	70.83				
Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 52. Balance-of-Knowledge Representation for Math Grade 7: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math Grade 7			
1 Number Sense	12	10.00	20.75	30	75	5.63	Y
2 Algebra and Functions	13	7.50	22.25	32	73	4.80	Y
3 Measurement and Geometry	12	8.50	15.75	23	73	5.87	Y
4 Statistics, Data Analysis, And Probability	3	2.50	4.75	7	81	8.32	Y
5 Mathematical Reasoning	14	2.33	8.00	11	70	13.33	Y
Total	54	30.83	71.50				
Percentage of standards with a balance of representation index of 70 or greater							100%

Mathematics: Grade 8 End-of-Course Exams

Categorical Concurrence

The tables below present the results for Algebra I, Algebra II, Geometry, and General Math end-of-course tests on categorical concurrence per standard. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion.

Table A- 53. Categorical Concurrence for Algebra I: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Algebra I		
1	Number Properties, Operations, and Linear Equations	17	16.25	1.98	Y
2	Graphing and Systems of Linear Equations	14	14.25	1.98	Y
3	Quadratics and Polynomials	21	21.63	1.77	Y
4	Functions and Rational Expressions	13	12.00	1.85	Y
	Total	65	64.13		
	Percent of standards with at least six items				100%

Table A- 54. Categorical Concurrence for Algebra II: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Algebra II		
1	Polynomials and Rational Expressions	19	14.00	2.00	Y
2	Quadratics, Conics, and Complex Numbers	16	11.60	0.55	Y
3	Exponents and Logarithms	16	24.40	1.14	Y
4	Series, Combinatorics, Probability and Statistics	14	13.80	1.30	Y
	Total	65	63.80		
	Percent of standards with at least six items				100%

Table A- 55. Categorical Concurrence for Geometry: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Geometry		
1	Logic and Geometric Proofs	23	21.71	2.36	Y
2	Volume and Area Formulas	11	11.29	1.50	Y
3	Angle Relationships, Constructions, and Lines	16	16.43	2.07	Y
4	Trigonometry	15	15.43	1.40	Y
	Total	65	64.86		
	Percent of standards with at least six items				100%

Table A- 56. Categorical Concurrence for General Math: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			General Math		
1	Rational Numbers	14	14.86	3.72	Y
2	Exponents, Powers, and Roots	10	10.43	2.88	Y
3	Quantitative Relationships and Evaluating Expressions	11	10.14	1.21	Y
4	Multistep Problems, Graphing, and Functions	9	9.00	0.82	Y
5	Measurement and Geometry	12	12.00	1.91	Y
6	Statistics, Data Analysis, and Probability	9	8.57	0.53	Y
	Total	65	65.00		
	Percent of standards with at least six items				100%

Depth-of-Knowledge Consistency

The tables below present the results for depth-of-knowledge consistency. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A- 57. Depth-of-Knowledge Consistency for Algebra I: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Algebra I				
1	Number Properties, Operations, and Linear Equations	16.25	32	0.20	51	0.15	17	0.11	Y
2	Graphing and Systems of Linear Equations	14.25	30	0.13	58	0.10	12	0.16	Y
3	Quadratics and Polynomials	21.63	29	0.14	56	0.09	15	0.14	Y
4	Functions and Rational Expressions	12.00	31	0.21	56	0.11	13	0.13	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 58. Depth-of-Knowledge Consistency for Algebra II: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Algebra II				
1	Polynomials and Rational Expressions	14.00	27	0.24	53	0.15	20	0.13	Y
2	Quadratics, Conics, and Complex Numbers	11.60	49	0.21	45	0.14	68	0.07	Y
3	Exponents and Logarithms	24.40	74	0.15	22	0.11	03	0.06	N
4	Series, Combinatorics, Probability and Statistics	13.80	16	0.18	59	0.15	25	0.12	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								75%

Table A- 59. Depth-of-Knowledge Consistency for Geometry: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Geometry				
1	Logic and Geometric Proofs	21.71	57	0.22	36	0.15	08	0.10	N
2	Volume and Area Formulas	11.29	39	0.30	36	0.18	25	0.19	Y
3	Angle Relationships, Constructions, and Lines	16.43	21	0.12	55	0.14	24	0.11	Y
4	Trigonometry	15.43	35	0.22	50	0.18	15	0.17	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								75%

Table A- 60. Depth-of-Knowledge Consistency for General Math: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					General Math				
1	Rational Numbers	14.86	14	0.11	54	0.15	32	0.18	Y
2	Exponents, Powers, and Roots	10.43	24	0.24	34	0.19	41	0.24	Y
3	Quantitative Relationships and Evaluating Expressions	10.14	52	0.20	38	0.17	10	0.15	N
4	Multistep Problems, Graphing, and Functions	9.00	55	0.18	28	0.13	18	0.15	N
5	Measurement and Geometry	12.00	46	0.19	34	0.11	20	0.13	Y
6	Statistics, Data Analysis, and Probability	8.57	20	0.20	51	0.13	29	0.16	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								67%

Range-of-Knowledge Correspondence

The tables below present the results on the range of content covered by the test items for Algebra I, Algebra II, Geometry, and General Math. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 61. Range-of-Knowledge for Algebra I: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
						Algebra I	
1	Number Properties, Operations, and Linear Equations	11	16.25	4.62	0.52	42	N
2	Graphing and Systems of Linear Equations	4	14.25	3.88	0.35	97	Y
3	Quadratics and Polynomials	8	21.63	7.62	0.52	95	Y
4	Functions and Rational Expressions	6	12.00	5.00	0.00	83	Y
	Total	29	64.13	21.12			
	Percentage of standards with 50% of objectives linked to at least one item						75%

Table A- 62. Range-of-Knowledge for Algebra II: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
						Math Grade	
1	Polynomials and Rational Expressions	5	14.00	4.20	0.45	84	Y
2	Quadratics, Conics, and Complex Numbers	7	11.60	4.40	0.89	63	Y
3	Exponents and Logarithms	7	24.40	5.20	0.45	74	Y
4	Series, Combinatorics, Probability and Statistics	10	13.80	5.00	0.00	50	Y
	Total	29	63.80	18.80			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 63. Range-of-Knowledge for Geometry: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Geometry			
1	Logic and Geometric Proofs	7	21.71	6.43	0.53	92	Y
2	Volume and Area Formulas	4	11.29	3.43	0.79	86	Y
3	Angle Relationships, Constructions, and Lines	6	16.43	5.00	0.82	83	Y
4	Trigonometry	9	15.43	4.86	0.38	54	Y
	Total	26	64.86	19.72			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Table A- 64. Range-of-Knowledge for General Math: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				General Math			
1	Rational Numbers	6	14.86	5.43	0.79	91	Y
2	Exponents, Powers, and Roots	5	10.43	4.86	0.38	97	Y
3	Quantitative Relationships and Evaluating Expressions	6	10.14	4.43	0.79	74	Y
4	Multistep Problems, Graphing, and Functions	5	9.00	4.14	0.69	83	Y
5	Measurement and Geometry	9	12.00	7.00	0.58	78	Y
6	Statistics, Data Analysis, and Probability	8	8.57	5.57	0.79	70	Y
	Total	39	65.00	31.43			
	Percentage of standards with 50% of objectives linked to at least one item						100%

Balance-of-Knowledge Representation

The tables below present the mean balance index calculated per standard for Algebra I, Algebra II, Geometry, and General Math. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 65. Balance-of-Knowledge Representation for Algebra I: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					Algebra I			
1	Number Properties, Operations, and Linear Equations	11	6.25	18.13	27.88	66.94	5.85	N
2	Graphing and Systems of Linear Equations	4	3.88	13.13	20.19	80.89	3.98	Y
3	Quadratics and Polynomials	8	7.63	20.88	32.12	81.70	5.44	Y
4	Functions and Rational Expressions	6	5.00	12.00	18.46	76.42	2.43	Y
	Total	29	22.76	64.14				
	Percentage of standards with a balance of representation index of 70 or greater							80%

Table A- 66. Balance-of-Knowledge Representation for Algebra II: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Algebra II			
1 Polynomials and Rational Expressions	5	5.00	20.40	31.38	82.65	3.90	Y
2 Quadratics, Conics, and Complex Numbers	7	6.20	16.40	25.23	76.31	5.87	Y
3 Exponents and Logarithms	7	4.40	13.20	20.31	83.26	5.37	Y
4 Series, Combinatorics, Probability and Statistics	10	7.20	13.80	21.23	84.73	5.85	Y
Total	29	22.80	63.80				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 67. Balance-of-Knowledge Representation for Geometry: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Geometry			
1 Logic and Geometric Proofs	7	6.57	22.00	33.85	64.96	10.00	N
2 Volume and Area Formulas	4	3.43	11.43	17.58	81.09	11.98	Y
3 Angle Relationships, Constructions, and Lines	6	4.71	16.14	24.84	73.90	8.26	Y
4 Trigonometry	9	4.86	15.43	23.74	81.76	3.97	Y
Total	26	19.57	65.00				
Percentage of standards with a balance of representation index of 70 or greater							75%

Table A- 68. Balance-of-Knowledge Representation for General Math: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				General Math			
1 Rational Numbers	6	5.43	14.57	22.53	68.90	5.85	N
2 Exponents, Powers, and Roots	5	4.86	10.29	15.88	77.00	4.44	Y
3 Quantitative Relationships and Evaluating Expressions	6	4.57	10.14	15.68	80.09	7.08	Y
4 Multistep Problems, Graphing, and Functions	5	4.14	9.00	13.90	77.74	4.88	Y
5 Measurement and Geometry	9	7.14	11.57	17.89	77.46	4.87	Y
6 Statistics, Data Analysis, and Probability	8	5.57	8.57	13.25	78.75	1.98	Y
Total	39	31.71	64.14				
Percentage of standards with a balance of representation index of 70 or greater							83%

Mathematics: Integrated Math I, II, and III

Categorical Concurrence

The tables below present the results for the Integrated Math I, II, and III end-of-course tests on categorical concurrence per standard. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion.

