New England Common Assessment Program

2008–2009

Technical Report

June 2009

Measured Progress

100 Education Way, Dover, NH 03820 (800) 431-8901

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW 1
1.1 Purpose of the New England Common Assessment Program 1
1.2 Purpose of this Report 1
1.3 Organization of this Report 2

SECTION I—DESCRIPTION OF THE 2008–09 NECAP TEST 3

CHAPTER 2. DEVELOPMENT AND TEST DESIGN 3
2.1 Operational Development Process 3
2.1.1 Item Development 3
2.1.2 Grade-Level and Grade-Span Expectations 3
2.1.3 Internal Item Review 3
2.1.4 External Item Review 4
2.1.5 Bias and Sensitivity Review 5
2.1.6 Item Editing 5
2.1.7 Reviewing and Refining 6
2.1.8 Operational Test Assembly 6
2.1.9 Editing Drafts of Operational Tests 7
2.1.10 Braille and Large-Print Translation 7
2.2 Item Types 7
2.3 Operational Test Designs and Blueprints 8
2.3.1 Embedded Equating and Field Test Items 9
2.3.2 Test Booklet Design 9
2.4 Reading Test Designs 9
2.4.1 Reading Blueprint 10
2.5 Mathematics Test Design 12
2.5.1 The Use of Calculators on the NECAP 12
2.5.2 Mathematics Blueprint 13
2.6 Writing Test Design 14
2.6.1 Writing Blueprint: Grades 5 and 8 15
2.6.2 Writing Blueprint: Grade 11 16
2.7 Test Sessions 17
CHAPTER 3. TEST ADMINISTRATION 21
3.1 Responsibility for Administration 21
3.2 Administration Procedures 21
3.3 Participation Requirements and Documentation 21
3.4 Administrator Training 24
3.5 Documentation of Accommodations 24
3.6 Test Security 25
3.7 Test and Administration Irregularities 25
3.8 Test Administration Window 26
3.9 NECAP Service Center 27
CHAPTER 4. SCORING 29
4.1 Scoring of Standard Test Items 29
4.1.1 Machine-Scored Items 29
4.1.2 Hand-Scored Items 29
4.1.2.1 Scoring Location and Staff 30
4.1.2.2 Benchmarking Meetings with the NECAP State Specialists 31
4.1.2.3 Reader Recruitment and Qualifications 32
4.1.2.4 Methodology for Scoring Polytomous Items 32
4.1.2.5 Reader Training 33
4.1.2.6 Senior Quality Assurance Coordinator (QAC) and Senior Reader (SR) Training 35
4.1.2.7 Monitoring of Scoring Quality Control and Consistency 35
4.1.2.8 Reports Generated During Scoring 39

SECTION II—STATISTICAL AND PSYCHOMETRIC SUMMARIES 41

CHAPTER 5. CLASSICAL ITEM ANALYSES 41
5.1 Classical Difficulty and Discrimination Indices 41
5.2 Differential Item Functioning (DIF) 43
5.3 Dimensionality Analyses 44
CHAPTER 6. IRT SCALING AND EQUATING 48
6.1 Item Response Theory Scaling 48

6.2 Item Response Theory Analyses *50*
6.3 Equating *51*
6.4 Equating Results *52*
6.5 Achievement Standards *53*
6.6 Reported Scale Scores *54*
6.6.1 Description of Scale 54
6.6.2 Calculations 55
6.6.3 Distributions 57
CHAPTER 7. RELIABILITY 61
7.1 Reliability and Standard Errors of Measurement *62*
7.2 Subgroup Reliability *63*
7.3 Item Type Reliability *63*
7.4 Reporting Categories Reliability *64*
7.5 Reliability of Achievement Level Categorization *65*
7.5.1 Accuracy and Consistency 65
7.5.2 Calculating Accuracy 66
7.5.3 Calculating Consistency 66
7.5.4 Calculating Kappa 66
7.5.5 Results of Accuracy, Consistency, and Kappa Analyses 67
CHAPTER 8. VALIDITY 69
8.1 Questionnaire Data *70*
8.2 Validity Studies Agenda *75*
8.2.1 External Validity 75
8.2.2 Convergent and Discriminant Validity 75
8.2.3 Structural Validity 76
8.2.4 Procedural Validity 77

SECTION III—NECAP REPORTING 79

CHAPTER 9. SCORE REPORTING 79
9.1 Teaching Year vs. Testing Year Reporting *79*
9.2 Primary Reports *79*
9.3 Student Report *79*
9.4 Item Analysis Reports *80*
9.5 School and District Results Reports *81*
9.6 School and District Summary Reports *84*
9.7 Decision Rules *84*
9.8 Quality Assurance *85*

SECTION IV—REFERENCES 87

APPENDICES 89

Appendix A Committee Membership
Appendix B Accommodation Frequencies by Content Area
Appendix C Table of Standard Test Accommodations
Appendix D Appropriateness of the Accommodations Allowed in NECAP General Assessment, and Their Impact on Student Results
Appendix E Double-blind Interrater Agreement
Appendix F Item Level Classical Statistics Results
Appendix G Classical Item Statistic Descriptives Summarized by Grade, Content, and Form
Appendix H Number of Items Classified into DIF Categories by Subgroup, Test Form, and Item Type
Appendix I Common Item DIF Category Counts in the Male vs. Female Comparison
Appendix J Item Response Theory Parameters, TCCs and TIFs
Appendix K Delta Analyses and Rescore Analysis Results
Appendix L Raw to Scaled Score Look-up Tables
Appendix M Scaled Score Percentages and Cumulative Percentages
Appendix N Detailed Alpha Coefficient Results
Appendix O Decision Accuracy and Consistency Results
Appendix P Student Questionnaire Data
Appendix Q Sample Reports
Appendix R Decision Rules

Chapter 1. OVERVIEW

1.1 Purpose of the New England Common Assessment Program

The New England Common Assessment Program (NECAP) is the result of collaboration among New Hampshire (NH), Rhode Island (RI), and Vermont (VT) to build a set of tests for grades 3 through 8 and 11 to meet the requirements of the No Child Left Behind Act (NCLB). The purposes of the tests are as follows: (1) provide data on student achievement in reading/language arts and mathematics to meet the requirements of NCLB; (2) provide information to support program evaluation and improvement; and (3) provide information regarding student and school performance to both parents and the public. The tests are constructed to meet rigorous technical criteria, to include universal design elements and accommodations to allow all students access to test content, and to gather reliable student demographic information for accurate reporting. School improvement is supported by

- providing a transparent test design through the elementary and middle school grade-level expectations (GLEs), the high school grade-span expectations (GSEs), distributions of emphasis, and practice tests
- reporting results by GLE/GSE subtopics, released items, and subgroups
- hosting report interpretation workshops to foster understanding of results

It is important to note that the NECAP tests in reading, mathematics, and writing are administered in the fall at the beginning of the school year and test student achievement based on the *prior year's* GLEs/GSEs. Student-level results are provided to schools and families for use as one piece of evidence about progress and learning that occurred on the prior year's GLEs/GSEs. The results are a status report of a student's performance against GLEs/GSEs and should be used cautiously in concert with local data.

1.2 Purpose of this Report

The purpose of this report is to document the technical aspects of the 2008–09 NECAP. In October of 2008, students in grades 3 through 8 and 11 participated in the administration of the NECAP in reading and mathematics. Students in grades 5, 8, and 11 also participated in writing. This report provides information about the technical quality of those tests, including a description of the processes used to develop, administer, and score the tests and to analyze the test results. This report is intended to serve as a guide for replicating and/or improving the procedures in subsequent years.

Though some parts of this technical report may be used by educated laypersons, the intended audience is experts in psychometrics and educational research. The report assumes a working knowledge of measurement concepts, such as "reliability" and "validity," and statistical concepts, such as "correlation" and "central tendency." In some chapters, knowledge on more advanced topics is required.

1.3 Organization of this Report

The organization of this report is based on the conceptual flow of a test's life span; the report begins with the initial test specification and addresses all the intermediate steps that lead to final score reporting. Section I provides a description of the NECAP test, and consists of four chapters – test design and development process; test administration; scoring; scaling and equating. Section II provides statistical and psychometric summaries, and consists of three chapters – item analysis; reliability; and validity. Section III covers NECAP score reporting. Section IV provides references, and Section V contains appendices to the report.

SECTION I—DESCRIPTION OF THE 2008–09 NECAP TEST

Chapter 2. DEVELOPMENT AND TEST DESIGN

2.1 Operational Development Process

2.1.1 Item Development

Curriculum and assessment content specialists at Measured Progress begin the item development process by selecting passages and graphics and developing items and scoring guides according to guidance and specifications provided by the NECAP states. During this process, the curriculum and assessment specialists do the following:

- Work in close collabaoration with the states from conceptualization to production of the final camera ready copies
- Review the Grade-Level, (GLEs), Grade-Span Expectations (GSEs), and test specifications to ensure that the items developed truly measure student learning and meet each state's goals for accountability
- Write and edit items that adhere to the NECAP test specifications
- Lead Item Review Committees
- Participate in benchmarking to ensure NECAP scoring reflects the true intent of the items
- Select items to create tests that are educationally significant, as well as valid and reliable for the purposes of reporting data

2.1.2 Grade-Level and Grade-Span Expectations

NECAP test items are directly linked to the *content standards* and *performance indicators* described in the GLEs/GSEs. The content standards for each grade are grouped into content clusters for the purpose of reporting results; the performance indicators are used by the content specialists to help guide the development of test questions. An item may address one, several, or all of the performance indicators.

2.1.3 Internal Item Review

- The lead Measured Progress test developer within the content area reviewed the formatted item, open response (OR) scoring guide, and any reading selections and graphics.
- The lead developer considered item "integrity," content, and structure; appropriateness to designated content area; item format; clarity; possible ambiguity; answer cueing; appropriateness and quality of reading selections and graphics; and appropriateness of scoring guide descriptions

and distinctions (in relation to each item and across all items within the guide). The item reviewer also ensured that, for each item, there was only one correct answer.

- The lead developer also considered scorability and evaluated whether the scoring guide adequately addressed performance on the item.
- Fundamental questions that the lead developer considered, but was not limited to, included the following:

  - What is the item asking?
  - Is the key the only possible key? (Is there only *one* correct answer?)
  - Is the OR item scorable as written (were the correct words used to elicit the response defined by the guide)?
  - Is the wording of the scoring guide appropriate and parallel to the item wording?
  - Is the item complete (i.e., includes scoring guide, content codes, key, grade level, depth of knowledge (DOK) and identified contract)?
  - Is the item appropriate for the designated grade level?

2.1.4 External Item Review

Item Review Committees (IRCs) were formed by the states to provide an external review of items. The committees included teachers, curriculum supervisors, and higher-education faculty from all three states, with committee members serving rotating terms. (A list of IRC member names and affiliations is included in Appendix A.) The committees role is to review test items for the NECAP, provide feedback , and make recommendations about which items should be selected for program use. The 2008–09 NECAP IRCs for each content area in grade levels 3 through 8 and 11 met in the spring of 2008. Committee members reviewed the entire set of embedded field-test items proposed for the 2008–09 operational test and made recommendations about selecting, revising, or eliminating specific items from the item pool. Members reviewed each item against the following criteria:

- Grade-Level/Grade-Span Expectation Alignment
  - Is the test item aligned to the appropriate GLE/GSE?
  - If not, which GLE/GSE or grade level is more appropriate?

- Correctness
  - Are the items and distractors correct with respect to content accuracy and developmental appropriateness?
  - Are the scoring guides consistent with GLE/GSE wording and developmental appropriateness?

- Depth of Knowledge[1]
  - Are the items coded to the appropriate Depth of Knowledge?
  - If consensus cannot be reached, is there clarity around why the item might be on the borderline of two levels?

- Language
  - Is the item language clear?
  - Is the item language accurate (syntax, grammar, conventions)?

- Universal Design
  - Is there an appropriate use of simplified language (does not interfere with the construct being assessed)?
  - Are charts, tables, and diagrams easy to read and understandable?
  - Are charts, tables, and diagrams necessary to the item?
  - Are instructions easy to follow?
  - Is the item amenable to accommodations—read-aloud, signed, or Brailled?

2.1.5 Bias and Sensitivity Review

Bias review is an essential part of the development process. During the bias review process, NECAP items were reviewed by a committee of teachers, English Language Learner (ELL) specialists, special-education teachers, and other educators and members of major constituency groups who represent the interests of legally protected and/or educationally disadvantaged groups. (A list of bias and sensitivity review committee member names and affiliations is included in Appendix A.) Items were examined for issues that might offend or dismay students, teachers, or parents. Including such groups in the development of test items and materials can prevent many unduly controversial issues, and can allay unfounded concerns before the test forms are produced.

2.1.6 Item Editing

Measured Progress editors reviewed and edited the items to ensure uniform style (based on *The Chicago Manual of Style,* 15th edition) and adherence to sound testing principles. These principles included the stipulation that items

- were correct with regard to grammar, punctuation, usage, and spelling;
- were written in a clear, concise style;

[1] NECAP employed the work of Dr. Norman Webb to guide the development process with respect to Depth of Knowledge. Test specification documents identified ceilings and targets for Depth of Knowledge coding.

- contained unambiguous explanations to students detailing what is required to attain a maximum score;
- were written at a reading level that would allow the student to demonstrate his or her knowledge of the tested subject matter, regardless of reading ability;
- exhibited high technical quality in terms of psychometric characteristics;
- had appropriate answer options or score-point descriptors; and
- were free of potentially sensitive content.

2.1.7 Reviewing and Refining

Test developers presented item sets to the IRCs who then recommended which items should be included in the embedded field-test portions of the test. The NH, RI, and VT Departments of Education content specialists made the final selections with the assistance of Measured Progress at a final face-to-face meeting.

2.1.8 Operational Test Assembly

At Measured Progress, test assembly is the sorting and laying out of item sets into test forms. Criteria considered during this process for the 2008–09 NECAP included the following:

- *Content coverage/match to test design*. The Measured Progress test developers completed an initial sorting of items into sets based on a balance of reporting categories across sessions and forms, as well as a match to the test design (e.g., number of multiple-choice (MC), short-answer (SA), and constructed-response (CR) items).
- *Item difficulty and complexity*. Item statistics drawn from the data analysis of previously tested items were used to ensure similar levels of difficulty and complexity across forms.
- *Visual balance*. Item sets were reviewed to ensure that each reflected similar length and "density" of selected items (e.g., length/complexity of reading selections, number of graphics).
- *Option balance*. Each item set was checked to verify that it contained a roughly equivalent number of key options (As, Bs, Cs, and Ds).
- *Name balance*. Item sets were reviewed to ensure that a diversity of student names was used.
- *Bias.* Each item set was reviewed to ensure fairness and balance based on gender, ethnicity, religion, socioeconomic status, and other factors.
- *Page fit.* Item placement was modified to ensure the best fit and arrangement of items on any given page.
- *Facing-page issues*. For multiple items associated with a single stimulus (a graphic or reading selection), consideration was given both to whether those items needed to begin on a left- or right-hand page and to the nature and amount of material that needed to be placed on facing

pages. These considerations served to minimize the amount of "page flipping" required of students.

- *Relationship between forms*. Although embedded field-test items differ from form to form, they must take up the same number of pages in each form so that sessions and content areas begin on the same page in every form. Therefore, the number of pages needed for the longest form often determined the layout of each form.
- *Visual appeal*. The visual accessibility of each page of the form was always taken into consideration, including such aspects as the amount of "white space," the density of the text, and the number of graphics.

2.1.9 Editing Drafts of Operational Tests

Any changes made by a test construction specialist were reviewed and approved by a lead developer. After a form was laid out in what was considered its final form, it was reviewed to identify any final considerations, including the following:

- *Editorial changes*. All text was scrutinized for editorial accuracy, including consistency of instructional language, grammar, spelling, punctuation, and layout. (based on Measured Progress's publishing standards in *The Chicago Manual of Style*, 15th edition).
- *"Keying" items*. Items were reviewed for any information that might "key" or provide information that would help to answer another item. Decisions about moving keying items are based on the severity of the "key-in" and the placement of the items in relation to each other within the form.
- *Key patterns*. The final sequence of keys was reviewed to ensure that their order appeared random (i.e., no recognizable pattern and no more than three of the same key in a row).

2.1.10 Braille and Large-Print Translation

Common items for grades 3 through 8 and 11were translated into Braille by a subcontractor that specializes in test materials for blind and visually impaired students. In addition, Form 1 for each grade was adapted into a large-print version.

2.2 Item Types

The item types used and the functions of each are described below.

Multiple-Choice (MC) items were administered in grades 3 through 8 and 11 in reading and mathematics, and in grades 5 and 8 in writing to provide breadth of coverage of the GLEs/GSEs. Because they require approximately one minute for most students to answer, these items make efficient use of limited

testing time and allow coverage of a wide range of knowledge and skills, including, for example, word identification and vocabulary skills.

Short-Answer (SA) items were administered in grades 3 through 8 and 11 mathematics to assess students' skills and their ability to work with brief, well-structured problems with one solution or a very limited number of solutions. SA items require approximately two to five minutes for most students to answer. The advantage of this item type is that it requires students to demonstrate knowledge and skills by generating rather than merely selecting, an answer. SA planning boxes accompanied writing prompts in grades 5 and 8 as noted below.

Constructed-Response (CR) items typically require students to use higher-order thinking skills such as summary, evaluation, and analysis in constructing a satisfactory response. CR items require approximately five to ten minutes for most students to complete. These items were administered in grades 3 through 8 and 11 in reading, in grades 5 and 8 in writing, and in grades 5 through 8 and 11 in mathematics.

Writing Prompt (WP). A single common writing prompt with three SA planning box items were administered in grades 5 and 8. A single common writing prompt and one additional matrix writing prompt per form were administered in grade 11. Students were given 45 minutes (plus additional time if necessary) to compose an extended response for the common prompt that was scored by two independent readers both on quality of the stylistic and rhetorical aspects of the writing and on the use of standard English conventions. Students were encouraged to write a rough draft and were advised by the test administrator when to begin copying their final draft into their student answer booklets.

Approximately twenty-five percent of the common NECAP items were released to the public in 2008–09. The released NECAP items are posted on a Web site hosted by Measured Progress and on the Department of Education Web sites. Schools are encouraged to incorporate the use of released items in their instructional activities so that students will be familiar with the types of questions found on the NECAP assessment.

2.3 Operational Test Designs and Blueprints

Since the beginning of the program, the goal of the NECAP has been to measure what students know and are able to do by using a variety of test item types. The program was structured to use both common and matrix-sampled items. (Common items are those taken by all students at a given grade level. Matrix-sampled items comprise a pool that is divided among the multiple forms of the test at each grade level. Their purpose is described in the next section.) This design provides reliable and valid results at the student level, and breadth of coverage of a content area at the school results level while minimizing testing time. (Note: Only common items count toward students' scaled scores.)

2.3.1 Embedded Equating and Field Test Items

To ensure that NECAP scores obtained from different test forms and different years are equivalent to each other, a set of equating items is matrixed across forms of the reading and mathematics tests. Chapter 5 presents more detail on the equating process. (Note: Equating items are not counted toward students' scaled scores.)

The NECAP also includes embedded field-test items in all content areas except grades 5 and 8 writing. Because the field-test items are taken by many students, the sample is sufficient to produce reliable data from which to inform the process of selecting items for future tests. Embedding field-test items achieves two other objectives. First, it creates a pool of replacement items in reading and mathematics that are needed due to the release of common items each year. Second, embedding field-test items into the operational test ensures that students take the items under operational conditions. (Note: As with the matrixed equating items, field-test items are not counted toward students' scaled scores.)

2.3.2 Test Booklet Design

To accommodate the embedded equating and field-test items in the 2008–09 NECAP, there were nine unique test forms in grades 3 through 8 and eight unique forms in grade 11. In all reading and mathematics test sessions, the equating and field-test items were distributed among the common items in a way that was not evident to test takers. The grades 5 and 8 writing design called for one common test form that was made up of a single writing prompt with three SA planning box items, four CR items, and ten MC items. The grade 11 writing design called for each student to respond to two writing prompts. The first writing prompt was common for all students and the second writing prompt was either a matrix prompt or a field-test prompt, depending on the particular test form.

2.4 Reading Test Designs

Table 2-1 summarizes the number and types of items that were used in the 2008–09 NECAP reading test for grades 3 through 8. Note that in reading, all students received the common items and one of either the equating or field-test forms. Each MC item was worth one point, and each CR item was worth four points.

Table 2-1. 2008–09 NECAP: Number of Items by Item Type and Number of Items—Reading Grades 3–8

Common: 2 long and 2 short passages plus 4 stand-alone MC		*Matrix – Equating Forms 1–3: 1 long and 1 short passage plus 2 stand-alone MC*		*Matrix – FT Forms 4–7: 1 long and 1 short passage plus 2 stand-alone MC*		*Matrix – FT Forms 8–9: 3 short passages plus 2 stand-alone MC*		*Total per student: 3 long and 3 short or 2 long and 5 short passages plus 6 stand-alone MC*	
MC	CR	MC	CR	MC	CR	MC	CR	MC	CR
28	6	14	3	14	3	14	3	42	9

Long passages have 8 MC and 2 CR items; short passages have 4 MC and 1 CR items MC = multiple choice; CR = constructed response; FT = field-test

Table 2-2 summarizes the numbers and types of items that were used in the 2008–09 NECAP reading test for grade 11. Note that in reading, all students received the common items and one of either the equating or field-test forms. Each MC item was worth one point, and each CR item was worth four points.

Table 2-2. Number of Items by Item Type and Number of Items—Reading Grade 11

Common: 2 long and 2 short passages plus 4 stand-alone MC		*Matrix – Equating Forms 1–2: 1 long and 1 short passage plus 2 stand-alone MC*		*Matrix – FT Forms 3–8: 1 long and 1 short passage plus 2 stand-alone MC*		*Total per student 3 long and 3 short passages plus 6 stand-alone MC*	
MC	CR	MC	CR	MC	CR	MC	CR
28	6	14	3	14	3	42	9

Long passages have 8 MC and 2 CR items; short passages have 4 MC and 1 CR items; MC = multiple choice; CR = constructed response; FT = field-test

2.4.1 Reading Blueprint

As indicated earlier, the test framework for reading in grades 3 through 8 was based on the NECAP GLEs, and all items on the NECAP test were designed to measure a specific GLE. The test framework for reading in grade 11 was based on the NECAP GSEs, and all items on the NECAP test were designed to measure a specific GSE. The reading passages on all the NECAP tests are broken down into the following categories:

- Literary passages, representing a variety of forms: modern narratives; diary entries; drama; poetry; biographies; essays; excerpts from novels; short stories; and traditional narratives, such as fables, tall tales, myths, and folktales.
- Informational passages/factual text, often dealing with areas of science and social studies. These passages are taken from such sources as newspapers, magazines, and book excerpts. Informational text could also be directions, manuals, recipes, etc. The passages are authentic texts selected from grade level appropriate reading sources that students would be likely to encounter in both classroom and independent reading. All passages are collected from published works.