Table A- 69. Categorical Concurrence for Integrated Math I: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		Integrated Math I			
1	Number Properties, Operations, and Linear Equations	15	13.71	1.70	Y
2	Graphing	9	9.00	0.00	Y
3	Quadratics and Polynomials	14	14.71	2.43	Y
4	Functions and Rational Expressions	7	6.57	1.13	Y
5	Geometry	20	20.00	1.53	Y
	Total	65	64.00		
	Percent of standards with at least six items				100%

Table A- 70. Categorical Concurrence for Integrated Math II: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		Integrated Math II			
1	Algebra I	20	18.57	1.51	Y
2	Logic and Geometric Proofs	22	20.71	2.69	Y
3	Angle Relationships, Constructions, and Lines	8	9.71	2.14	Y
4	Trigonometry	10	10.14	0.38	Y
5	Algebra II/Probability and Statistics	5	5.00	0.00	N
	Total	65	64.00		
	Percent of standards with at least six items				80%

Table A- 71. Categorical Concurrence for Integrated Math III: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Integrated Math III		
1	Geometry	1	4.86	0.38	N
2	Polynomials and Rational Expressions	23	18.71	2.06	Y
3	Quadratics, Conics, and Complex Numbers	16	15.86	1.43	Y
4	Exponents and Logarithms	16	13.43	1.40	Y
5	Series, Combinatorics, Probability and Statistics	9	9.14	2.54	Y
	Total	65	62.00		
	Percent of standards with at least six items				80%

Depth-of-Knowledge Consistency

The tables below present the results for depth-of-knowledge consistency. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A- 72. Depth-of-Knowledge Consistency for Integrated Math I: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Integrated Math I				
1	Number Properties, Operations, and Linear Equations	13.71	29	0.22	59	0.23	11	0.07	Y
2	Graphing	9.00	62	0.13	32	0.12	6	0.06	N
3	Quadratics and Polynomials	14.71	51	0.21	43	0.18	6	0.09	N
4	Functions and Rational Expressions	6.57	62	0.21	31	0.12	7	0.13	N
5	Geometry	20.00	38	0.19	45	0.13	17	0.12	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								40%

Table A- 73. Depth-of-Knowledge Consistency for Integrated Math II: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Integrated Math II				
1	Algebra I	18.57	57	0.20	42	0.20	1	0.04	N
2	Logic and Geometric Proofs	20.71	70	0.22	30	0.22	0	0	N
3	Angle Relationships, Constructions, and Lines	9.71	38	0.19	49	0.22	13	0.07	Y
4	Trigonometry	10.14	46	0.34	53	0.32	1	0.04	Y
5	Algebra II/Probability and Statistics	5.00	26	0.34	51	0.32	23	0.35	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								60%

Table A- 74. Depth-of-Knowledge Consistency for Integrated Math III: Mean Percent of Core Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			Integrated Math III						
1	Geometry	4.86	97	0.08	3	0.08	0	0	N
2	Polynomials and Rational Expressions	18.71	23	0.17	68	0.13	9	0.10	Y
3	Quadratics, Conics, and Complex Numbers	15.86	61	0.11	36	0.09	3	0.04	N
4	Exponents and Logarithms	13.43	60	0.13	35	0.12	5	0.07	N
5	Series, Combinatorics, Probability and Statistics	9.14	42	0.13	46	0.13	12	0.08	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								40%

Range-of-Knowledge Correspondence

The tables below present the results on the range of content covered by the test items for Integrated Math I, II, and III. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 75. Range-of-Knowledge for Integrated Math I: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Integrated Math			
1	Number Properties, Operations, and Linear Equations	7	13.71	3.86	0.69	55	Y
2	Graphing	3	9.00	3.00	0.00	100	Y
3	Quadratics and Polynomials	5	14.71	4.57	0.53	91	Y
4	Functions and Rational Expressions	2	6.57	1.86	0.38	19	N
5	Geometry	10	20.00	6.00	1.00	10	N
	Total	27	64.00	19.29			
	Percentage of standards with 50% of objectives linked to at least one item						60%

Table A- 76. Range-of-Knowledge for Integrated Math II: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Integrated Math			
1	Algebra I	13	18.57	9.14	0.69	70	Y
2	Logic and Geometric Proofs	6	20.71	5.00	1.00	83	Y
3	Angle Relationships, Constructions, and Lines	5	9.71	3.00	0.00	60	Y
4	Trigonometry	4	10.14	3.71	0.49	93	Y
5	Algebra II/Probability and Statistics	3	5.00	2.57	0.53	18	N
	Total	31	64.00	23.42			
	Percentage of standards with 50% of objectives linked to at least one item						80%

Table A- 77. Range-of-Knowledge for Integrated Math III: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Integrated Math			
1	Geometry	1	4.86	1.00	0.00	100	Y
2	Polynomials and Rational Expressions	5	18.71	4.86	0.38	97	Y
3	Quadratics, Conics, and Complex Numbers	7	15.86	5.71	0.95	14	N
4	Exponents and Logarithms	6	13.43	5.14	0.38	6	N
5	Series, Combinatorics, Probability and Statistics	7	9.14	4.57	0.98	14	N
	Total	26	62.00	21.28			
	Percentage of standards with 50% of objectives linked to at least one item						40%

Balance-of-Knowledge Representation

The tables below present the mean balance index calculated per standard for Integrated Math I, II, and III. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 78. Balance-of-Knowledge Representation for Integrated Math I: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Integrated Math I			
1 Number Properties, Operations, and Linear Equations	7	3.86	13.71	21.10	79.96	4.90	Y
2 Graphing	3	3.00	9.00	13.85	74.60	5.42	Y
3 Quadratics and Polynomials	5	4.57	14.71	22.64	76.59	9.89	Y
4 Functions and Rational Expressions	2	1.86	6.57	10.11	88.15	10.64	Y
5 Geometry	10	6.00	20.00	30.77	70.54	5.93	Y
Total	27	19.29	63.99				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 79. Balance-of-Knowledge Representation for Integrated Math II: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					Integrated Math II			
1	Algebra I	13	9.14	18.57	28.57	73.27	4.78	Y
2	Logic and Geometric Proofs	6	5.00	20.71	31.87	70.30	8.25	Y
3	Angle Relationships, Constructions, and Lines	5	3.00	9.71	14.95	83.57	9.64	Y
4	Trigonometry	4	3.71	10.14	15.60	84.85	6.22	Y
5	Algebra II/Probability and Statistics	3	2.57	5.00	7.69	78.57	8.79	Y
	Total	31	23.42	64.13				
	Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 80. Balance-of-Knowledge Representation for Integrated Math III: Mean Balance Index per Standard

	Title of Standard	Balance-of-Knowledge Representation						
		Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
			M	M	M	M	S.D.	
					Integrated Math III			
1	Geometry	1	1.00	4.86	7.47	100.00	0.00	Y
2	Polynomials and Rational Expressions	5	4.86	18.71	28.79	76.65	7.89	Y
3	Quadratics, Conics, and Complex Numbers	7	5.71	15.86	24.40	81.13	7.51	Y
4	Exponents and Logarithms	6	5.14	13.43	20.66	82.86	7.22	Y
5	Series, Combinatorics, Probability and Statistics	7	4.57	9.14	14.07	90.15	9.05	Y
	Total	26	21.28	62.00				
	Percentage of standards with a balance of representation index of 70 or greater							100%

Science: Grades 5, 8, and 10

Categorical Concurrence

The tables below present the results for Science 5, 8, and 10 on categorical concurrence for each standard separated by grade level. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row shows the percentage of standards that met the minimum alignment criterion.

Table A- 81. Categorical Concurrence for Science Grade 5: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
			Science Grade 5		
1	Physical Sciences	18	17.14	1.07	Y
2	Life Sciences	18	17.00	0.00	Y
3	Earth Sciences	18	17.86	0.69	Y
4	Investigation and Experimentation	6	8.00	1.63	Y
5	Physical Sciences	18	17.14	1.07	Y
	Total	60	60.00		
	Percent of standards with at least six items				100%

Table A- 82. Categorical Concurrence for Science Grade 8: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		Science Grade 8			
1	Motion	8	9.00	0.58	Y
2	Forces	8	8.14	0.69	Y
3	Structure of Matter	8	7.86	0.90	Y
4	Earth in the Solar System (Earth Science)	7	7.00	0.00	Y
5	Reactions	7	7.57	1.13	Y
6	Chemistry of Living Systems (Life Science)	3	2.86	0.38	N
7	Periodic Table	7	7.29	0.49	Y
8	Density and Buoyancy	5	6.43	1.13	Y
9	Investigation and Experimentation	6	3.57	1.27	N
	Total	60	59.72		
	Percent of standards with at least six items				78%

Table A- 83. Categorical Concurrence for Science Grade 10: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		Science Grade 10			
1	Cell Biology	10	9.00	2.38	Y
2	Genetics	12	12.71	1.89	Y
3	Ecology	11	11.29	2.69	Y
4	Evolution	11	10.57	1.62	Y
5	Physiology	10	11.44	0.98	Y
6	Investigation and Experimentation	6	4.72	2.56	N
	Total	60	59.73		
	Percent of standards with at least six items				84%

Depth-of-Knowledge Consistency

The tables below present the results from the comparison between the depth-of-knowledge expected in the standards compared to the depth-of-knowledge assessed by items for Science grades 5, 8, and 10. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A- 84. Depth-of-Knowledge Consistency for Science Grade 5: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Science Grade 5				
1	Physical Sciences	17.14	25	6.26	37	16.39	38	17.59	Y
2	Life Sciences	17.00	49	17.90	31	8.52	19	11.55	Y
3	Earth Sciences	17.86	28	12.64	46	12.77	26	24.41	Y
4	Investigation and Experimentation	8.00	35	9.80	36	13.57	29	10.44	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 85. Depth-of-Knowledge Consistency for Science Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Science Grade 8				
1	Motion	9.00	14	4.85	38	17.85	48	17.07	Y
2	Forces	8.14	25	3.18	54	16.75	21	17.83	Y
3	Structure of Matter	7.86	24	13.59	59	9.86	16	12.85	Y
4	Earth in the Solar System (Earth Science)	7.00	29	21.61	40	24.23	31	31.18	Y
5	Reactions	7.57	45	24.93	14	17.82	40	23.29	Y
6	Chemistry of Living Systems (Life Science)	2.86	14	16.50	49	28.79	36	30.93	Y
7	Periodic Table	7.29	8	10.16	35	17.03	57	23.45	Y
8	Density and Buoyancy	6.43	40	25.65	42	36.32	18	19.50	Y
9	Investigation and Experimentation	3.57	14	4.85	38	17.85	48	17.07	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table A- 86. Depth-of-Knowledge Consistency for Science Grade 10: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
					Science Grade 10				
1	Cell Biology	9.00	42	20.91	42	24.79	16	20.42	Y
2	Genetics	12.71	42	17.11	39	10.11	20	13.79	Y
3	Ecology	11.29	60	25.44	34	22.33	6	5.45	N
4	Evolution	10.57	31	17.63	48	13.02	22	27.37	Y
5	Physiology	11.44	60	28.79	28	15.70	12	17.33	N
6	Investigation and Experimentation	4.71	33	22.01	44	28.28	23	17.28	Y