Reading comprehension is assessed on the NECAP test by items that are dually-categorized by the type of text and by the level of comprehension measured. The level of comprehension is designated as either "Initial Understanding" or "Analysis and Interpretation." Word identification and vocabulary skills are assessed at each grade level primarily through MC items. The distribution of emphasis for reading is shown in Table 2-3.

Table 2-3. 2008–09 NECAP: Distribution of Emphasis Across Reporting Subcategories In Terms of Targeted Percentage of Test, by Grade—Reading Grades 3–8 and 11

Subcategory	*GLE/GSE Grade (Grade Tested)*						
	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	7 (8)	9–10 (11)
Word Identification Skills and Strategies	20%	15%	0%	0%	0%	0%	0%
Vocabulary Strategies/Breadth of Vocabulary	20%	20%	20%	20%	20%	20%	20%
Initial Understanding of Literary Text	20%	20%	20%	20%	15%	15%	15%
Initial Understanding of Informational Text	20%	20%	20%	20%	20%	20%	20%
Analysis and Interpretation of Literary Text	10%	15%	20%	20%	25%	25%	25%
Analysis and Interpretation of Informational Text	10%	10%	20%	20%	20%	20%	20%
Total	100%	100%	100%	100%	100%	100%	100%

Table 2-4 shows the content category reporting structure for reading and the maximum possible number of raw score points that students could earn. (With the exception of word identification/vocabulary items, reading items were reported in two ways: type of text and level of comprehension.) Note: Because only common items are counted toward students' scaled scores, only common items are reflected in this table.

Table 2-4. 2008–09 NECAP: Reporting Subcategories and Possible Raw Score Points by Grade—Reading Grades 3–8 and 11

Subcategory	*Grade Tested*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Word ID/Vocabulary	20	19	10	10	10	10	10
Type of Text							
Literary	16	16	21	21	21	21	21
Informational	16	17	21	21	21	21	21
Level of Comprehension							
Initial Understanding	19	19	18	19	19	18	18
Analysis and Interpretation.	13	14	24	23	23	24	24
Total	52	52	52	52	52	52	52

Total possible points in reading equals the sum of Word ID/Vocabulary points and the total points from either Type of Text or Level of Comprehension, (since reading comprehension items are dually-categorized by type of text and level of comprehension).

Table 2-5 lists the percentage of total score points assigned to each Depth of Knowledge (DOK) level in reading.

Table 2-5.2008–09 NECAP: Depth of Knowledge (DOK) In Terms of Targeted Percentage of Test, by Grade—Reading Grades 3–8 and 11

DOK	*Grade*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Level 1	65%	53%	24%	29%	32%	26%	29%
Level 2	32%	44%	65%	65%	53%	65%	65%
Level 3	3%	3%	11%	6%	15%	9%	6%
Total	100%	100%	100%	100%	100%	100%	100%

2.5 Mathematics Test Design

Tables 2-6 through 2-8 summarize the numbers and types of items that were used in the 2008–09 NECAP mathematics tests for grades 3 and 4, 5 through 8, and grade 11, respectively. Note that all students received the common items plus equating and field-test items in their forms. Each MC item was worth one point, each SA item either one or two points, and each CR item four points. Score points within a grade level were evenly divided, so that MC items represented approximately fifty percent of possible score points, and SA and CR items together represented approximately fifty percent of score points.

Table 2-6.2008–09 NECAP: Item Type and Number of Items—Mathematics Grades 3 and 4

Common			*Matrix – Equating (9 Forms)*			*Matrix – FT (9 Forms)*			*Total per Student*		
MC	SA1	SA2	MC	SA1	SA2	MC	SA1	SA2	MC	SA1	SA2
35	10	10	6	2	2	3	1	1	44	13	13

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer; FT = field test
Total of 9 forms, 6 unique. Forms 1, 2, and 3 are repeated as forms 7, 8, and 9.

Table 2-7. 2008–09 NECAP: Item Type and Number of Items—Mathematics Grades 5–8

Common				*Matrix – Equating (9 Forms)*				*Matrix – FT (9 Forms)*				*Total per Student*			
MC	SA1	SA2	CR	MC	SA1	SA2	CR	MC	SA1	SA2	CR	MC	SA1	SA2	CR
32	6	6	4	6	2	2	1	3	1	1	1	41	9	9	6

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer; FT = field test
Total of 9 forms, 6 unique. Forms 1, 2, and 3 are repeated as forms 7, 8, and 9.

Table 2-8. 2008–09 NECAP: Item Type and Number of Items—Mathematics Grade 11

Common				*Matrix – Equating (8 Forms)*				*Matrix – FT (8 Forms)*				*Total per Student*			
MC	SA1	SA2	CR	MC	SA1	SA2	CR	MC	SA1	SA2	CR	MC	SA1	SA2	CR
24	12	6	4	4	2	1	1	4	2	1	1*	32	16	8	6

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer; FT = field test
* = 4 unique CR items are field-tested across forms
Total of 8 forms, 6 unique. Forms 1 and 2 are repeated as forms 7 and 8.

2.5.1 The Use of Calculators on the NECAP

The mathematics specialists from the NH, RI, and VT Departments of Education who designed the mathematics test acknowledge the importance of mastering arithmetic algorithms. At the same time, they understand that the use of calculators is a necessary and important skill. Calculators can save time and prevent error in the measurement of some higher-order thinking skills, and in turn allow students to work on more sophisticated and intricate problems. For these reasons, it was decided that at grades 3 through 8 calculators should be prohibited in the first of the three sessions of the NECAP mathematics test and permitted in the remaining two sessions. At grade 11, it was decided that calculators should be prohibited in the first of the two sessions and permitted in the second session. (Test sessions are discussed in greater detail at the end of this chapter.)

2.5.2 Mathematics Blueprint

The test framework for mathematics at grades 3 through 8 was based on the NECAP GLEs, and all items on the grades 3 through 8 NECAP tests were designed to measure a specific GLE. The test framework for mathematics at grade 11 was based on the NECAP GSEs, and all items on the grade 11 NECAP test were designed to measure a specific GSE. The mathematics items are organized into the four content strands as follows:

- Numbers and Operations: Students understand and demonstrate a sense of what numbers mean and how they are used. Students understand and demonstrate computation skills.
- Geometry and Measurement: Students understand and apply concepts from geometry. Students understand and demonstrate measurement skills.
- Functions and Algebra: Students understand that mathematics is the science of patterns, relationships, and functions. Students understand and apply algebraic concepts.
- Data, Statistics, and Probability: Students understand and apply concepts of data analysis. Students understand and apply concepts of probability.

Additionally, problem solving, reasoning, connections, and communication are embedded throughout the GLEs/GSEs. The distribution of emphasis for mathematics reporting subcategories is shown in Table 2-9.

Table 2-9. 2008–09 NECAP: Distribution of Emphasis (in targeted percentage of test)—Mathematics Grades 3–8 and 11

Subcategory	*Grade*						
	2 (3)	*3 (4)*	*4 (5)*	*5 (6)*	*6 (7)*	*7 (8)*	*9–10 (1)1*
Numbers and Operations	55%	50%	45%	40%	30%	20%	15%
Geometry and Measurement	15%	20%	20%	25%	25%	25%	30%
Functions and Algebra	15%	15%	20%	20%	30%	40%	40%
Data, Statistics, and Probability	15%	15%	15%	15%	15%	15%	15%
Total	100%	100%	100%	100%	100%	100%	100%

Table 2-10 shows the subcategory reporting structure for mathematics and the maximum possible number of raw score points that students could earn. The goal for distribution of score points or balance of representation, across the four content strands varies from grade to grade. Note: Only common items are reflected in this table, as only they are counted toward students' scaled scores.

Table 2-10. 2008–09 NECAP: Reporting Subcategories and Possible Raw Score Points Grade —Mathematics Grades 3–8 and 11

Subcategory	*Grade Tested*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Numbers and Operations	35	32	30	26	20	13	10
Geometry and Measurement	10	13	13	17	16	17	19
Functions and Algebra	10	10	13	13	20	26	26
Data, Statistics, and Probability	10	10	10	10	10	10	9
Total	65	65	66	66	66	66	64

Table 2-11 lists the percentage of total score points assigned to each level of DOK in mathematics.

Table 2-11. 2008–09 NECAP: Depth of Knowledge (DOK) by Grade (in percentage of test)—Mathematics Grades 3–8 and 11

DOK	*Grade*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Level 1	25%	27%	27%	23%	44%	27%	35%
Level 2	71%	66%	69%	71%	50%	63%	61%
Level 3	4%	7%	4%	6%	6%	10%	4%
Total	100%	100%	100%	100%	100%	100%	100%

2.6 Writing Test Design

Table 2-12 summarizes the numbers and types of items that were used in the 2008–09 NECAP writing test for grades 5 and 8. Note that all items on the grades 5 and 8 writing tests were common. Each MC item was worth one point, each CR item four points, each SA item one point, and the writing prompt 12 points.

Table 2-12.2008–09 NECAP: Number of Items by Item Type (All Items Common) and Number of Items—Writing Grades 5 and 8

MC	*CR*	*SA1*	*WP*
10	3	3	1

MC = multiple-choice; CR = constructed response; SA1 = 1-point short answer; WP = writing prompt

Table 2-13 summarizes the test design used in the 2008–09 NECAP writing test for grade 11. Each grade 11 student responded to two different writing prompts, one common and either one matrix-equating or one field-test prompt. The common prompt was worth 12 points.

Table 2-13. 2008–09 NECAP: Number of Items by Item Type and Number of Items—Writing Grade 11 (8 forms)

Common	*Matrix – Equating (5 Forms)*	*Matrix – Field Test (3 Forms)*
1 Writing Prompt	1 Writing Prompt	1 Writing Prompt

2.6.1 Writing Blueprint: Grades 5 and 8

The test framework for grades 5 and 8 writing was based on the *NECAP Grade-Level Expectations* (GLEs), and all items on the NECAP test were designed to measure a specific GLE. The content standards for grades 5 and 8 writing identify four major genres that are assessed in the writing portion of the NECAP test each year.

- Writing in response to literary text
- Writing in response to informational text
- Narratives
- Informational writing (report/procedure text for grade 5 and persuasive essay for grade 8)

The writing prompt and the three CR items each address a different genre. In addition, structures of language and writing conventions are assessed through MC items and throughout the student-writing test. The prompts and CR items were developed with the following criteria as guidelines:

- the prompts must be interesting to students
- the prompts must be accessible to all students (i.e., all students would have something to say about the topic)
- the prompts must generate sufficient text to be effectively scored

The category reporting structure for grades 5 and 8 writing is shown in Table 2-14. The table provides the maximum possible number of raw score points that students could earn. The content category "Short Responses" lists the total raw score points from the three CR items; the reporting category "Extended Response" lists the total raw score points from the three SA items and the writing prompt.

Table 2-14. 2008–09 NECAP: Reporting Subcategory and Possible Raw Score Points Possible by Grade—Writing Grades 5 and 8

Subcategory	*Grade Tested*	
	Grade 5	Grade 8
Structures of Language and Writing Conventions	10	10
Short Response	12	12
Extended Response	15	15
Total	37	37

Short response = CR items; Extended response = SA items and writing prompt

Table 2-15 lists the percentage of total score points assigned to each level of DOK in writing.

Table 2-15. 2008–09 NECAP: Depth of Knowledge (DOK) by Grade (in percentage of test)—Writing Grades 5 and 8

DOK	*Grade Tested*	
	Grade 5	Grade 8
Level 1	35%	47%
Level 2	41%	29%
Level 3	24%	24%
Total	100%	100%

2.6.2 Writing Blueprint: Grade 11

The test framework for grade 11 writing was based on the NECAP GSEs, and all items on the NECAP test were designed to measure a specific GSE. The content standards for grade 11 writing identify six genres that are grouped into three major strands:

- Writing in response to text (literary and informational)
- Informational writing (report, procedure, and persuasive essay)
- Expressive writing (reflective essay)

The writing prompts (common, matrix equating, and field test) in combination, address each of the different genres. The prompts were developed using the following criteria as guidelines:

- the prompt must be interesting to students
- the prompt must be accessible to all students (i.e., all students would have something to write about the topic)
- the prompt must generate sufficient text to be effectively scored

There is only one reporting category, "Extended Response," for grade 11 writing, as shown in Table 2-16.

Table 2-16. 2008–09 NECAP: Reporting Subcategories and Possible Raw Score Points—Writing Grade 11

Subcategory	*Grade 11*
Extended Response	12
Total	12

Table 2-17 lists the percentage of total score points assigned to each level of DOK in writing.

Table 2-17. 2008–09 NECAP: Depth of Knowledge (DOK) In Terms of Targeted Percentage of Test, by Grade—Writing Grade 11

DOK	*Grade 11*
Level 1	0%
Level 2	0%
Level 3	100%
Total	100%

2.7 Test Sessions

The NECAP tests were administered October 1–23, 2008 to grades 3 through 8 and 11. During the testing window, schools were able to schedule testing sessions at any time, but were instructed to follow the sequence in the scheduling guidelines as detailed in the test administration manual. It was also mandatory that all testing classrooms within a grade level be on the same schedule. A third week during the testing window was reserved for make-up testing of students who were absent during initial test sessions.

The timing and scheduling guidelines for the NECAP tests were based on estimates of the time it would take an average student to respond to each type of item on the test:

- multiple-choice – 1 minute
- short answer (1 point) – 1 minute
- short answer (2 point) – 2 minutes
- constructed response – 10 minutes
- long writing prompt – 45 minutes

For the reading sessions, the scheduling guidelines estimate that reading the stimulus material (passage) will take approximately 10 minutes. Tables 2-18 through 2-24 show the distribution of items across the test sessions for each content area and grade level.

Table 2-18.2008–09 NECAP: Item Type by Session—Reading Grades 3–8

	Session 1	*Session 2*	*Session 3*
MC	14	14	14
CR	3	3	3

MC = multiple-choice; CR = constructed response

Table 2-19. 2008–09 NECAP: Item Type by Session—Reading Grade 11

Item Type1	*Session 1*	*Session 2*
MC	22	20
CR	4	5

MC = multiple-choice; CR = constructed response

Table 2-20. 2008–09 NECAP: Item Type by Session—Mathematics Grades 3 and 4

Item Type	*Session 1*	*Session 2*	*Session 3*
MC	15	15	14
SA1	4	3	6
SA2	4	5	4

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer

Table 2-21. 2008–09 NECAP: Item Type by Session—Mathematics Grades 5–8

Item Type	*Session 1*	*Session 2*	*Session 3*
MC	14	14	13
SA1	3	3	3
SA2	3	3	3
CR	2	2	2

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer; CR = constructed response

Table 2-22. 2008–09 NECAP: Item Type by Session—Mathematics Grade 11

Item Type	*Session 1*	*Session 2*
MC	16	16
SA1	6	6
SA2	6	6
CR	3	3

MC = multiple-choice; SA1 = 1-point short answer; SA2 = 2-point short answer; CR = constructed response

Table 2-23. 2008–09 NECAP: Item Type by Session—Writing Grades 5 and 8

Item Type	*Session 1*	*Session 2*
MC	10	0
CR	3	0
SA1	0	3
WP	0	1

MC = multiple-choice; CR = constructed response; SA1 = 1-point short answer; WP = writing prompt

Table 2-24. 2008–09 NECAP: Item Type by Session—Writing Grade 11

Item Type	*Session 1*	*Session 2*
MC	0	0
CR	0	0
SA1	0	0
WP	1	1

MC = multiple-choice; CR = constructed response; SA1 = 1-point short answer; WP = writing prompt

Although the scheduling guidelines are based on the assumption that most students will complete the test within the estimated time, each test session allowed additional time for students who may have needed it. Up to one-hundred percent additional time was allocated for each session (i.e., a 50-minute session could be extended by an additional 50 minutes).

If classroom space was not available for students who required additional time to complete the tests, schools were allowed to consider using another space for this purpose, such as a guidance office. If additional areas were not available, it was recommended that each classroom used for test administration be scheduled for the maximum amount of time. Detailed instructions regarding test administration and scheduling were provided in the test coordinators' and administrators' manuals.

Chapter 3. TEST ADMINISTRATION

3.1 Responsibility for Administration

The 2008 *NECAP Principal/Test Coordinator Manual* indicated that principals and/or their designated NECAP test coordinator were responsible for the proper administration of the NECAP. Uniformity of administration procedures from school to school was ensured by using manuals that contained explicit directions and scripts to be read aloud to students by test administrators.

3.2 Administration Procedures

Principals and/or the schools' designated NECAP test coordinator were instructed to read the *Principal/Test Coordinator Manual* before testing and to be familiar with the instructions provided in the *Test Administrator Manual*. The *Principal/Test Coordinator Manuals* included a section highlighting aspects of test administration that were new for the year and checklists to help prepare for testing. The checklists outlined tasks to be performed by school staff before, during, and after test administration. In addition to these checklists, the *Principal/Test Coordinator Manuals* described the testing material sent to each school, how to inventory it, track it during administration, and return it after testing was complete. The *Test Administrator Manual* included checklists for the administrators to use to prepare themselves, their classrooms, and the students for the administration of the test. The *Test Administrator Manual* contained sections that detailed the procedures to be followed for each test session and instructions for preparing the material before the principal/test coordinator would return it to Measured Progress.

3.3 Participation Requirements and Documentation

The Department of Education's intent is for *all* students in grades 3 though 8 and 11 to participate in the NECAP through standard administration, administration with accommodations, or alternate assessment. Furthermore, any student who is absent during any session of the NECAP is expected to take a make-up test within the three-week testing window.

Schools were required to return a student answer booklet for every enrolled student in the grade level, with the exception of students who took an alternate assessment in the 2007–08 school year and therefore were not required participating in the NECAP in 2008–09. On those occasions when it was deemed impossible to test a particular student, school personnel were required to inform their Department of Education. A grid was included on the student answer booklets that listed the approved reasons why a booklet could be returned blank for one or more sessions of the test:

- Student is new to the United States after October 1, 2007 and is LEP (reading and writing only)

First-year LEP students who took the ACCESS test of English language proficiency, as scheduled in their states, were not required to take the reading and writing tests in 2008–09. However, these students were required to take the mathematics test in 2008–09.

- Student withdrew from school after October 1, 2008

  If a student withdrew after October 1, 2008 but before completing all of the test sessions, school personnel were instructed to code this reason on the student's answer booklet.

- Student enrolled in school after October 1, 2008

  If a student enrolled after October 1, 2008 and was unable to complete all of the test sessions before the end of the test administration window, school personnel were instructed to code this reason on the student's answer booklet.

- State-approved special consideration

  Each state Department of Education had a process for documenting and approving circumstances that made it impossible or not advisable for a student to participate in testing.

- Student was enrolled in school on October 1, 2008 and did not complete test for reasons other than those listed above

  If a student was not tested for a reason other than those stated above, school personnel were instructed to code this reason on the student's answer booklet. These "Other" categories were considered "not state-approved."

Tables 3-1, 3-2, and 3-3 list the participation rates of the three states combined in reading, mathematics, and writing.

Table 3-1. 2008–09 NECAP: Participation Rates—Reading

Category	*Description*	*Enrollment*	*Not Tested State-Approved*	*Not Tested Other*	*Number Tested*	*Percent Tested*
All	All Students	231,012	3,136	2,352	225,524	98
Gender	Male	119,552	1,913	1,455	116,184	97
	Female	111,425	1,220	896	109,309	98
	Not Reported	35	3	1	31	89
Ethnicity	Am. Indian	1,006	20	24	962	96
	Asian	5,480	156	77	5,247	96
	Black	9,745	237	155	9,353	96
	Hispanic	18,197	482	250	17,465	96
	NHPI	97	4	2	91	94
	White	195,395	2,205	1,807	191,383	98
	Not Reported	1,092	32	37	1,023	94
LEP	Current	5,898	533	156	5,209	88
	Monitoring Year 1	1,089	4	4	1,081	99
	Monitoring Year 2	1,198	2	8	1,188	99
	Other	222,827	2,597	2,184	218,046	98
IEP	IEP	37,426	2,200	907	34,319	92
	Other	193,586	936	1,445	191,205	99
SES	SES	66,593	1,335	912	64,346	97
	Other	164,419	1,801	1,440	161,178	98
Migrant	Migrant	86	2	2	82	95
	Other	230,926	3,134	2,350	225,442	98
Title 1	Title 1	34,210	651	247	33,312	97
	Other	196,802	2,485	2,105	192,212	98
Plan 504	Plan 504	1,232	6	6	1,220	99
	Other	229,780	3,130	2,346	224,304	98

Table 3-2. 2008–09 NECAP: Participation Rates—Mathematics

Category	*Description*	*Enrollment*	*Not Tested State-Approved*	*Not Tested Other*	*Number Tested*	*Percent Tested*
All	All Students	231,012	2,673	2,431	225,908	98
Gender	Male	119,552	1,655	1,519	116,378	97
	Female	111,425	1,016	911	109,498	98
	Not Reported	35	2	1	32	91
Ethnicity	Am. Indian	1,006	20	26	960	95
	Asian	5,480	56	59	5,365	98
	Black	9,745	153	156	9,436	97
	Hispanic	18,197	223	231	17,743	98
	NHPI	97	2	2	93	96
	White	195,395	2,197	1,919	191,279	98
	Not Reported	1,092	22	38	1,032	95
LEP	Current	5,898	49	93	5,756	98
	Monitoring Year 1	1,089	3	4	1,082	99
	Monitoring Year 2	1,198	3	9	1,186	99
	Other	222,827	2,618	2,325	217,884	98
IEP	IEP	37,426	2,212	994	34,220	91
	Other	193,586	461	1,437	191,688	99
SES	SES	66,593	1,114	942	64,537	97
	Other	164,419	1,559	1,489	161,371	98
Migrant	Migrant	86	2	2	82	95
	Other	230,926	2,671	2,429	225,826	98
Title 1	Title 1	31,812	391	245	31,176	98
	Other	199,200	2,282	2,186	194,732	98
Plan 504	Plan 504	1,232	6	5	1221	99
	Other	229,780	2,667	2,426	224,687	98

Table 3-3. 2008–09 NECAP: Participation Rates—Writing

Category	*Description*	*Enrollment*	*Not Tested State-Approved*	*Not Tested Other*	*Number Tested*	*Percent Tested*
All	*All Students*	*102,467*	*1,344*	*1,979*	*99,144*	*97*
Gender	Male	52,516	805	1,268	50,443	96
	Female	49,932	539	710	48,683	97
	Not Reported	19	0	1	18	95
Ethnicity	Am. Indian	452	9	22	421	93
	Asian	2,256	64	54	2,138	95
	Black	4,214	93	122	3,999	95
	Hispanic	7,781	181	195	7,405	95
	NHPI	46	0	2	44	96
	White	87,278	983	1,555	84,740	97
	Not Reported	440	14	29	397	90
LEP	Current	2,078	191	98	1,789	86
	Monitoring Year 1	400	2	4	394	99
	Monitoring Year 2	510	1	4	505	99
	Other	99,479	1,150	1,873	96,456	97
IEP	IEP	16,881	925	774	15,182	90
	Other	85,586	419	1,205	83,962	98
SES	SES	27,345	585	745	26,015	95
	Other	75,122	759	1,234	73,129	97
Migrant	Migrant	24	1	1	22	92
	Other	102,443	1,343	1,978	99,122	97
Title 1	Title 1	12,246	242	184	11,820	97
	Other	90,221	1,102	1,795	87,324	97
Plan 504	Plan 504	627	5	3	619	99
	Other	101,840	1,339	1,976	98,525	97

3.4 Administrator Training

In addition to distributing the *Principal/Test Coordinator Manual* and *Test Administrator Manual*, the NH, RI, and VT Departments of Education, along with Measured Progress, conducted test administration workshops in regional locations in each state to inform school personnel about the NECAP and to provide training on the policies and procedures regarding administration of the tests. RI and VT reduced the number of workshops conducted in each state, but produced a CD of the workshop PowerPoint presentation with audio, which was produced and sent to each school in these states to use for additional training.