Percent of standards with 50% of item DOK at or above objective DOK: 67%

Range-of-Knowledge Correspondence

The tables below present the results on the range of content covered by the test items for Science Grades 5, 8, and 10. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 87. Range-of-Knowledge for Science Grade 5: Mean Percent Objectives per Standard Linked with Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Science Grade 5			
1 Physical Sciences	16	17.14	14.00	1.53	88	Y
2 Life Sciences	14	17.00	11.43	1.13	82	Y
3 Earth Sciences	17	17.86	13.43	1.62	79	Y
4 Investigation and Experimentation	15	8.00	6.43	1.51	43	N
Total	62	60.00	45.29			
Percentage of standards with 50% of objectives linked to at least one item						75%

Table A- 88. Range-of-Knowledge for Science Grade 8: Mean Percent of Objectives per Standard Linked with Core Items

Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Science Grade 8			
1 Motion	6	9.00	5.57	0.53	93	Y
2 Forces	7	8.14	6.29	0.76	90	Y
3 Structure and Matter	6	7.86	5.14	0.69	86	Y
4 Earth in the Solar System (Earth Science)	5	7.00	5.00	0.00	100	Y
5 Reactions	5	7.57	5.00	0.00	100	Y
6 Chemistry of Living Systems (Life Science)	3	2.86	2.57	0.53	86	Y
7 Periodic Table	3	7.29	3.00	0.58	100	Y
8 Density and Buoyancy	4	6.43	3.71	0.49	93	Y
9 Investigation and Experimentation	7	3.57	2.71	1.25	39	N
Total	46	59.72	38.99			
Percentage of standards with 50% of objectives linked to at least one item						89%

Table A- 89. Range-of-Knowledge for Science Grade 10: Mean Percent of Objectives per Standard Linked with Core Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				Science Grade 10			
1	Cell Biology	8	9.00	7.00	1.15	88	Y
2	Genetics	10	12.71	8.29	1.25	83	Y
3	Ecology	9	11.29	7.43	0.79	83	Y
4	Evolution	10	10.57	7.43	0.98	74	Y
5	Physiology	8	11.44	7.14	0.38	89	Y
6	Investigation and Experimentation	9	4.72	3.67	0.82	41	N
	Total	54	59.73	40.96			
	Percentage of standards with 50% of objectives linked to at least one item						83%

Balance-of-Knowledge Representation

The tables below present the mean balance index calculated per standard for Science per grade level. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 90. Balance-of-Knowledge Representation for Science Grade 5: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Science Grade 5			
1 Physical Sciences	16	14.00	17.14	28.57	85.56	4.97	Y
2 Life Sciences	14	11.43	17.00	28.33	77.90	3.55	Y
3 Earth Sciences	17	13.43	17.86	29.76	81.74	2.21	Y
4 Investigation and Experimentation	15	6.43	8.00	13.33	90.26	7.51	Y
Total	62	45.29	60.00				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 91. Balance-of-Knowledge Representation for Science Grade 8: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Science Grade 8			
1 Motion	6	5.57	9.00	15.00	77.42	1.99	Y
2 Forces	7	6.29	8.14	13.57	85.11	6.96	Y
3 Structure and Matter	6	5.14	7.86	13.10	88.24	2.54	Y
4 Earth in the Solar System (Earth Science)	5	5.00	7.00	11.67	82.04	2.16	Y
5 Reactions	5	5.00	7.57	12.62	81.18	4.37	Y
6 Chemistry of Living Systems (Life Science)	3	2.57	2.86	4.76	95.24	8.13	Y
7 Periodic Table	3	3.00	7.29	12.14	81.46	6.73	Y
8 Density and Buoyancy	4	3.71	6.43	10.71	76.53	8.40	Y
9 Investigation and Experimentation	7	2.71	3.57	5.95	90.95	8.54	Y
Total	46	38.99	59.72				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 92. Balance-of-Knowledge Representation for Science Grade 10: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Science Grade 10			
1 Cell Biology	8	7.00	9.00	15.00	85.79	7.92	Y
2 Genetics	10	8.29	12.71	21.19	78.44	5.09	Y
3 Ecology	9	7.43	11.29	18.81	82.53	4.66	Y
4 Evolution	10	7.43	10.57	17.62	82.48	9.06	Y
5 Physiology	8	7.14	11.43	19.05	77.29	3.80	Y
6 Investigation and Experimentation	9	3.67	5.50	9.17	86.67	11.30	Y
Total	54	40.96	60.5				
Percentage of standards with a balance of representation index of 70 or greater							100%

History-Social Science: Grades 8, 10, and 11

Categorical Concurrence

The results for Grades 8, 10, and 11 History-Social Science on categorical concurrence are presented below. Results are separated by grade level. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion.

Items matched to the History and Social Science Analysis Skills (used as a secondary standard by all panelists) were included in the analyses with the same weight as the primary standard. As a result, the mean number of items matched often exceeds the target number of items.

Table A- 93. Categorical Concurrence for History and Social Science Grade 8: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		History-Social Science Grade 8			
1	World History and Geography: Ancient Civilizations	16	16.71	1.25	Y
2	Late Antiquity and the Middle Ages	14	13.57	0.53	Y
3	Renaissance/Reformation	10	9.57	0.79	Y
4	U.S. Constitution and the Early Republic	22	21.86	0.90	Y
5	Civil War and its Aftermath	13	13.14	0.90	Y
6	History and Social Science Analysis Skills	Embedded	10.00	6.02	Y
	Total	75	84.85		
	Percent of standards with at least six items				100%

Table A- 94. Categorical Concurrence for History and Social Science Grade 10: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		History-Social Science Grade 10			
1	Development of Modern Political Thought	13	12.86	0.38	Y
2	Industrial Expansion and Imperialism	10	10.00	0.82	Y
3	Causes and Effects of the First World War	14	13.86	1.21	Y
4	Causes and Effects of the Second World War	13	13.00	1.15	Y
5	International Developments in the Post-World War II Era	10	9.14	0.38	Y
6	History and Social Science Analysis Skills	Embedded	3.00	2.08	
	Total	60	61.86		
	Percent of standards with at least six items				100%

Table A- 95. Categorical Concurrence for History and Social Science Grade 11: Mean Number of Items per Standard

	Title of Standard	Number of Items Per Standard			At Least Six Items
		Target Number	Mean Number Matched	Standard Deviation	
		History-Social Science Grade 11			
1	Foundations of American Political and Social Thought	10	6.00	0.82	Y
2	Industrialization and the U.S. Role as a World Power	13	15.86	0.90	Y
3	United States: Between the World Wars	12	13.00	0.82	Y
4	World War II and Foreign Affairs	12	12.86	1.355	Y
5	Post-World War II Domestic Issues	13	12.29	1.80	Y
6	History and Social Science Analysis Skills	Embedded	3.14	3.67	
	Total	60	63.15		
	Percent of standards with at least six items				100%

Depth-of-Knowledge Consistency

The results for Grades 8, 10, and 11 History-Social Science on the comparison between the depth-of-knowledge expected in the standards and the depth-of-knowledge assessed by items. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards along with the corresponding standard deviations. Results are separated by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table A- 96. Depth-of-Knowledge Consistency for History and Social Science Grade 8: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

Title of Standard		Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			History-Social Science Grade						
1	World History and Geography: Ancient Civilizations	16.71	61	9.22	30	9.16	9	11.98	N
2	Late Antiquity and the Middle Ages	13.57	58	25.22	35	14.67	7	11.04	N
3	Renaissance/Reformation	9.57	44	34.49	44	21.35	12	18.04	Y
4	U.S. Constitution and the Early Republic	21.86	47	20.67	41	11.70	11	13.39	Y
5	Civil War and its Aftermath	13.14	49	30.31	41	25.62	9	16.28	Y
6	History and Social Science Analysis Skills	Embedded	54	34.45	43	33.80	3	8.17	N
Percent of standards with 50% of item DOK at or above objective DOK:									50%

Table A- 97. Depth-of-Knowledge Consistency for History and Social Science Grade 10: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			History-Social Science Grade						
1	Development of Modern Political Thought	12.86	21	13.74	56	14.77	23	12.44	Y
2	Industrial Expansion and Imperialism	10.00	49	25.16	46	21.45	4	5.41	Y
3	Causes and Effects of the First World War	13.86	65	11.07	35	11.07	0	0.00	N
4	Causes and Effects of the Second World War	13.00	64	15.27	33	16.26	2	3.91	N
5	International Developments in the Post-World War II Era	9.14	78	11.60	22	11.60	0	0.00	N
6	History and Social Science Analysis Skills	Embedded	81	20.30	18	20.30	0	0.00	N
	Percent of standards with 50% of item DOK at or above objective DOK:								33%

Table A- 98. Depth-of-Knowledge Consistency for History and Social Science Grade 11: Mean Percent of Items with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Items per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Items At or Above)
			% Items Below		% Items At Same Level		% Items Above		
			M	S.D.	M	S.D.	M	S.D.	
			History-Social Science Grade						
1	Foundations of American Political and Social Thought	6.00	44	25.94	44	20.99	12	12.77	Y
2	Industrialization and the U.S. Role as a World Power	15.86	37	14.40	59	12.07	4	5.48	Y
3	United States: Between the World Wars	13.00	68	19.38	31	20.22	1	2.70	N
4	World War II and Foreign Affairs	12.86	67	20.18	29	18.51	4	7.63	N
5	Post-World War II Domestic Issues	12.29	51	24.29	44	19.69	5	7.06	N
6	History and Social Science Analysis Skills	Embedded	91	11.03	9	11.03	0	0.00	N
	Percent of standards with 50% of item DOK at or above objective DOK:								33%

Range-of-Knowledge Correspondence

The results for Grades 8, 10, and 11 History-Social Science on Range-of-Knowledge are presented below. Results are separated by grade level. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table A- 99. Range-of-Knowledge for History and Social Science Grade 8: Mean Percent Objectives per Standard Linked with Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				History-Social Science Grade 8			
1	World History and Geography: Ancient Civilizations	48	16.71	15.29	1.11	32	N
2	Late Antiquity and the Middle Ages	40	13.57	11.57	0.53	29	N
3	Renaissance/Reformation	21	9.57	8.71	0.95	41	N
4	U.S. Constitution and the Early Republic	42	21.86	17.71	0.76	42	N
5	Civil War and its Aftermath	27	13.14	10.29	1.11	38	N
6	History and Social Science Analysis Skills	Embedded	10.00	2.33	0.81	33*	N
	Total	178	84.85	63.57			
	Percentage of standards with 50% of objectives linked to at least one item						0%

* Constructed from an estimate based on a target of seven, or half, of the content objectives for this standard.