3.5 Documentation of Accommodations

The *Principal/Test Coordinator Manual* and *Test Administrator Manual* provided directions for coding information related to accommodations and modifications on page 2 of the student answer booklet. All accommodations used during any test session were required to be coded by authorized school personnel—not students—after testing was completed.

An *Accommodations, Guidelines, and Procedures: Administrator Training Guide* was also produced to provide detailed information on planning and implementing accommodations. This guide was available on each state's Department of Education Web site. The states collectively made the decision that

accommodations be made available to all students based on individual need regardless of disability status. Decisions regarding accommodations were to be made by the students' educational team on an individual basis and were to be consistent with those used during the students' regular classroom instruction. Making accommodations decisions for a group rather than on an individual basis was not permitted. If the decision made by a student's educational team required an accommodation not listed in the state-approved Table of Standard Test Accommodations, schools were instructed to contact the Department of Education in advance of testing for specific instructions for coding in the "Other Accommodations (E)" and/or "Modifications (F)" sections.

Appendix B shows the accommodation frequencies by content area for the October 2008 NECAP test administration. The accommodation codes are defined in the Table of Standard Test Accommodations, which can be found in Appendix C. Information on the appropriateness and impact of accommodations may be found in Appendix D.

3.6 Test Security

Maintaining test security is critical to the success of the NECAP and the continued partnership among the three states. The *Principal/Test Coordinator Manual* and the *Test Administrator Manual*s explain in detail all test security measures and test administration procedures. School personnel were informed that any concerns about breaches in test security were to be reported to the school's test coordinator and/or principal immediately. The test coordinator and/or principal were responsible for immediately reporting the concern to the District Superintendent and the State Assessment Director at the Department of Education. Test security was also strongly emphasized at test administration workshops that were conducted in all three states. The three states also required principals to log on to a secure Web site to complete the *Principal's Certification of Proper Test Administration* form for each grade level tested at their school. Principals were requested to provide the number of secure tests received from Measured Progress, the number of tests administered to students, and the number of secure test materials that they were returning to Measured Progress. Principals were instructed to submit the form by entering a unique password, which acted as their digital signature. By signing and submitting the form, the principal was certifying that the tests were administered according to the test administration procedures outlined in the *Principal/Test Coordinator Manual* and *Test Administrator Manual*, that the security of the tests was maintained, that no secure material was duplicated or in any way retained in the school, and that all test materials had been accounted for and returned to Measured Progress.

3.7 Test and Administration Irregularities

During test administration, a printing error was discovered in some of the grade 3 and grade 11 NECAP test booklets. These booklets were produced by two different print vendors. The printing irregularities for grade 3 were not the same across booklets and occurred mainly in forms 3 and 7. The most common errors included pages inserted upside down or pages 3–10 missing from a booklet. The print vendor

determined that the errors occurred while the binding machine was being loaded: page was either loaded incorrectly or upside down. (Cameras are in place to check for specific images from a page, and the machine stops if the correct image is not captured three times in a row. An operator is then tasked with finding the error and pulling the incorrect booklets.) The vendor believes this was an isolated incident due to human error.

The grade 11 printing error occurred only in form 7 and only five defective booklets were reported. The booklets were missing pages 17–24, and pages 25–40 were repeated. The print vendor determined that an operator loaded an incorrect group of pages into one of the pockets in the binding machine. A camera checks for correct pages in each booklet as binding takes place. When a problem is detected, the machine shuts down, the operator removes the incorrect booklets and inspects the booklets immediately ahead on the production line. The vendor believes the incorrect booklets were missed by the operator while correcting the problem.

During login, printing errors were found and reported in seven grade 3 booklets and five grade 11 booklets, either by schools or by Measured Progress. All affected schools either replaced the defective test booklets with extra test booklets they already had available or Measured Progress immediately sent new test booklets to the school when an error was reported. NECAP reports were not affected by these irregularities.

Before the test administration window began but after test forms were printed and shipped to schools, an error was discovered in forms 3 and 4 of the grade 5 test booklets. During item selection by test developers and the test form production process, a mistake was made in the slotting of item types into each form. Form 3 was constructed with items 19 and 20 both as SA2 items (2-point short-answer items), and form 4 was constructed with items 19 and 20 as CR items (4-point constructed-response items). There should have been one of each item type in each form. The scannable answer booklets were designed as such. In form 3, there was sufficient space to answer the SA2 items. But in form 4, the short-answer item space in position 19 was insufficient to answer the CR item put errantly in that position. (It should be noted that the CR item was a matrix item and not used for computing student scores.)

On September 29th, Measured Progress sent both a letter and revised administration instructions to all schools with grade 5 students informing them of the error in the grade 5 form 4 test booklets. Test administrators were asked to find all form 4 booklets and place an X over item 19 and the corresponding answer space. Administrators were asked to follow a script alerting students with form 4 tests that they would find item 19 had been crossed out in their booklets. They were told that there was not enough room for them to answer the question and they should skip it. Otherwise, administrators were advised to continue with the test session as outlined in the *Test Administrator Manual.* No student scores or reports were affected by this error.

3.8 Test Administration Window

The test administration window was October 1–23, 2008.

3.9 NECAP Service Center

To provide additional support to schools before, during, and after testing, Measured Progress established the NECAP Service Center. The support of the Service Center is essential to the successful administration of any statewide test program. It provides a centralized location to which individuals in the field can call using a toll-free number to ask specific questions or report any problems they may be experiencing. Representatives are responsible for receiving, responding to and tracking calls, then routing issues to the appropriate person(s) for resolution. All calls are logged into a database which includes notes regarding the issue and resolution of each call.

The Service Center was staffed year-round by representatives depending upon need and call volume, and available from 8:00 AM to 4:00 PM. Extra representatives were available beginning two weeks before the start of testing and ending two weeks after testing.

Chapter 4. SCORING

4.1 Scoring of Standard Test Items

Upon receipt of used NECAP answer booklets following testing, Measured Progress scanned all student responses, along with student identification and demographic information. Imaged data for multiple-choice responses were machine-scored. Images of open-response items were processed and organized by iScore, a secure, server-to-server electronic scoring software designed by Measured Progress, for hand-scoring.

Student responses that could not be physically scanned (e.g., answer documents damaged during shipping), and typed responses submitted according to applicable test accommodations, were physically reviewed and scored on an individual basis by trained, qualified readers. These scores were linked to the student's demographic data and merged with the student's scoring file by Measured Progress's data processing department.

4.1.1 Machine-Scored Items

MC item responses were compared to scoring keys using item analysis software. Correct answers were assigned a score of one point and incorrect answers were assigned zero points. Student responses with multiple marks and blank responses were also assigned zero points.

The hardware elements of the scanners monitor themselves continuously for correct read, and the software that drives these scanners also monitors correct data reads. Standard checks include recognition of a sheet that does not belong, is upside down, or is backwards and identification of critical data that are missing (e.g., a student ID number), test forms that are out of range or missing, and page or document sequence errors. When a problem is detected, the scanner stops and displays an error message directing the operator to investigate and to correct the situation.

4.1.2 Hand-Scored Items

The images of student responses to CR items were hand-scored through the iScore system. Use of iScore minimizes the need for readers to physically handle answer booklets and related scoring materials. Student confidentiality was easily maintained, since all NECAP scoring was "blind" (i.e., district, school, and student names were not visible to readers) the iScore system maintained the linkage between the student response images and their associated test booklet numbers.

Through iScore, qualified readers at computer terminals accessed electronically scanned images of student responses. Readers evaluated each response and recorded each score via keypad or mouse entry through the iScore system. When a reader finished one response, the next response appeared immediately on the computer screen.

Imaged responses from all answer booklets were sorted into item-specific groups for scoring purposes. Readers reviewed responses from only one item at a time; however, imaged responses from a student's entire booklet were always available for viewing when necessary, and the physical booklet was also available to the Chief Reader onsite. (Chief Reader and other scoring roles are described in the section that follows.)

The use of iScore also helped ensure that access to student response images was limited to only those who were scoring or working for Measured Progress in a scoring management capacity.

4.1.2.1 Scoring Location and Staff

Scoring Location

The iScore database, its operation, and its administrative controls are all based in Dover, NH; however, Table 4-1 presents the locations where 2008–09 NECAP test item responses by grade and content area were scored.

Table 4-1.2008–09 NECAP: Scoring Locations by Content and Grade

Content /Grade Level	*Louisville, KY*	*Dover, NH*	*Troy, NY*	*Longmont, CO*
Reading Grade 3		X		
Reading Grade 4		X		
Reading Grade 5	X			
Reading Grade 6	X			
Reading Grade 7	X			
Reading Grade 8				X
Reading Grade 11				X
Mathematics Grade 3	X			
Mathematics Grade 4	X			
Mathematics Grade 5	X			
Mathematics Grade 6	X			
Mathematics Grade 7	X			
Mathematics Grade 8			X	
Mathematics Grade 11			X	
Writing Grade 5				X
Writing Grade 8				X
Writing Grade 11				X

The iScore system monitored accuracy, reliability, and consistency across all scoring sites. Constant daily communication and coordination were accomplished through e-mail, telephone, faxes, and secure Web sites, to ensure that critical information and scoring modifications were shared and implemented across all scoring sites.

Staff Positions

The following staff members were involved with scoring the 2008–09 NECAP responses:

- The NECAP Scoring Project Manager, an employee of Measured Progress, was located in Dover, NH and oversaw communication and coordination of scoring across all scoring sites.
- The iScore Operational Manager and iScore administrators, employees of Measured Progress, were located in Dover, NH and coordinated technical communication across all scoring sites.
- A Chief Reader (CR) in each content area (mathematics, reading, and writing) ensured consistency of scoring across all scoring sites for all grades tested in that content area. CRs also provided read-behind activities (defined in a later section) for Quality Assurance Coordinators. CRs are employees of Measured Progress.
- Numerous Quality Assurance Coordinators (QACs), selected from a pool of experienced Senior Readers for their ability to score accurately and their ability to instruct and train readers, participated in benchmarking activities for each specific grade and content area. QACs provided read-behind activities (defined in a later section) for Senior Readers at their sites. The ratio of QACs and Senior Readers to Readers was approximately 1:11.
- Numerous Senior Readers (SRs) selected from a pool of skilled and experienced Readers, provided read-behind activities (defined in a later section) for the Readers at their scoring tables (2–12 Readers at each table). The ratio of QACs and SRs to Readers was approximately 1:11.
- Readers at scoring sites scored operational and field-test NECAP 2008–09 student responses. Recruitment of Readers is described in Section 4.1.2.3.

4.1.2.2 Benchmarking Meetings with the NECAP State Specialists

In preparation for implementing NECAP scoring guidelines, Measured Progress scoring staff prepared and facilitated benchmarking meetings held with NECAP state specialists from their respective departments of education. The purpose of these meetings was to establish guidelines for scoring NECAP items during the current field-test scoring session and for future operational scoring sessions.

Several dozen student responses for each item were selected that CRs identified as illustrative mid-range examples of the respective score points. CRs presented these responses to the NECAP content specialists during benchmarking meetings and worked collaboratively with them to finalize an authoritative set of score-point exemplars for each field-test item. As a matter of practice, these sets are included in the scoring training materials each time an item is administered.

This repeated use of NECAP-approved sets of mid-range score point exemplars helps ensure that Readers follow established guidelines each time a particular NECAP item is scored.

4.1.2.3 Reader Recruitment and Qualifications

For scoring the 2008–09 NECAP, Measured Progress actively sought a diverse scoring pool representative of the population of the three NECAP states. The broad range of Reader backgrounds included scientists, editors, business professionals, authors, teachers, graduate school students, and retired educators. Demographic information about Readers (e.g., gender, race, educational background) was electronically captured for reporting.

Although a four-year college degree or higher was preferred, Readers were required to have successfully completed at least two years of college and to have demonstrated knowledge of the content area they scored. This permitted recruiting Readers currently enrolled in a college program, a sector of the population with relatively recent exposure to current classroom practices and trends in their fields. In all cases, potential Readers were required to submit documentation (e.g., resume and/or transcripts) of their qualifications.

Table 4-2 summarizes the qualifications of the 2008–09 NECAP scoring leadership and Readers.

Table 4-2. 2008–09 NECAP: Qualifications of Scoring Leadership and Readers—Fall Administration

Scoring Responsibility	*Educational Credentials*				*Total*
	Doctorate	*Masters*	*Bachelors*	*Other*	
Scoring Leadership	5.3%	28.3%	57.2%	9.2%*	100.0%
Readers	4.2%	24.8%	56.2%	14.9%**	100.0%

Scoring Leadership = CRs, QACs, and SRs
*9 QAC/SRs had an Associate's degree and 5 at least 48+ college credits
**106 Readers had an Associate's degree and 69 at least 48+ college credits

Readers were either temporary Measured Progress employees or were secured through temporary employment agencies. All Readers were required to sign a nondisclosure/confidentiality agreement.

4.1.2.4 Methodology for Scoring Polytomous Items

Possible Score Points

The ranges of possible score points for the different polytomous items are shown in Table 4-3.

Table 4-3. 2008–09 NECAP: Possible Score Points for Polytomous Item Types

Polytomous Item Type	*Possible Score Point Range*
Writing Prompt (WP)	0–6
Constructed-Response (CR)	0–4
2-point Short-Answer (SA2)	0–2
1-point Short-Answer (SA1)	0–1
Non-Scorable Items	0

Non-Scorable Items.

Readers could designate a response as non-scorable for any of the following reasons:

- response was blank (no attempt to respond to the question)
- response was unreadable (illegible, too faint to see, or only partially legible/visible)—*see note below*
- response was written in the wrong location (seemed to be a legitimate answer to a different question)—*see note below*
- response was written in a language other than English
- response was completely off-task or off-topic
- response included an insufficient amount of material to make scoring possible
- response was an exact copy of the assignment
- response was incomprehensible
- student made a statement refusing to write a response to the question

Note**:** "Unreadable" and "wrong location" responses were eventually resolved, whenever possible, by researching the actual answer document (electronic copy or hard copy, as needed) to identify the correct location (in the answer document) or to more closely examine the response and then assign a score.

Scoring Procedures

Scoring procedures for polytomous items included both single-scoring and double-scoring. Single-scored items were scored by one Reader. Double-scored items were scored independently by two Readers, whose scores were tracked for "interrater agreement" (for further discussion of double-scoring and interrater agreement, see Section 4.1.2.7 and Appendix E).

4.1.2.5 Reader Training

Reader training began with an introduction of the onsite scoring staff and providing an overview of the NECAP program's purpose and goals (including discussion about the security, confidentiality, and proprietary nature of testing materials, scoring materials, and procedures).

Next, Readers thoroughly reviewed and discussed the scoring guides for each item to be scored. Each item-specific scoring guide included the item itself and score point descriptions.

Following review of an item's scoring guide, Readers reviewing or scoring the particular response set organized for that training: Anchor Sets, Training Sets, and Qualifying Sets. (These are defined below.)

During training, Readers could highlight or mark hard copies of the Anchor, Training, and first Qualifying Sets, even if all or part of the set was also presented online via computer.

Anchor Set

Readers first reviewed an Anchor Set of exemplary responses for an item. This is a set approved by the Reading, Writing, and Mathematics content specialists representing the three NECAP state departments of

education. Responses in Anchor Sets are typical, rather than unusual or uncommon; solid, rather than controversial or borderline; and true, meaning that they had scores that could not be changed by anyone other than the NECAP client and Measured Progress test development staff. Each contains one client-approved sample response per score point considered to be a mid-range exemplar. The set includes a second sample response if there is more than one plausible way to illustrate the merits and intent of a score point.

Responses were read aloud to the room of Readers in descending score order. Announcing the true score of each anchor response, trainers facilitated group discussion of responses in relation to score point descriptions to help Readers internalize the typical characteristics of score points.

This Anchor Set continued to serve as a reference for Readers as they went on to calibration, scoring, and recalibration activities for that item.

Training Set

Next, Readers practiced applying the scoring guide and anchors to responses in the Training Set. The Training Set typically included 10 to 15 student responses designed to help establish both the full score point range and the range of possible responses within each score point. The Training Set often included unusual responses that were less clear or solid (e.g., shorter than normal, employing atypical approaches, simultaneously containing very low and very high attributes, and written in ways difficult to decipher). Responses in the Training Set were presented in randomized score point order.

After Readers independently read and scored a Training Set response, trainers would poll Readers or use online training system reports to record their initial range of scores. Trainers then led group discussion of one or two responses, directing Reader attention to difficult scoring issues (e.g., the borderline between two score points). Trainers modeled for Readers throughout how to discuss scores by referring to the Anchor Set and to scoring guides.

Qualifying Set

After the Training Set had been completed, Readers were required to score responses accurately and reliably in Qualifying Sets assembled for CR items, WP items, and all SA2 items for grades 3 and 4 mathematics. The ten responses in each Qualifying Set were selected from an array of responses that clearly illustrated the range of score points for that item as reviewed and approved by the state specialists. Hard copies of the responses were also available to Readers so that they could make notes and refer back during the post-qualifying discussion.

To be eligible to live score one of the above items, Readers were required to demonstrate scoring accuracy rates of at least 80% exact agreement (i.e. to exactly match the pre-determined score on at least 8 of the 10 responses) and at least 90% exact-or-adjacent agreement (i.e., to exactly match or be within one score point of the pre-determined score on 9 or 10 of the 10 responses), except 70% and 90%, respectively, for 6-point WP responses. In other words, Readers were allowed 1 discrepant score (i.e., 1 score of 10 that was

more than one score point from the pre-determined score) provided they had at least 8 exact scores (7 for WP items).

To be eligible to score SA1 mathematics items (which were benchmarked "right" or "wrong"), and SA2 mathematics items for Grades 5–8 and 11, Readers had to qualify on at least one other mathematics item for that grade.

Retraining

Readers who did not pass the first Qualifying Set were retrained as a group by reviewing their performance with scoring leadership and then scoring a second Qualifying Set of responses. If they achieved the required accuracy rate on the second Qualifying Set, they were allowed to score operational responses.

Readers who did not achieve the required scoring accuracy rates on the second Qualifying Set were not allowed to score responses for that item. Instead, they either began training on a different item or were dismissed from scoring for that day.

4.1.2.6 Senior Quality Assurance Coordinator (QAC) and Senior Reader (SR) Training

QACs and select SRs were trained in a separate training session immediately prior to Reader training. In addition to discussing the items and their responses, QAC and SR training included greater detail on the client's rationale behind the score points than that covered with regular Readers in order to better equip QACs and SRs to handle questions from the latter.

4.1.2.7 Monitoring of Scoring Quality Control and Consistency

Readers were monitored for continued accuracy and consistency throughout the scoring process, using the following methods and tools (which are defined in this section):

- Embedded Committee-Reviewed Responses (CRRs)
- "Read-Behind" Procedures
- Double-Blind Scoring
- Recalibration Sets
- Scoring Reports

It should be noted that any Reader whose accuracy rate fell below the expected rate for a particular item and monitoring method was retrained on that item. Upon approval by the QAC or CR as appropriate (see below), the Reader was allowed to resume scoring. Readers who met or exceeded the expected accuracy rates continued scoring.

Furthermore, the accuracy rate required of a Reader to *qualify* to score responses live was more strict than that required to *continue* to score responses live. The reason for the difference is that an "exact score" in double-blind scoring requires that *two* Readers choose the same score for a response (in other words, is dependent on peer agreement), whereas an "exact score" in qualification requires only that a *single* Reader

match a score pre-established by scoring leadership. Reader accuracy and the use of multiple monitoring techniques is critical to the process of live scoring.

Embedded Committee-Reviewed Responses (CRRs)

Committee-Reviewed Responses (CRRs) are previously scored responses that are loaded ("embedded") by scoring leadership into iScore and distributed "blindly" to Readers during scoring. Embedded CRRs may be chosen either before or during scoring, and are inserted into the scoring queue so that they appear the same as all other live student responses.

Between 5 and 30 Embedded CRRs were distributed at random points throughout the first full day of scoring to ensure that Readers were sufficiently calibrated at the beginning of the scoring period. Individual Readers often received up to 20 Embedded CRRs within the first 100 responses scored and up to 10 additional responses within the next 100 responses scored on that first day of scoring.

Any Reader who fell below the required scoring accuracy rate was retrained before being allowed by the QAC to continue scoring. Once allowed to resume scoring, scoring leadership carefully monitored these Readers by increasing the number of Read-Behinds (defined in the next section).

Embedded CRRs were employed for all CR items. They were not used for WP items, because these are 100% Double-Blind scored (defined below). Embedded CRRs were also not used for SA2 items, because Read-Behind and Double-Blind techniques are more informative and cost effective for these items.

"Read-Behind" Procedures

Read-Behind scoring refers to scoring leadership (usually a SR) scoring a response after a Reader has already scored the response. The practice was applied to all open-ended item types.

Responses placed into the Read-Behind queue were randomly selected by scoring leadership; Readers were not aware which of their responses would be reviewed by their SR. The iScore system allowed 1, 2, or 3 responses per Reader to be placed into the Read-Behind queue at a time.

The SR entered his or her score into iScore before being allowed to see the Reader's score. The SR then compared the two scores, and the score of record (i.e., the reported score) was determined as follows:

- If there was exact agreement between the scores, no action was necessary; the regular Reader's score remained.
- If the scores were adjacent (i.e., differed by 1 point), the SR's score became the score of record. (A significant number of adjacent scores for a Reader triggered an individual scoring consultation with the SR, after which the QAC determined whether or when the Reader could resume scoring.)
- If the scores were discrepant (i.e., differed by more than 1 point), the SR's score became the score of record. (This triggered an individual consultation with the SR, after which the QAC determined whether or when the reader could resume scoring.)