Table A- 100. Range-of-Knowledge for History and Social Science Grade 10: Mean Percent of Objectives per Standard Linked with Core Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				History-Social Science Grade 10			
1	Development of Modern Political Thought	8	12.86	6.00	0.00	75	Y
2	Industrial Expansion and Imperialism	11	10.00	7.43	1.13	68	Y
3	Causes and Effects of the First World War	9	13.86	7.14	0.90	79	Y
4	Causes and Effects of the Second World War	9	13.00	7.71	1.11	86	Y
5	International Developments in the Post-World War II Era	11	9.14	5.71	0.49	52	Y
6	History and Social Science Analysis Skills	Embedded	3.00	1.83	0.75	26	N
	Total	48	61.86	35.82			
	Percentage of standards with 50% of objectives linked to at least one item						83%

* Constructed from an estimate based on a target of seven, or half, of the content objectives for this standard.

Table A-101. Range-of-Knowledge for History and Social Science Grade 11: Mean Percent of Objectives per Standard Linked with Core Items

	Title of Standard	Number of Objectives	Mean Items per Standard	Range of Objectives: Objectives with At Least One Item		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
				M	S.D.	M	
				History-Social Science Grade 11			
1	Foundations of American Political and Social Thought	9	6.00	4.86	1.07	54	Y
2	Industrialization and the U.S. Role as a World Power	15	15.86	9.57	0.53	64	Y
3	United States: Between the World Wars	12	13.00	7.86	0.69	66	Y
4	World War II and Foreign Affairs	15	12.86	8.29	0.76	55	Y
5	Post-World War II Domestic Issues	22	12.29	9.57	1.13	44	N
6	History and Social Science Analysis Skills	Embedded	3.14	2.00	1.60	29	N
	Total	73	63.15	42.15			
	Percentage of standards with 50% of objectives linked to at least one item						67%

* Constructed from an estimate based on a target of seven, or half, of the content objectives for this standard.

Balance-of-Knowledge Representation

The results for Grades 8, 10, and 11 History-Social Science on Balance-of-Knowledge Representation are presented below. Results are separated by grade level. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table A- 102. Balance-of-Knowledge Representation for History and Social Science Grade 8: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
	History-Social Science Grade 8						
1 World History and Geography: Ancient Civilizations	48	15.29	16.71	20.86	92.39	2.35	Y
2 Late Antiquity and the Middle Ages	40	11.57	13.57	16.93	88.07	2.52	Y
3 Renaissance/Reformation	21	8.71	9.57	11.98	92.00	3.59	Y
4 U.S. Constitution and the Early Republic	42	17.71	21.86	27.35	85.36	2.85	Y
5 Civil War and its Aftermath	27	10.29	13.14	16.36	84.76	2.11	Y
6 History and Social Science Analysis Skills	Embedded	2.00	9.50	11.11	75.28	17.22	Y
Total	178	65.57	84.35				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 103. Balance-of-Knowledge Representation for History and Social Science Grade 10: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
	History-Social Science Grade 10						
1 Development of Modern Political Thought	8	6.00	12.86	21.33	76.47	6.99	Y
2 Industrial Expansion and Imperialism	11	7.43	10.00	16.61	84.11	4.07	Y
3 Causes and Effects of the First World War	9	7.14	13.86	22.99	78.06	4.00	Y
4 Causes and Effects of the Second World War	9	7.71	13.00	21.58	77.71	4.48	Y
5 International Developments in the Post-World War II Era	11	5.71	9.14	15.18	80.00	3.80	Y
6 History and Social Science Analysis Skills	Embedded	1.67	3.00	4.84	88.89	9.62	Y
Total	48	35.66	61.86				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table A- 104. Balance-of-Knowledge Representation for History and Social Science Grade 11: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Items	Mean Items per Standard	Mean % of Items Linked to Standard (out of total items)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
	History-Social Science Grade 11						
1 Foundations of American Political and Social Thought	9	4.86	6.00	9.84	84.93	5.22	Y
2 Industrialization and the U.S. Role as a World Power	15	9.57	15.86	26.01	75.79	1.03	Y
3 United States: Between the World Wars	12	7.86	13.00	21.33	72.71	2.46	Y
4 World War II and Foreign Affairs	15	8.29	12.86	21.07	74.49	4.94	Y
5 Post-World War II Domestic Issues	22	9.57	12.29	20.18	83.75	3.19	Y
6 History and Social Science Analysis Skills	Embedded	1.50	3.50	5.46	85.00	21.21	Y
Total	73	41.65	63.51				
Percentage of standards with a balance of representation index of 70 or greater							100%

Appendix B:
Alignment Results on Webb Criteria by Content Area and Grade Level for the CAPA

CAPA – English-Language Arts

Categorical Concurrence

The results for Levels I through V of the CAPA ELA assessment for categorical concurrence are presented below. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion. In comparison to the categorical concurrence tables for the CSTs, the tables for the CAPA do not include a column listing the Target Number of Tasks. This is because we only reviewed the new field-test items.

Table B- 1. Categorical Concurrence for ELA CAPA Level I: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
	ELA CAPA Level I			
1	Reading	14.50	1.73	Y
2	Writing	4.00	0.00	Y
3	Listening and Speaking	5.24	1.5	Y
	Total	23.75		
	Percent of standards with at least one task			100%

Table B- 2. Categorical Concurrence for ELA CAPA Level II: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
	ELA CAPA Level II			
1	Reading	7.50	0.58	Y
2	Writing	4.00	0.00	Y
3	Listening and Speaking	4.50	0.58	Y
	Total	16		
	Percent of standards with at least one task			100%

Table B- 3. Categorical Concurrence for ELA CAPA Level III: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
	ELA CAPA Level III			
1	Reading	10.50	0.58	Y
2	Writing	5.25	0.96	Y
	Total	15.75		
	Percent of standards with at least one task			100%

Table B- 4. Categorical Concurrence for ELA CAPA Level IV: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
	ELA CAPA Level IV			
1	Reading	15.00	0.82	Y
2	Writing	6.75	1.50	Y
3	Listening and Speaking	4.50	0.58	Y
	Total	26.25		
	Percent of standards with at least one task			100%

Table B- 5. Categorical Concurrence for ELA CAPA Level V: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
	ELA CAPA Level V			
1	Reading	11.75	2.63	Y
2	Writing	7.00	1.41	Y
3	Listening and Speaking	5.25	1.71	Y
	Total	24.00		
	Percent of standards with at least one task			100%

Depth-of-Knowledge Consistency

The results for Levels I through V of the CAPA ELA assessment for Depth-of-Knowledge consistency are presented below. The tables present the results from the comparison between the depth-of-knowledge expected in the standards and the depth-of-knowledge assessed by items. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards along with the corresponding standard deviations. Results are separated by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table B- 6. Depth-of-Knowledge Consistency for ELA CAPA Level I: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA CAPA Level I						
1	Reading	14.50	69	10.47	30	11.65	2	3.57	N
2	Writing	4.00	6	12.50	88	14.43	6	12.50	Y
3	Listening and Speaking	5.24	50	57.74	50	57.74	0	0.00	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								67%

Table B- 7. Depth-of-Knowledge Consistency for ELA CAPA Level II: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA CAPA Level II						
1	Reading	7.50	10	6.60	58	22.13	33	20.84	Y
2	Writing	4.00	0	0.00	50	20.41	50	20.41	Y
3	Listening and Speaking	4.50	5	10.00	25	50.00	70	47.60	Y*
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

* Note that 70% of tasks were rated as assessing student knowledge above the level expected in the standards.

Table B- 8. Depth-of-Knowledge Consistency for ELA CAPA Level III: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA CAPA Level III						
1	Reading	10.50	41	13.81	45	12.00	14	5.57	Y
2	Writing	5.25	6	12.50	66	12.29	28	6.58	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table B- 9. Depth-of-Knowledge Consistency for ELA CAPA Level IV: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA CAPA Level IV						
1	Reading	15.00	27	20.51	53	15.23	20	8.24	Y
2	Writing	6.75	9	11.97	33	13.44	58	15.00	Y*
3	Listening and Speaking	4.50	28	32.02	50	24.49	23	20.62	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

* Note that over half of the tasks were rated as assessing student knowledge above the level expected in the standards.

Table B- 10. Depth-of-Knowledge Consistency for ELA CAPA Level V: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			ELA CAPA Level V						
1	Reading	11.75	28	10.87	60	11.77	12	10.61	Y
2	Writing	7.00	11	7.37	31	19.15	59	23.01	Y*
3	Listening and Speaking	5.25	12	15.79	50	19.27	38	27.78	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

* Note that over half of the tasks were rated as assessing student knowledge above the level expected in the standards.

Range-of-Knowledge Correspondence

The results for Levels I through V of the CAPA ELA assessment for Range-of-Knowledge correspondence are presented below. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table B- 11. Range-of-Knowledge for ELA CAPA Level I: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			ELA CAPA Level I			
1 Reading	6	14.50	4.75	1.50	79	Y
2 Writing	4	4.00	2.00	0.00	50	Y
3 Listening and Speaking	8	5.24	1.25	0.50	16	N
Total	18	23.75	8.00			
Percentage of standards with 50% of objectives linked to at least one item						67%

Table B- 12. Range-of-Knowledge for ELA CAPA Level II: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			ELA CAPA Level II			
1 Reading	8	7.50	5.25	0.50	65	Y
2 Writing	6	4.00	4.00	0.00	67	Y
3 Listening and Speaking	4	4.50	1.50	0.58	25	N
Total	18	16	10.75			
Percentage of standards with 50% of objectives linked to at least one item						67%

Table B- 13. Range-of-Knowledge for ELA CAPA Level III: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			ELA CAPA Level III			
1 Reading	10	10.50	7.51	0.58	75	Y
2 Writing	6	5.25	3.75	0.50	63	Y
Total	16	15.75	11.26			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 14. Range-of-Knowledge for ELA CAPA Level IV: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			ELA CAPA Level IV			
1 Reading	11	15.00	8	0.82	73	Y
2 Writing	4	6.75	3.5	0.58	88	Y
3 Listening and Speaking	5	4.50	3.5	0.58	70	Y
Total	20	26.25	15			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 15. Range-of-Knowledge for ELA CAPA Level V: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			ELA CAPA Level V			
1 Reading	7	11.75	5.25	0.96	75	Y
2 Writing	3	7.00	3.00	0.00	100	Y
3 Listening and Speaking	3	5.25	2.75	0.50	92	Y
Total	13	24.00	11.00			
Percentage of standards with 50% of objectives linked to at least one item						100%