Table 4-4 illustrates how scores were resolved by Read-Behind.

Table 4-4. 2008–09 NECAP: Examples of Read-Behind Scoring Resolutions

Reader Score	*QAC/SR Score*	*Score of Record*
4	4	4
4	3	3*
4	2	2*

* QAC/SR's score.

SRs were tasked with conducting, on average, five Read-Behinds per Reader throughout each half-scoring day; however, SRs conducted a proportionally greater number of Read-Behinds for Readers who seemed to be struggling to maintain, or who fell below, accuracy standards.

In addition to regular Read-Behinds, scoring leadership could choose to do Read-Behinds on any Reader at any point during the scoring process to gain an immediate, real-time "snapshot" of a Reader's accuracy.

Double-Blind Scoring

Double-Blind scoring refers to two Readers independently scoring a response without knowing it. The practice was applied to all open-ended item types. Table 4-5 shows by which method(s) both common and equating open-ended item responses for each operational test were scored.

Table 4-5. 2008–09 NECAP: Frequency of Double-Blind Scoring by Grade and Content

Grade	*Content*	*Responses Double-Blind Scored*
3–8, 11	Reading	2% randomly
3–8, 11	Mathematics	2% randomly
5, 8, 11	Writing (WP)	100%
5, 8	Writing (CR)	2% randomly
All	Unreadable responses	100%
All	Blank responses	100%

If there was a discrepancy (a difference greater than 1 score point) between Double-Blind scores, the response was placed into an arbitration queue. Arbitration responses were reviewed by scoring leadership (SR or QAC) without knowledge of the two Readers' scores. Scoring leadership assigned the final score. Appendix E provides the NECAP 2008–09 percentages of agreement between Readers for each common item for each grade and content area.

Scoring leadership consulted individually with any Reader whose scoring rate fell below the required accuracy rate, and the QAC determined whether or when the reader could resume scoring. Once the reader was allowed to resume scoring, scoring leadership carefully monitored the Reader by increasing the number of Read-Behinds .

Recalibration Sets

To determine whether reinstated Readers were still calibrated to the scoring standard, Readers were required to take an online Recalibration Set at the start and midpoint of the shift of their resumption of scoring.

Each Recalibration Set consisted of five responses representing the entire range of possible scores, including some with a score point of 0.

- Readers who were discrepant on 2 of 5 responses of the first Recalibration Set, or exact on 2 or fewer, were not permitted to score on that item that day and were either assigned to a different item or dismissed for the day.
- Readers who were discrepant on only 1 of 5 responses of the first Recalibration Set, and/or exact on 3, were retrained by their SR by discussing the Recalibration Set responses in terms of the score point descriptions and the original Anchor Set. After this retraining, such Readers began scoring operational responses under the proviso that the Reader's scores for that day and that item would be kept only if the Reader was exact on all 5 of 5 responses of the second Recalibration Set administered at the shift midpoint. The QAC determined whether or when these Readers had received enough retraining to resume scoring operational responses. Scoring leadership also carefully monitored the accuracy of such Readers by significantly increasing the number of their Read-Behinds.
- Readers who were not discrepant on any response of the first Recalibration Set, and exact on at least 4, was allowed to begin scoring operational responses immediately, under the proviso that this Recalibration performance would be combined with that of the second Recalibration Set administered at the shift midpoint.

The results of both Recalibration Sets were combined with the expectation that Readers would have achieved an overall 80%-exact and 90%-adjacent standard for that item for that day.

The Scoring Project Manager voided all scores posted on that item for that day by Readers who did not meet the accuracy requirement. Responses associated with voided scores were reset and redistributed to Readers with demonstrated accuracy for that item.

Recalibration Sets were employed for all CR items. They were not used for WP items, which were 100% Double-Blind scored. They were also not used for SA2 items, for which Read-Behind and Double-Blind techniques are more informative and cost effective.

Scoring Reports

Measured Progress's electronic scoring software, iScore, generated multiple reports that were used by scoring leadership to measure and monitor Readers for scoring accuracy, consistency, and productivity. These reports are further discussed in the following section.

4.1.2.8 Reports Generated During Scoring

Due to the complexity of the 2008–09 NECAP administration, computer-generated reports were necessary to ensure that

- overall group-level accuracy, consistency, and reliability of scoring were maintained at acceptable levels
- immediate, real-time individual Reader data were available to allow early intervention when necessary
- scoring schedules were maintained

The following reports were produced by iScore:

- **The Read-Behind Summary** showed the total number of Read-Behind responses for each Reader and noted the number and percentages of exact, adjacent, and discrepant scores with the SR/QAC. Scoring leadership could choose to generate this report by choosing options (such as "Today" "Past Week" and "Cumulative") from a pull-down menu. The report could also be filtered to select data for a particular item or across all items. This report was used in conjunction with other reports to determine whether a Reader's scores would be voided (i.e., sent back out to the floor to be rescored by other Readers). The benefit of this report is that it can reveal the degree to which an individual Reader agrees with their QAC or SR on how best to score live responses.
- **The Double-Blind Summary** showed the total number of double-scored responses of each Reader, and noted the number and percentages of exact, adjacent, and discrepant scores with second Readers. This report was used in conjunction with other reports to determine whether a Reader's scores should be voided (i.e., sent back out to the floor to be rescored by other Readers). The benefit of this report is that it can reveal the degree to which Readers are in agreement with each other about how best to score live responses.
- **The Accuracy Summary** combined Read-Behind and Double-Blind data, showing the total number for the Readers, their accuracy rates, and their score-point distributions.
- **The Embedded CRR Summary** showed, for each Reader (by item or across all items), the total number of responses scored, the number of CRRs scored, and the numbers and percentages of exact, adjacent, and discrepant scores with the CR. This report was used in conjunction with other

reports to determine whether a Reader's scores should be voided (i.e., sent back out to the floor to be rescored by other Readers). The benefit of this report is that it can reveal the degree to which an individual Reader agrees with their CR on how to best score live responses. Also, since embedded CRRs are administered during the first hours of scoring, this report can provide an early illustration of agreement between Readers and CRs.

- **The Qualification Statistics** listed each Reader by name and ID number, identified which Qualifying Set(s) they did and did not take and, for the ones taken, their pass rate. In addition to the pass rates of individuals, the report also showed numbers of Readers passing or failing a particular Qualifying Set. The QAC could use this report to determine how Readers within their scoring group performed on specific Qualifying Sets.

- **The Summary** showed the total number of student responses for an item, and identified, for the time at which the report was generated, the following:
    - the number of single and Double-Blind scorings that had been performed
    - the number of single and Double-Blind scorings yet to be performed

SECTION II—STATISTICAL AND PSYCHOMETRIC SUMMARIES

Chapter 5. CLASSICAL ITEM ANALYSES

As noted in Brown (1983), "A test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each item. Both the *Standards for Educational and Psychological Testing* (AERA, 1999) and the *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 1988) include standards for identifying high-quality items. Test items should assess only knowledge or skills that are identified as part of the domain being measured and should avoid assessing irrelevant factors. They should also be unambiguous, and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. Further, items must not unfairly disadvantage test takers from particular racial, ethnic, or gender groups.

Both qualitative and quantitative approaches were taken to ensure that 2008–09 NECAP items met these standards. Qualitative work was discussed in Chapter 2 ("Development and Test Design"). This chapter summarizes several types of quantitative analyses that were carried out on the 2008–09 NECAP items:

- classical item statistics
- Differential Item Functioning (subgroup differences in item performance)
- dimensionality analyses

All analyses presented are based on the statewide administration of the 2008–09 NECAP in fall 2008. The numbers of students who participated in the assessment at each grade level were about 30,400 in grade 3, 30,450 in grade 4, 32,150 in grade 5, 32,400 in grade 6, 33,000 in grade 7, 33,600 in grade 8, and 33,700 in grade 11.

5.1 Classical Difficulty and Discrimination Indices

All 2008–09 NECAP items were evaluated in terms of item difficulty according to standard Classical Test Theory (CTT) practice. The expected item difficulty, also known as the p-value, is the main index of item difficulty under the CTT framework. This index measures an item's difficulty by averaging the proportion of points received across all students who took the item. MC items were scored dichotomously (correct vs. incorrect), so for these items, the difficulty index is simply the proportion of students who correctly answered the item. CR items were scored polytomously, where a student can achieve a score of 0, 1, 2, 3, or 4. SA1 items were scored 0 or 1 and SA2 items 0, 1, or 2. By computing the CR and SA2 difficulty indices as the average proportion of points achieved, the indices for all item types are placed on a similar scale and ranges from 0.00 to 1.00. Although the p-value is traditionally called a measure of difficulty, it is

properly interpreted as an easiness index, because larger values indicate easier items. An index of 1.00 indicates that every student received full credit for the item; such items provide little information about differences in student ability, but do indicate knowledge or skills that have been mastered by most students. Similarly, an index of 0.00 indicates that no student received credit for the item; such items provide little information about differences in student ability, but may indicate knowledge or skills that have not yet been mastered by most students.

To provide best measurement, difficulty indices generally should range from near-chance performance (i.e., 0.25 for four-option, multiple-choice items; essentially 0.00 for open-response items) to 0.90. Indices outside this range indicate items that were either too difficult or too easy for the target population. Nonetheless, on a standards-referenced assessment such as NECAP, it may be appropriate to include some items with very low or very high item difficulty values to ensure sufficient content coverage.

Another desirable feature of an item is that the higher-achieving students perform better on the item than do lower-achieving students. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of an item. Within classical test theory, the item-test correlation is referred to as the item's discrimination, because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. For CR items, the item discrimination index used was the Pearson product-moment correlation; for dichotomous items (MC and SA1), this statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is –1.00 to +1.00 and their typical observed range is 0.20 to 0.60.

A discrimination index can be thought of as measuring how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, as a measure of construct consistency. In light of this interpretation, the selection of an appropriate criterion total score is crucial to the interpretation of the discrimination index. Because each form of the 2008–09 NECAP was constructed to be parallel in content, the criterion score selected for each item was the raw score total for each form. The analyses were conducted for each form separately.

Difficulty and discrimination indices (i.e., item level classical stats) for each item are provided in Appendix F. Item level statistics are summarized by form in Appendix G. The item difficulty and discrimination indices are within acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that students who performed well on individual items tended to perform well overall. There were a small number of items with near-zero discrimination indices, but none were negative.

Attempting to compare CTT difficulty indices across content areas or grade levels is a thorny proposition, because the statistics are population dependent. Such direct comparisons would require that either items or students were common across comparisons, and since that is not the case, it can not be determined whether differences in performance are due to real differences in student ability or differences in item difficulty or both. With this caveat in mind, it appears generally that students in higher grades found their

mathematics items more difficult than did students in lower grades. Mathematics items also appeared to be more difficult than items in other content areas across grades. Comparing difficulty indices across item types is also suspect, because MC items can be answered correctly by guessing. That the difficulty indices for the dichotomous items tended to be higher (i.e., the items are easier) than those for the polytomous items is not surprising. Similarly, discrimination indices for the polytomous items were larger than those for the dichotomous items due to the greater variability of the former (i.e., the partial credit these items allow) and the tendency for higher correlation coefficients given greater variances of the correlates

5.2 Differential Item Functioning (DIF)

The *Code of Fair Testing Practices in Education* (2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit, and actions should be taken to make certain that differences in performance are due to construct-relevant, rather than irrelevant, factors. The *Standards for Educational and Psychological Testing* (AERA, 1999) includes similar guidelines.