Balance-of-Knowledge Representation

The results for Levels I through V of the CAPA ELA assessment for Balance-of-Knowledge representation are presented below. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table B- 16. Balance-of-Knowledge Representation for ELA CAPA Level I: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA CAPA Level I			
1 Reading	6	4.75	14.5	61	72	7.41	Y
2 Writing	4	2.00	4.00	17	94	12.50	Y
3 Listening and Speaking	8	1.25	5.24	22	100	0.00	Y
Total	18	8.00	23.75				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 17. Balance-of-Knowledge Representation for ELA CAPA Level II: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA CAPA Level II			
1 Reading	8	5.25	7.50	47	79	2.61	Y
2 Writing	6	4.00	4.00	25	100	0.00	Y
3 Listening and Speaking	4	1.50	4.50	26	85	17.32	Y
Total	18	10.75	16.00				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 18. Balance-of-Knowledge Representation for ELA CAPA Level III: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA CAPA Level III			
1 Reading	10	7.51	10.50	67	84	1.07	Y
2 Writing	6	3.75	5.25	33	80	5.34	Y
Total	16	11.26	15.75				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 19. Balance-of-Knowledge Representation for ELA CAPA Level IV: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA CAPA Level IV			
1 Reading	11	8	15.00	57	80	2.22	Y
2 Writing	4	3.5	6.75	26	90	6.40	Y
3 Listening and Speaking	5	3.5	4.50	17	84	.96	Y
Total	20	15	26.25				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 20. Balance-of-Knowledge Representation for ELA CAPA Level V: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				ELA CAPA Level V			
1 Reading	7	5.25	11.75	49	83	3.39	Y
2 Writing	3	3.00	7.00	29	87	3.72	Y
3 Listening and Speaking	3	2.75	5.25	22	87	9.98	Y
Total	13	11.00	24.00				
Percentage of standards with a balance of representation index of 70 or greater							100%

CAPA – Mathematics

Categorical Concurrence

The results for Levels I through V of the CAPA mathematics assessment for categorical concurrence are presented below. Each table includes: the target number of items from the test blueprint; the mean number of items matched by panelists; the standard deviation among panelists' ratings; and, the final alignment conclusion (Yes or No). The bottom row indicates the percentage of standards that met the minimum alignment criterion. In comparison to the categorical concurrence tables for the CSTs, the tables for the CAPA do not include a column listing the Target Number of Tasks. This is because we only reviewed the new field-test items.

Table B- 21. Categorical Concurrence for Math CAPA Level I: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
		Math CAPA Level I		
1	Number Sense	9.00	0.82	Y
2	Algebra and Functions	3.75	0.96	Y
3	Measurement and Geometry	8.00	0.82	Y
4	Statistics, Data Analysis, and Probability	3.00	0.82	Y
	Total	23.75		
	Percent of standards with at least one task			100%

Table B- 22. Categorical Concurrence for Math CAPA Level II: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
		Math CAPA Level II		
1	Number Sense	7.25	0.50	Y
2	Algebra and Functions	3.00	0.82	Y
3	Measurement and Geometry	4.00	0.00	Y
4	Statistics, Data Analysis, and Probability	1.75	0.96	Y
	Total	16.00		
	Percent of standards with at least one task			100%

Table B- 23. Categorical Concurrence for Math CAPA Level III: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
		Math CAPA Level III		
1	Number Sense	5.00	0.00	Y
2	Algebra and Functions	2.50	1.29	Y
3	Measurement and Geometry	4.00	0.00	Y
4	Statistics, Data Analysis, and Probability	4.5	1.29	Y
	Total	16.00		
	Percent of standards with at least one task			100%

Table B- 24. Categorical Concurrence for Math CAPA Level IV: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
		Math CAPA Level IV		
1	Number Sense	14.25	0.50	Y
2	Algebra and Functions	8.25	0.96	Y
3	Measurement and Geometry	4.00	0.00	Y
	Total	26.50		
	Percent of standards with at least one task			100%

Table B- 25. Categorical Concurrence for Math CAPA Level V: Mean Number of Performance Tasks per Standard

	Title of Standard	Number of Tasks per Standard		At Least One Task per Standard
		Mean Tasks Matched	Standard Deviation	
		Math CAPA Level V		
1	Number Sense	19.25	0.96	Y
3	Measurement and Geometry	3.75	0.50	Y
	Total	23.00		
	Percent of standards with at least one task			100%

Depth-of-Knowledge Consistency

The results for Levels I through V of the CAPA mathematics assessment for Depth-of-Knowledge consistency are presented below. The tables present the results from the comparison between the depth-of-knowledge expected in the standards and the depth-of-knowledge assessed by items. The tables include the mean percentage of items rated as below, at the same level, or above the DOK level of the content standards along with the corresponding standard deviations. Results are separated by grade level. Standards with at least 50% of items at the same (or above) DOK level met the minimum criterion.

Table B- 26. Depth-of-Knowledge Consistency for Math CAPA Level I: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

Title of Standard		Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks At Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math CAPA Level I						
1	Number Sense	9.00	32	16.44	55	16.11	14	10.02	Y
2	Algebra and Functions	3.75	30	24.73	70	24.73	0	0.00	Y
3	Measurement and Geometry	8.00	19	21.65	81	21.65	0	0.00	Y
4	Statistics, Data Analysis, and Probability	3.00	1	0.00	0	0.00	0	0.00	N
	Percent of standards with 50% of item DOK at or above objective DOK:								75%

Table B- 27. Depth-of-Knowledge Consistency for Math CAPA Level II: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

Title of Standard		Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks At Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math CAPA Level II						
1	Number Sense	7.25	0	0.00	76	16.32	24	16.32	Y
2	Algebra and Functions	3.00	29	34.36	71	34.36	0	0.00	Y
3	Measurement and Geometry	4.00	31	12.5	56	23.93	13	14.43	Y
4	Statistics, Data Analysis, and Probability	1.75	21	25.00	79	25.00	0	0.00	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table B- 28. Depth-of-Knowledge Consistency for Math CAPA Level III: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks At Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math CAPA Level III						
1	Number Sense	5.00	6	12.50	75	35.36	19	23.94	Y
2	Algebra and Functions	2.50	13	25.00	88	25.00	0	0.00	Y
3	Measurement and Geometry	4.00	0	0.00	81	37.50	19	37.50	Y
4	Statistics, Data Analysis, and Probability	4.50	13	16.33	66	26.89	20	16.69	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table B- 29. Depth-of-Knowledge Consistency for Math CAPA Level IV: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks At Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math CAPA Level IV						
1	Number Sense	14.25	7	14.29	77	23.00	16	9.06	Y
2	Algebra and Functions	8.25	20	14.23	53	21.48	27	20.66	Y
3	Measurement and Geometry	4.00	0	0.00	88	25.00	13	25.00	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Table B- 30. Depth-of-Knowledge Consistency for Math CAPA Level V: Mean Percent of Performance Tasks with DOK Below, At, and Above DOK Level of Objectives

	Title of Standard	Mean Tasks per Standard	Depth-of-Knowledge Consistency						DOK Consistency (min 50% of Tasks At or Above)
			% Tasks Below		% Tasks At Same Level		% Tasks Above		
			M	S.D.	M	S.D.	M	S.D.	
			Math CAPA Level V						
1	Number Sense	7	14	15.43	59	31.38	27	16.64	Y
3	Measurement and Geometry	1	13	14.43	44	42.70	44	42.70	Y
	Percent of standards with 50% of item DOK at or above objective DOK:								100%

Range-of-Knowledge Correspondence

The results for Levels I through V of the CAPA mathematics assessment for Range-of-Knowledge correspondence are presented below. The tables include the mean number and percentage of objectives by standard. For acceptable range-of-knowledge correspondence, a minimum of 50% of content objectives within each standard should be matched to at least one item.

Table B- 31. Range-of-Knowledge for Math CAPA Level I: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math CAPA Level I			
1 Number Sense	7	9.00	6.00	0.82	86	Y
2 Algebra and Functions	5	3.75	3.25	0.50	65	Y
3 Measurement and Geometry	7	8.00	4.00	0.00	57	Y
4 Statistics, Data Analysis, and Probability	1	3.00	1.00	0.00	100	Y
Total	20	23.75	14.25			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 32. Range-of-Knowledge for Math CAPA Level II: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math CAPA Level II			
1 Number Sense	14	7.25	5.00	0.00	36	N
2 Algebra and Functions	4	3.00	2.50	0.58	63	Y
3 Measurement and Geometry	6	4.00	3.00	0.00	50	Y
4 Statistics, Data Analysis, and Probability	2	1.75	1.00	0.00	50	Y
Total	26	16.00	11.50			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 33. Range-of-Knowledge for Math CAPA Level III: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math CAPA Level III			
1 Number Sense	11	5.00	5.00	0.00	45	N
2 Algebra and Functions	2	2.50	1.75	0.50	88	Y
3 Measurement and Geometry	6	4.00	4.00	0.00	67	Y
4 Statistics, Data Analysis, and Probability	6	4.50	3.75	0.50	63	Y
Total	14	16.00	14.50			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 34. Range-of-Knowledge for Math CAPA Level IV: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math CAPA Level IV			
1 Number Sense	6	14.25	5.75	0.50	96	Y
2 Algebra and Functions	2	8.25	2.00	0.00	100	Y
3 Measurement and Geometry	1	4.00	1.00	0.00	100	Y
Total	9	26.50	8.75			
Percentage of standards with 50% of objectives linked to at least one item						100%

Table B- 35. Range-of-Knowledge for Math CAPA Level V: Mean Percent of Objectives per Standard Linked with Performance Tasks

Title of Standard	Number of Objectives	Mean Tasks per Standard	Range of Objectives: Objectives with At Least One Task		Range of Objectives: % of Total Objectives per Standard	Range-of-Knowledge Correspondence
			M	S.D.	M	
			Math CAPA Level V			
1 Number Sense	8	19.25	6.75	0.96	84	Y
3 Measurement and Geometry	1	3.75	1.00	0.00	100	Y
Total	9	23.00	7.75			
Percentage of standards with 50% of objectives linked to at least one item						100%

Balance-of-Knowledge Representation

The results for Levels I through V of the CAPA mathematics assessment for Balance-of-Knowledge representation are presented below. The tables also include the percentage of items linked to each standard. The minimum acceptable balance index is a 70 out of 100.

Table B- 36. Balance of Representation for Math CAPA Level I: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math CAPA Level I			
1 Number Sense	7	6.00	9.00	38	83	1.84	Y
2 Algebra and Functions	5	3.25	3.75	16	97	6.67	Y
3 Measurement and Geometry	7	4.00	8.00	34	77	2.76	Y
4 Statistics, Data Analysis, and Probability	1	1.00	3.00	13	100	0.00	Y
Total	20	14.25	23.75				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 37. Balance of Representation for Math CAPA Level II: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math CAPA Level II			
1 Number Sense	14	5.0	7.25	45	83	1.07	
2 Algebra and Functions	4	2.50	3.00	19	92	9.62	
3 Measurement and Geometry	6	3.00	4.00	25	83	0.00	
4 Statistics, Data Analysis, and Probability	2	1.00	1.75	11	100	0.00	
Total	26	11.50	16.00				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 38. Balance of Representation for Math CAPA Level III: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math CAPA Level III			
1 Number Sense	11	5.00	5.00	31	100	0.00	Y
2 Algebra and Functions	2	1.75	2.50	16	90	12.5	Y
3 Measurement and Geometry	6	4.00	4.00	25	100	0.00	Y
4 Statistics, Data Analysis, and Probability	6	3.75	4.50	28	90	12.25	Y
Total	14	14.5	16				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 39. Balance of Representation for Math CAPA Level IV: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math CAPA Level IV			
1 Number Sense	6	5.75	14.25	54	81	1.96	Y
2 Algebra and Functions	2	2.00	8.25	31	92	3.29	Y
3 Measurement and Geometry	1	1.00	4.00	15	100	0.00	Y
Total	9	8.75	26.50				
Percentage of standards with a balance of representation index of 70 or greater							100%

Table B- 40. Balance of Representation for Math CAPA Level V: Mean Balance Index per Standard

Title of Standard	Balance-of-Knowledge Representation						
	Objectives per Standard	Mean Obj. Linked with Tasks	Mean Tasks per Standard	Mean % of Tasks Linked to Standard (out of total tasks)	Mean Balance Index		Acceptability of Balance Index (70 or above)
		M	M	M	M	S.D.	
				Math CAPA Level V			
1 Number Sense	8	6.75	19.25	84	84	8.25	Y
3 Measurement and Geometry	1	1.00	3.75	16	100	0.00	Y
Total	9	7.75	23.00				
Percentage of standards with a balance of representation index of 70 or greater							100%

Appendix C:
Panelist Comments on Assessment Items

CST Panelist Comments on Individual Items

The comments below reflect panelists' perspectives on the assessment items for the CSTs. Due to the need to maintain security, individual item identifiers are not presented. However, these comments have been given to the test developer for review and consideration.

Table C- 1. ELA Grades 2 to 8: Summary of Panelist Comments on Items

Grade*	Type of Comment	Percent of panelists	Number of panelists with comment
2	No standard matches this item.	100%	5
	Item addresses a standard from a lower grade. (1st grade standard)	100%	5
3	Item is worded confusingly.	50%	3
	Item does not directly address the standard (dividing a word into syllables is not the same as decoding.)	50%	3
4	No fourth grade standard addresses identifying nouns.	100%	6
	Item requires basic recall and matches no standard.	83%	5
6	Must stretch definitions to match this item to a standard (include dialogue as a literary device.)	40%	2
7	No good match between item and any 7th grade standard	50%	2
8	Item explanation uses terminology that causes it to not align with standard. (should say 'directions' rather than 'advertisement'.)	33%	2
	"Rural" is not a time period, but a setting.	33%	2

* No repeated comments in grade 5 analysis.

Table C- 2. Math Grades 2 to 7: Summary of Panelist Comments on Items

Grade	Type of Comment	Percent of panelists	Number of panelists with comment
2	The question does not give an array in which students can work.	50%	2
	Graphic (of a clock face) should be added to this item.	50%	2
3	Item requires student to read and write numbers in word form, which is above their grade level.	60%	3
	Distractors are not well done. (too easy to find answer with estimation, make the item more confusing, etc.)	60%	3
	Item assesses a math fact, measured by standard 2.2 which is listed as N/A in our code column.	60%	3
	Item assesses something for which there is no CA standard.	80%	4
4	There is no standard for multiplying or dividing decimals/money.	75%	3
	Item goes beyond what is stated in the standard.	75%	3
5	The item is worded in a confusing way, a picture or bullets could be used instead.	50%	2
	The value of pi is a recall item; the skill being tested (rounding) is not necessary to get the correct answer.	50%	2
6	The item would be easier to understand with a picture, table, or graph.	50%	2
	The item is too wordy, or wording is confusing.	75%	3
7	Item requires student to have knowledge above their grade level	100%	4
	There is no standard that directly relates to this item.	100%	4

Table C- 3. Math End-of-Course Tests: Summary of Panelist Comments on Items

Course*	Type of Comment	Percent of panelists	Number of panelists with comment
Alg I	Item is poorly worded or confusing.	25%	2
	Item assumes prior knowledge.	25%	2
	Item does not align to standard.	63%	5
Alg II	Item does not align with any Alg II standard. (or aligns best with an Alg I or Geometry standard.)	100%	4
	Item is poorly worded, confusing, or vague.	100%	4
Geometry	Item wording is confusing.	43%	3
General Math	Item wording is confusing, or item is too wordy..	43%	3

Table C- 4. Math Integrated Math Tests I, II, and III: Summary of Panelist Comments on Items

Course*	Type of Comment	Percent of panelists	Number of panelists with comment
Int Math I	Item does not align with any Int Math I standard	43%	3
Int Math II	Item addresses a math standard not included in Int Math II.	43%	3
	Shading needed to make a selection is missing from the answer choices.	43%	3
Int Math III	Item does not align with any standard, or does not align well with any Int Math III standard..	43%	3

Table C- 5. Science Grades 5, 8 and 10: Summary of Panelist Comments on Items

Grade*	Type of Comment	Percent of panelists	Number of panelists with comment
5	Item assumes prior knowledge beyond what is covered in the standards.	29%	2
	Item is unclearly worded.	29%	2
8	Item is worded unclearly or is too wordy.	43%	3
10	Question assumes knowledge beyond what is in the standards.	29%	2

Table C- 6. History-Social Science Grades 8, 10, and 11: Summary of Panelist Comments on Items

Grade*	Type of Comment	Percent of panelists	Number of panelists with comment
8	Poor placement of item within the test, disrupts conceptual flow of the test..	86%	6
10	Item quotes person who is not widely known.	29%	2
	Spelling of word changes between the standard and the item.	43%	3

* No repeated comments in grade 11 analysis.

CAPA Panelist Comments on Individual Items

The comments below reflect panelists' perspectives on the assessment items for the CAPA. Due to the need to maintain security, individual item identifiers are not presented. However, these comments have been given to the test developer for review and consideration.

Table C- 7. CAPA ELA Levels I to V: Summary of Panelist Comments on Items

CAPA Level*	Type of Comment	Percent of panelists	Number of panelists with comment
I	Item requires a physical response, which some students would not be able to perform, or item needs to be adapted for hearing/vision/orally impaired students.	100%	4
	Requires prior experience many special ed students do not have (i.e. solid food for tube fed students).	50%	2
II	Requires prior experience many special ed students do not have (i.e. loud/quiet for hearing impaired students).	50%	2

In levels III, IV, and V comments similar to those in the table were made, but not by more than one reviewer per level.

Table C- 8. CAPA Math Levels I to V: Summary of Panelist Comments on Items

CAPA Level*	Type of Comment	Percent of panelists	Number of panelists with comment
I	The item is worded in a confusing way.	50%	2
III	The graph used is confusing (shading is poor, another type of graph would have been more clear).	100%	4
IV	The item is too complex, above many special ed students' developmental level.	75%	3
	The item is worded in a confusing way.	50%	2

In levels II and IV comments similar to those included in the table were made, but not by more than one reviewer per level.

Appendix D
Workshop Materials for Panelists

Example of Alignment Instructions for CST Review

Item Alignment Tasks for English-Language Arts

For the alignment review, we would like to ask you to perform several tasks to evaluate the test items compared to the California content standards. These tasks will be performed individually. You will be entering your ratings and responses in the Excel spreadsheet setup on the laptop provided. Please ask the HumRRO staff if you have a question.

Step 1: Rate the depth-of-knowledge of the California Content Standards for ELA.

There are several files on your laptop that include the ELA California content standards. Open the file labeled CA_ELA Standards_Grade… for the grade level that you will be working on currently. After you open the file, please enter your name and ID number under the appropriate columns. Your name will be deleted from the file once all of the data are merged. Enter the Grade Level that you will be working on currently.

Using the attached depth-of-knowledge (DOK) descriptions and examples, rate each content objective per standard on the degree of cognitive processing expected of students to demonstrate proficiency. In the CA_ELA Standards_Grade file, enter the DOK level (number) in the spreadsheet under the column labeled DOK Rating. Remember that *cognitive complexity* is related to *difficulty*, but these terms are not synonymous. If you find that a single content objective really requires several different tasks of varying complexity (i.e., “Students should be able to identify, distinguish, and explain…”), indicate the *highest* DOK level required by this content objective.

When you have finished rating the DOK level of the current grade level, you may move onto the next grade level. However, the HumRRO group leader will be stopping to review the grade level standards together after everyone has completed this grade. At this time, there will be group discussion to achieve majority agreement on the ratings.

Repeat the step above for EACH grade level of the ELA Content Standards.

Step 2: Rate the depth-of-knowledge (DOK) level of the item.

Next, open the spreadsheet labeled CA_ELA_ItemRatings_Grade… file for the current grade level assessment. After you open the file, please enter your name and ID number under the appropriate columns. Your name will be deleted from the file once all of the data are merged. Enter the Grade Level that you will be working on currently.

Using the attached depth-of-knowledge (DOK) descriptions and examples, rate each item on the degree of cognitive processing required of students to answer the item adequately. Enter the DOK level (number) in the spreadsheet under the column labeled Item DOK Rating. Remember that *cognitive complexity* is related to *difficulty*, but these terms are not synonymous. If you find that a single item really requires several different tasks of varying complexity (i.e., “Students should be able to identify, distinguish, and explain…”), indicate the *highest* DOK level required by this content objective.

Step 3: Match the item to a specific content objective.

Find the content standard that you think best reflects the content that the item is supposed to assess. Within the content standard, identify the specific content objective that the item targets. In the CA ELA Ratings file, enter the code for this content standard and objective (located on the left-hand column of your Coded Content Standards printout sheet next to each standard) into the Excel spreadsheet under the column labeled Content Standard/Objective 1.

If you find that an item assesses two or more content standards or objectives equally, you may include the additional standard and objective in the column labeled Content Standard/Objective 2. Please only enter a secondary standard if the item assesses this standard at an equal level to the first standard you chose.

Step 4: Rate the overall match level of the item to the standard and objective you chose.

Indicate how well you think that the item actually assesses the standard you selected. Please use the rating scale below to make your judgment. Enter the appropriate rating number from the scale into your spreadsheet.

1 Not aligned to any California content standard (Use ONLY if you did not assign a standard to the item).
2 Weakly aligned to this California content standard - Not a very good example of the standards.
3 Highly aligned this California content standard - Good and reasonable example of the standards.
4 Fully aligned to the California content standards - Exemplary item, clear example of standard for which it is matched.

Step 5: Rate the overall quality of the item.

Next, rate the overall quality of the item. Is the item clear and precise? Could you understand what the item is asking (NOT whether you are capable of answering the question correctly)? use the scale below to make your judgments.

1 Item is of poor overall quality (Rating requires annotation).
2 Item is of good quality, but has some easily repairable flaw (Rating requires annotation).
3 Item is of good quality, typical of what you would expect on this and similar tests.
4 Item is of exceptional quality (annotations encouraged).

Step 6: Perform the steps above for each item on the assessment.

Step 7: Return your current assessment form and receive the next assessment form (total of three to four assessments to review).

Depth-of-Knowledge (DoK) Descriptions for CSTs

DOK Levels for Reading

- **Reading Level 1 (recall)** item requires students to receive or recite facts or to use simple skills or abilities, such as word pronunciation, verbatim recitation of text, or definitions of recognition of figurative language.

  Keywords: Identify, list, determine, define

- **Reading Level 2 (skills/concepts)** item calls for engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Examples include using context cues to identify the meaning of unfamiliar words or summarizing major events in a narrative.

  Key words: summarize, interpret, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed.

- **Reading Level 3 (strategic thinking)** Students must synthesize ideas from the text to show understanding of ideas. They also may need to go beyond the text. Students must explain, generalize, or connect ideas. Items require reasoning and planning, and may involve abstract theme identification, inference across an entire passage, or students' application of prior knowledge.

  Keywords: Compare/contrast, analyze, explain, synthesize or connect ideas (single text), similarities and differences, apply, infer, support

- **Reading Level 4 (extended thinking)** Higher order thinking is central, such as complex, reasoning, planning, inference, and synthesis of ideas from multiple sources. Students may need to develop hypotheses, perform critical analysis, and make connections among texts. Items may require extended time and thinking.

NOTE: Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met.

Keywords: Predict, discuss, dispute, connect to self, critically analyze, synthesize or connect (multiple texts)

DOK Levels for Writing

- **Writing Level 1 (recall):** requires the student to write or recite simple facts. This writing or recitation does not include complex synthesis or analysis but basic ideas. The students are engaged in listing ideas or words as in a brainstorming activity prior to written composition, are engaged in a simple spelling or vocabulary assessment or are asked to write simple sentences. Students are expected to write and speak using Standard English conventions. This includes using appropriate grammar, punctuation, capitalization and spelling. Some examples that represent but do not constitute all of Level 1 performance are:

    1. Use punctuation marks correctly.
    2. Identify Standard English grammatical structures and refer to resources for correction.

    Keywords: Identify, list, determine, define

- **Writing Level 2 (skills/concepts):** requires some mental processing. At this level students are engaged in first draft writing or brief extemporaneous speaking for a limited number of purposes and audiences. Students are beginning to connect ideas using a simple organizational structure. For example, students may be engaged in note-taking, outlining or simple summaries. Text may be limited to one paragraph. Students demonstrate a basic understanding and appropriate use of such reference materials as a dictionary, thesaurus, or web site. Some examples that represent but do not constitute all of Level 2 performance are:

    1. Construct compound sentences.
    2. Use simple organizational strategies to structure written work.

    Key words: summarize, interpret, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed.

- **Writing Level 3 (strategic thinking):** requires some higher level mental processing. Students are engaged in developing compositions that include multiple paragraphs. These compositions may include complex sentence structure and may demonstrate some synthesis and analysis. Students show awareness of their audience and purpose through focus, organization and the use of appropriate compositional elements. The use of appropriate compositional elements includes such things as addressing chronological order in a narrative or including supporting facts and details in an informational report. At this stage students are engaged in editing and revising to improve the quality of the composition. Some examples that represent but do not constitute all of Level 3 performance are:

    1. Support ideas with details and examples.
    2. Use voice appropriate to the purpose and audience.
    3. Edit writing to produce a logical progression of ideas

Keywords: Compare/contrast, analyze, explain, synthesize or connect ideas (single text), similarities and differences, apply, infer, support

- **Writing Level 4 (extended thinking)**: Higher-level thinking is central to Level 4. The standard at this level is a multi- paragraph composition that demonstrates synthesis and analysis of complex ideas or themes. There is evidence of a deep awareness of purpose and audience. For example, informational papers include hypotheses and supporting evidence. Students are expected to create compositions that demonstrate a distinct voice and that stimulate the reader or listener to consider new perspectives on the addressed ideas and themes. An example that represents but does not constitute all of Level 4 performance is:

    1. Write an analysis of two selections, identifying the common theme and generating a purpose that is appropriate for both.

NOTE: Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met.

Keywords: Predict, discuss, dispute, connect to self, critically analyze, synthesize or connect (multiple texts)

Example Format of Depth-of-Knowledge (DoK) Rating Sheet for CST Standards

DOK Rating Form
ELA Content Standards
Grade 2

Enter rating of 1 to 4 in blanks below.

CALIFORNIA CONTENT STANDARDS: READING	CODE
1.0 WORD ANALYSIS, FLUENCY, AND SYSTEMATIC VOCABULARY DEVELOPMENT: Students understand the basic features of reading. They select letter patterns and know how to translate them into spoken language by using phonics, syllabication, and word parts. They apply this knowledge to achieve fluent oral and silent reading.	
1.1 **Word Recognition:** read narrative and expository text aloud with grade-appropriate fluency and accuracy and with appropriate pacing, intonation, and expression	**NA***
1.2 **Vocabulary and Concept Development:** apply knowledge of word origins, derivations, synonyms, antonyms, and idioms to determine the meaning of words and phrases	
1.3 **Vocabulary and Concept Development:** use knowledge of root words to determine the meaning of unknown words within a passage	
1.4 **Vocabulary and Concept Development:** know common roots and affixes derived from Greek and Latin and use this knowledge to analyze the meaning of complex words (e.g., international)	
1.5 **Vocabulary and Concept Development:** use a thesaurus to determine related words and concepts	
1.6 **Vocabulary and Concept Development:** distinguish and interpret multiple meaning words	
2.0 READING COMPREHENSION: Students read and understand grade-level-appropriate material. They draw upon a variety of comprehension strategies as needed (e.g., generating and responding to essential questions, making predictions, comparing information from several sources). The selections in *Recommended Readings in Literature, Kindergarten Through Grade Eight* illustrate the quality and complexity of the materials to be read by students. In addition to their regular school reading, students read one-half million words annually, including a good representation of grade-level-appropriate narrative and expository text (e.g., classic and contemporary literature, magazines, newspapers, online information).	
2.1 **Structural Features of Informational Materials:** identify structural patterns found in informational text (e.g., compare and contrast, cause and effect, sequential or chronological order, proposition and support) to strengthen comprehension	

CALIFORNIA CONTENT STANDARDS: READING	CODE
2.2 **Comprehension and Analysis of Grade-Level-Appropriate Text:** use appropriate strategies when reading for different purposes (e.g., full comprehension, location of information, personal enjoyment)	**NA***
2.3 **Comprehension and Analysis of Grade-Level-Appropriate Text:** make and confirm predictions about text by using prior knowledge and ideas presented in the text itself, including illustrations, titles, topic sentences, important words, and foreshadowing clues	
2.4 **Comprehension and Analysis of Grade-Level-Appropriate Text:** evaluate new information and hypotheses by testing them against known information and ideas	
2.5 **Comprehension and Analysis of Grade-Level-Appropriate Text:** compare and contrast information on the same topic after reading several passages or articles	
2.6 **Comprehension and Analysis of Grade-Level-Appropriate Text:** distinguish between cause and effect and between fact and opinion in expository text	
2.7 **Comprehension and Analysis of Grade-Level-Appropriate Text:** follow multiple-step instructions in a basic technical manual (e.g., how to use computer commands or video games)	
3.0 LITERARY RESPONSE AND ANALYSIS: Students read and respond to a wide variety of significant works of children's literature. They distinguish between the structural features of the text and the literary terms or elements (e.g., theme, plot, setting, characters). The selections in *Recommended Readings in Literature, Kindergarten Through Grade Eight* illustrate the quality and complexity of the materials to be read by students.	
3.1 **Structural Features of Literature:** describe the structural differences of various imaginative forms of literature, including fantasies, fables, myths, legends, and fairy tales	
3.2 **Narrative Analysis of Grade-Level-Appropriate Text:** identify the main events of the plot, their causes, and the influence of each event on future actions	
3.3 **Narrative Analysis of Grade-Level-Appropriate Text:** use knowledge of the situation and setting and of a character's traits and motivations to determine the causes for that character's actions	
3.4 **Narrative Analysis of Grade-Level-Appropriate Text:** compare and contrast tales from different cultures by tracing the exploits of one character type and develop theories to account for similar tales in diverse cultures (e.g., trickster tales)	
3.5 **Narrative Analysis of Grade-Level-Appropriate Text:** define figurative language (e.g., simile, metaphor, hyperbole, personification) and identify its use in literary works	

CALIFORNIA CONTENT STANDARDS: WRITING	CODE
1.0 WRITTEN AND ORAL ENGLISH LANGUAGE CONVENTIONS: Students write and speak with a command of standard English conventions appropriate to this grade level.	
1.1 **Sentence Structure:** use simple and compound sentences in writing and speaking	
1.2 **Sentence Structure:** combine short, related sentences with appositives, participial phrases, adjectives, adverbs, and prepositional phrases	
1.3 **Grammar:** identify and use regular and irregular verbs, adverbs, prepositions, and coordinating conjunctions in writing and speaking	
1.4 **Punctuation:** use parentheses, commas in direct quotations, apostrophes in the possessive case of nouns and in contractions	
1.5 **Punctuation:** use underlining, quotations marks, or italics to identify titles of documents	
1.6 **Capitalization:** capitalize names of magazines, newspapers, works of art, musical compositions, organizations, and the first word in quotations when appropriate	
1.7 **Spelling:** spell correctly roots, inflections, suffixes and prefixes, and syllable constructions	
2.0 WRITING STRATEGIES: Students write clear, coherent sentences and paragraphs that develop a central idea. Their writing shows they consider the audience and purpose. Students progress through the stages of the writing process (i.e., pre-writing, drafting, revising, editing successive versions).	
2.1 **Organization and Focus:** select a focus, an organizational structure, and a point of view based upon purpose, audience, length, and format requirements	
2.2 **Organization and Focus:** create multiple-paragraph compositions that	
1) provide an introductory paragraph	
2) establish and support a central idea with a topic sentence at or near the beginning of the first paragraph	
3) include supporting paragraphs with simple facts, details, and explanations	
4) conclude with a paragraph that summarizes the points	
5) use correct indentation	**NA***
2.3 **Organization and Focus:** use traditional structures for conveying information (e.g., chronological order, cause and effect, similarity and difference, and posing and answering a question)	
2.4 **Penmanship:** write fluidly and legibly in cursive or joined italic	**NA***
2.5 **Research and Technology:** quote or paraphrase information sources, citing them appropriately	
2.6 **Research and Technology:** locate information in reference texts by using organizational features (e.g., prefaces, appendices)	

CALIFORNIA CONTENT STANDARDS: WRITING	CODE
2.7 **Research and Technology:** use various reference materials (e.g., dictionary, thesaurus, card catalog, encyclopedia, on-line information) as an aid to writing	
2.8 **Research and Technology:** understand the organization of almanacs, newspapers, and periodicals and how to use those print materials	
2.9 **Research and Technology:** demonstrate basic keyboarding skills and familiarity with computer terminology (e.g., cursor, software, memory, disk drive, hard drive)	**NA***
2.10 **Evaluation and Revision:** edit and revise selected drafts to improve coherence and progression by adding, deleting, consolidating, and rearranging text	

Example of Item Rating Sheet for CST Review

1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						

Example of Alignment Instructions for CAPA Review

CAPA Alignment Tasks for English-language Arts

For the alignment review, we would like for you to evaluate the performance tasks for the California Alternate Performance Assessment (CAPA) compared to the alternate assessment standards. You will be reviewing the performance tasks for each level (Level I through Level V). These tasks will be performed individually. You will be entering your ratings and responses in the Excel spreadsheet setup on the laptop provided.

Step 1: Rate the depth-of-knowledge of the CAPA ELA standards.

There are several files on your laptop that include the Math California content standards. Open the file labeled CAPA_ELA_Level… for the level that you will be working on currently. After you open the file, please enter your name and ID number under the appropriate columns. Your name will be deleted from the file once all of the data are merged. Enter the Grade Level that you will be working on currently.

Using the attached depth-of-knowledge (DOK) descriptions and examples, rate each content objective per standard on the degree of cognitive processing expected of students to demonstrate proficiency. In the CAPA_ELA_Level…, enter the DOK level (number) in the spreadsheet under the column labeled DOK Rating. Remember that *cognitive complexity* is related to *difficulty*, but these terms are not synonymous. If you find that a single content objective really requires several different tasks of varying complexity (i.e., “Students should be able to identify, distinguish, and explain…”), indicate the *highest* DOK level required by this content objective.

Step 2: Rate the depth-of-knowledge (DOK) level of the item.

Next, open the spreadsheet labeled CAPA_ELA_ItemRatings_Grade… file for the current grade level assessment. After you open the file, please enter your name and ID number under the appropriate columns. Your name will be deleted from the file once all of the data are merged. Enter the Grade Level that you will be working on currently.

Using the attached depth-of-knowledge (DOK) descriptions and examples, rate each item on the degree of cognitive processing required of students to answer the item adequately. Enter the DOK level (number) in the spreadsheet under the column labeled Item DOK Rating. Remember that *cognitive complexity* is related to *difficulty*, but these terms are not synonymous. If you find that a single item really requires several different tasks of varying complexity (i.e., “Students should be able to identify, distinguish, and explain…”), indicate the *highest* DOK level required by this content objective.

When you have finished rating the DOK level of the current CAPA level, you may move onto the next level. However, the HumRRO group leader will be stopping to review each level together after everyone has completed it. At this time, there will be group discussion to achieve majority agreement on the ratings.

Repeat the step above for EACH level of the CAPA ELA standards.

Step 3: Match the item to a specific content objective.

Find the content standard that you think best reflects the content that the task is supposed to assess. Within the content standard, try identify at least one content objective that the task targets. In the CAPA ELA Ratings file, enter the code for this content standard and objective (located on the left-hand column of your Coded Content Standards printout sheet next to each standard) into the Excel spreadsheet under the column labeled Content Standard/Objective 1.

For the CAPA, you may find that a performance task assesses two or more content standards or objectives. In this case, include the additional standard and objective in the column labeled Content Standard/Objective 2.

Step 4: Rate the overall match level of the task to the standard and objective you chose.

Indicate how well you think that each task actually assesses the standard you selected. Please use the rating scale on the next page to make your judgment. Enter the appropriate rating number from the scale into your spreadsheet.

1 Not aligned to any California alternate content standard (Use ONLY if you did not assign a standard to the item).
2 Weakly aligned to this California alternate content standard - Not a very good example of the standards.
3 Highly aligned this California alternate content standard - Good and reasonable example of the standards.
4 Fully aligned to the California alternate content standards - Exemplary item, clear example of standard for which it is matched.

Step 5: Perform the steps above for <u>each task</u> in Level I of the alternate assessment.

Step 6: Return your current assessment form and receive the next assessment form (total of five levels of assessment tasks to review).

Additional Instructions and Considerations for Evaluation of the CAPA

(1) Please try to review and keep in mind the task preparations, cues, and scoring rubrics for each task as you review them.

(2) Several sets of Stimulus Cards are available to review for those CAPA items that require them. Please simply ask a HumRRO staff member for a copy.

Depth-of-Knowledge (DOK) Descriptions for CAPA

DOK Level 1

Requires students to recall or observe facts, definitions, terms. Involves simple one-step procedures. Involves computing simple algorithms (e.g., sum, quotient).

DOK Level 2

This level includes the engagement of some mental processing beyond a habitual response. The item requires students to make some decisions as to how to approach a problem or activity.

Keywords: classify, organize estimate, make observations, collect and display data, and compare data.

DOK Level 3

A multiple step 'behavioral event' is executed in more than one context. Requires reasoning, planning, or use of evidence to solve problem or algorithm. May involve activity with more than one possible answer. Requires conjecture or restructuring of problems. Involves drawing conclusions from observations, citing evidence and developing logical arguments for concepts. Uses concepts to solve non-routine problems.

DOK Level 4

The 'behavioral event' reflects an approach (of many) to completing the task. May require complex reasoning, planning, developing and thinking. Typically requires extended time to complete problem, but time spent not on repetitive tasks. Requires students to make several connections and apply one approach among many to solve the problem. Involves complex restructuring of data, establishing and evaluating criteria to solve problems.

Example Format of Depth-of-Knowledge (DoK) Rating Sheet for CAPA Standards

Enter rating of 1 to 4 in blanks below.

DOK Rating Form
ELA CAPA Content Standards
Levels I

CALIFORNIA CONTENT STANDARDS	
Reading Number of Tasks: 4 **Percentage of Test: 50%**	**CODE**
Kindergarten	
1.0 WORD ANALYSIS, FLUENCY, AND SYSTEMATIC VOCABULARY DEVELOPMENT: Students know about letters, words, and sounds. They apply this knowledge to read simple sentences.	
1.3 **Concepts About Print: Understand that printed materials provide information.**	
✓ Identify environmental symbols/signs/cues.	
✓ Match symbol or cue to activity or function.	
Grade 1	
1.0 WORD ANALYSIS, FLUENCY, AND SYSTEMATIC VOCABULARY DEVELOPMENT: Students understand the basic features of reading. They select letter patterns and know how to translate them into spoken language by using phonics, syllabication, and word parts. They apply this knowledge to achieve fluent oral and silent reading.	
1.17 **Vocabulary and Concept Development:** Classify grade-appropriate categories of words (e.g., concrete collections of animals, foods, toys).	
✓ Identify object by function.	
✓ Sort objects by function/use.	
✓ Identify picture by function.	
2.0 READING COMPREHENSION: Students read and understand grade-level-appropriate material. They draw upon a variety of comprehension strategies as needed (e.g., generating and responding to essential questions, making predictions, comparing information from several sources). The selections in *Recommended Readings in Literature, Kindergarten Through Grade Eight* illustrate the quality and complexity of the materials to be read by students. In addition to their regular school reading, by grade four, students read one-half million words annually, including a good representation of grade-level-appropriate narrative and expository text (e.g., classic and contemporary literature, magazines, newspapers, online information). In grade one, students begin to make progress toward this goal.	
2.3 **Comprehension and Analysis of Grade-Level-Appropriate Text:** Follow one-step written instructions.	
✓ Identify a picture/object/word cue.	

Writing Number of Tasks: 1 **Percentage of Test: 12.5%**	
Kindergarten	
1.0 WRITING STRATEGIES: Students write words and brief sentences that are legible.	
1.3 **Organization and Focus:** Write by moving from left to right and from top to bottom.	
✓ Demonstrate left to right/top to bottom sequencing in a variety of activities.	
✓ Hold writing implement.	
✓ Make marks on paper.	
✓ Trace/copy purposeful marks on paper.	
Listening and Speaking Number of Tasks: 3 **Percentage of Test: 37.5%**	
Kindergarten	
1.0 LISTENING AND SPEAKING STRATEGIES: Students listen and respond to oral communication. They speak in clear and coherent sentences.	
1.1 **Comprehension:** Understand and follow one-and-two-step oral directions.	
✓ Orient in direction of speaker.	
✓ Respond to voice by stopping activity or going to source of sound.	
✓ Attend to speaker for duration of activity.	
1.2 **Comprehension:** Share information and ideas, speaking audibly in complete, coherent sentences.	
✓ Communicate wants/needs using a gesture, action, voice output device or vocalization.	
✓ Communicate choice using a gesture, action, voice output device or vocalization.	
Grade 1	
1.0 LISTENING AND SPEAKING STRATEGIES: Students listen critically and respond appropriately to oral communication. They speak in manner that guides the listener to understand important ideas by using proper phrasing, pitch, and modulation.	
1.1 **Comprehension:** Listen attentively.	
✓ Orient in direction of speaker.	
✓ Respond to voice by stopping activity or going to source of sound.	
✓ Attend to speaker for duration of activity.	
Total Level I Tasks: Total Number of Tasks: 8 **Percentage of Test: 100%**	

Example of Item Rating Sheet for CAPA Review

Name:			ID:		Content Area: CAPA ELA	
Item Number	Version	Depth-of-Knowledge	Content Strand/Objective 1	Content Strand/Objective 2	Overall Alignment	Explanation
(Number Listed in Test Form)		(Enter Level 1to 4)	(Enter Standard ID Code)	(Enter Standard ID Code)	(Enter Scale of 1 to 4)	Use ONLY IF you entered a 'Source of Challenge'
	1					
	1					
	1					
	1					
	2					
	2					
	2					
	2					
	3					
	3					
	3					
	3					
	4					
	4					
	4					
	4					
	5					
	5					
	5					

Name:			ID:		Content Area: CAPA ELA	
Item Number	Version	Depth-of-Knowledge	Content Strand/Objective 1	Content Strand/Objective 2	Overall Alignment	Explanation
(Number Listed in Test Form)		(Enter Level 1to 4)	(Enter Standard ID Code)	(Enter Standard ID Code)	(Enter Scale of 1 to 4)	Use ONLY IF you entered a 'Source of Challenge'
	5					
	6					
	6					
	6					
	6					
	7					
	7					
	7					
	7					