The standardization differential item functioning (DIF) procedure (Dorans and Kulick, 1986) is designed to identify items for which subgroups of interest perform differently, beyond the impact of differences in overall achievement. The DIF procedure calculates the difference in item performance for two groups of students (at a time) matched for achievement on the total test. Specifically, average item performance is calculated for students at every total score. Then an overall average is calculated, weighting the total score distribution so that it is the same for the two groups. The criterion (matching) score for 2008–09 NECAP was computed two ways. For common items, total score was the sum of scores on common items. Total score for matrix items was the sum of item scores on common and matrix items (excluding field-test items). Based on experience, this dual definition of criterion scores has worked well in identifying problematic common and matrix items.

When differential performance between two groups occurs on an item (i.e., a DIF index in the "low" or "high" categories, explained below), it may or may not be indicative of item bias. Course-taking patterns or differences in school curricula can lead to DIF but for construct-relevant reasons. On the other hand, if subgroup differences in performance could be traced to differential experience (such as geographical living conditions or access to technology), the inclusion of such items should be re-considered.

Computed DIF indices have a theoretical range from –1.0 to 1.0 for multiple-choice and short-answer items, and the index is adjusted to the same scale for constructed-response items. Dorans and Holland (1993) suggested that index values between –0.05 and 0.05 should be considered negligible. The preponderance of [contract] items fell within this range. Dorans and Holland further stated that items with values between –0.10 and –0.05 and between 0.05 and 0.10 (i.e., "low" DIF) should be inspected to ensure that no possible effect is

overlooked, and that items with values outside the [–0.10, 0.10] range (i.e., "high" DIF) are more unusual and should be examined very carefully[2].

For the 2008–09 NECAP tests, three subgroup comparisons were evaluated for DIF:

- Male versus female
- White versus African American
- White versus Hispanic

Other race/ethnicity groups (e.g., Asians) were not analyzed using DIF procedures, because limited sample sizes would have inflated type I error rates. Appendix H presents the number of items classified into each DIF category by test form and item type. Appendix I presents the number of items classified into each DIF category that favor males or females, by item type.

5.3 Dimensionality Analyses

Because tests are constructed with multiple content area categories and their associated knowledge and skills, the potential exists for a large number of dimensions being invoked beyond the common primary dimension. Generally, the reporting categories are highly correlated with each other; therefore, the primary dimension they share typically explains an overwhelming majority of variance in test scores. In fact, the presence of just such a dominant primary dimension is the psychometric assumption that provides the foundation for the unidimensional IRT models that are used for calibrating, linking, scaling, and equating the NECAP test forms.

The purpose of dimensionality analysis is to investigate whether violation of the assumption of test unidimensionality is statistically detectable and, if so, (a) the degree to which unidimensionality is violated and (b) the nature of the multidimensionality. Findings from dimensionality (DIM) analyses performed on the 2008–09 NECAP common items for mathematics, reading, and writing are reported below. (Note: Only common items were analyzed since they are used for score reporting.)

The DIM analyses were conducted using the nonparametric IRT-based methods DIMTEST (Stout, 1987; Stout, Froelich, & Gao, 2001) and DETECT (Zhang & Stout, 1999). Both of these methods use as their basic statistical building block the estimated average conditional covariances for item pairs. A conditional covariance is the covariance between two items conditioned on total score for the rest of the test, and the average conditional covariance is obtained by averaging over all possible conditioning scores. When a test is strictly unidimensional, all conditional covariances are expected to take on values within random noise of zero, indicating statistically independent item responses for examinees with equal expected scores. Non-zero conditional covariances are essentially violations of the principle of local independence; local dependence

[2] It should be pointed out here that DIF for items is evaluated initially at the time of field-testing. If an item displays high DIF, it is flagged for review by a Measured Progress content specialist. The content specialist consults with the Department to determine whether to include the flagged item in a future operational test administration.

would imply multidimensionality. Thus, non-random patterns of positive and negative conditional covariances are indicative of multidimensionality.

DIMTEST is a hypothesis-testing procedure for detecting violations of local independence. The data are first randomly divided into a training sample and a cross-validation sample. Then an exploratory analysis of the conditional covariances is conducted on the training sample data to find the cluster of items that displays the greatest evidence of local dependence. The cross-validation sample is then used to test whether the conditional covariances of the selected cluster of items display local dependence, conditioning on total score on the non-clustered items. The DIMTEST statistic follows a standard normal distribution under the null hypothesis of unidimensionality.

DETECT is an effect-size measure of multidimensionality. As with DIMTEST, the data are first randomly divided into a training sample and a cross-validation sample (these samples are drawn independently of those used with DIMTEST). The training sample is used to find a set of mutually exclusive and collectively exhaustive clusters of items that best fit a systematic pattern of positive conditional covariances for pairs of items from the same cluster and negative conditional covariances from different clusters. Next, the clusters from the training sample are used with the cross-validation sample data to average the conditional covariances: within-cluster conditional covariances are summed, from this sum the between-cluster conditional covariances are subtracted, this difference is divided by the total number of item pairs, and this average is multiplied by 100 to yield an index of the average violation of local independence for an item pair. DETECT values less than 0.20 indicate very weak multidimensionality (or near unidimensionality) values of 0.20 to 0.40 weak to moderate multidimensionality; values of 0.40 to 1.00 moderate to strong multidimensionality and values greater than 1.0 very strong multidimensionality.

DIMTEST and DETECT were applied to the 2008–09 NECAP data. For each grade and content area, the data were split into a training sample and a cross-validation sample. Every grade content combination had at least 30,000 student examinees. Because DIMTEST is limited to using 24,000 students, the training and cross-validation samples were constructed by randomly sampling 12,000 students each from the total sample. DETECT has an upper limit of 50,000 students, so all students were selected randomly into training and cross-validation samples of 15,000 or more students. DIMTEST was then applied to every grade content, and DETECT was applied in order to estimate the effect size of the multidimensionality to those datasets for which the DIMTEST null hypothesis was rejected.

The DIMTEST null hypothesis was strongly rejected for every data set ($p = .0002$ for writing grade 5 and $p < 0.00005$ for all other grade contents). Because strict unidimensionality is an idealization that almost never holds exactly for a given data set, these DIMTEST results were not surprising. Indeed, because of the very large NECAP sample sizes, DIMTEST would be sensitive to even quite small violations of unidimensionality. It was therefore important to use DETECT to estimate the effect size of the violations of local independence. Table 5-1 displays the multidimensional effect size estimates from DETECT for the 2008–09 NECAP along with the estimates from 2007–08 for comparison purposes.

Table 5-1. 2008–09 NECAP: Multidimensionality Effect Sizes by Grade and Content (2007–08 Data Also Shown for Comparison)

Grade	Content	Multidimensionality Effect Size 2007–08*	2008–09
3	Mathematics	0.16	0.12
	Reading	0.13	0.22
4	Mathematics	0.17	0.14
	Reading	0.24	0.35
5	Mathematics	0.12	0.21
	Reading	0.24	0.18
	Writing	0.21	0.20
6	Mathematics	0.11	0.17
	Reading	0.19	0.24
7	Mathematics	0.14	0.19
	Reading	0.28	0.23
8	Mathematics	0.20	0.15
	Reading	0.24	0.20
	Writing	0.18	0.18
11	Mathematics	0.16	0.17
	Reading	0.23	0.31

*(2007–08 Data Also Shown for Comparison)

All DETECT values indicated very weak to weak multidimensionality, except grade 4 reading, whose value of 0.35 is on the borderline between weak and moderate. Reading tests tended to show slightly greater multidimensionality than did mathematics tests. The average DETECT value for reading was 0.25 and 0.16 for mathematics (both of these values are towards the weak end of the 0.20 to 0.40 range). It can be seen in the table that the 2008–09 effect sizes for mathematics and reading were very similar to those from 2007–08 (0.22 and 0.15 for reading and mathematics, respectively). DETECT values for writing were also nearly the same in the two years.

Also investigated was how DETECT divided the 2008–09 NECAP grade content tests into clusters to see if there were any discernable patterns with respect to item type. Mathematics clusters showed no discernable patterns. For both reading and writing, however, there was a strong tendency for the MC items to cluster separately from the remaining items. Despite this multidimensionality between the item types in reading and writing, the effect sizes were not strong enough to warrant further investigation. (Note: These trends and conclusions are exactly the same as reported for the 2007–08 tests.)

Chapter 6. IRT SCALING AND EQUATING

6.1 Item Response Theory Scaling

All NECAP items were calibrated using Item Response Theory (IRT). IRT uses mathematical models to define a relationship between an unobserved measure of student performance, usually referred to as theta (θ), and the probability (p) of getting a dichotomous item correct or of getting a particular score on a polytomous item. In IRT, it is assumed that all items are independent measures of the same construct (i.e., of the same θ). Another way to think of θ is as a mathematical representation of the latent trait of interest. Several common IRT models are used to specify the relationship between θ and p (Hambleton and van der Linden, 1997; Hambleton and Swaminathan, 1985). The process of determining the specific mathematical relationship between θ and p is called item calibration. After items are calibrated, they are defined by a set of parameters that specify a nonlinear, monotonically increasing relationship between θ and p. Once the item parameters are known, an estimate of θ for each student can be calculated. This estimate $\hat{\theta}$, is considered to be an estimate of the student's true score or a general representation of student performance. It has characteristics that may be preferable to those of raw scores for equating purposes.

For NECAP 2008–09, the three-parameter logistic (3PL) model was used for dichotomous items and the graded-response model (GRM) was used for polytomous items. The 3PL model for dichotomous items can be defined as follows (note that for SA1 items, the c parameter is set to zero, thus the model becomes 2PL):

$$P_i\left(1\middle|\theta_j,\xi_i\right)=c_i+\left(1-c_i\right)\frac{\exp\left[Da_i\left(\theta_j-b_i\right)\right]}{1+\exp\left[Da_i\left(\theta_j-b_i\right)\right]}$$

Where:
i indexes the items,
j indexes students,
a represents item discrimination,
b represents item difficulty,
c is the pseudo-guessing parameter,
ξ_i represents the set of item parameters (a, b, and c), and
D is a normalizing constant equal to 1.701.

In the GRM for polytomous items, an item is scored in k+1 graded categories that can be viewed as a set of k dichotomies. At each point of dichotomization (i.e., at each threshold), a two-parameter model can be used. This implies that a polytomous item with k+1 categories can be characterized by k item category threshold curves (ICTC) of the two-parameter logistic form:

$$P_{ik}^{*}\left(1\middle|\theta_{j,}a_i,b_i,d_{ik}\right)=\frac{\exp\left[Da_i\left(\theta_j-b_i+d_{ik}\right)\right]}{1+\exp\left[Da_i\left(\theta_j-b_i+d_{ik}\right)\right]}$$

Where:

i indexes the items,

j indexes students,

k indexes threshold,

a represents item discrimination,

b represents item difficulty,

d represents threshold, and

D is a normalizing constant equal to 1.701.

After computing *k* ICTCs in the GRM, *k*+1 item category characteristic curves (ICCCs) are derived by subtracting adjacent ICTCs:

$$P_{ik}(1\,|\,\theta_j)=P_{i(k-1)}^{*}(1\,|\,\theta_j)-P_{ik}^{*}(1\,|\,\theta_j)$$

Where:

P_{ik} represents the probability that the score on item *i* falls in category *k*, and

P_{ik}^{*} represents the probability that the score on item *i* falls above the threshold *k*

($P_{i0}^{*}=1$ and $P_{i(m+1)}^{*}=0$).

The GRM is also commonly expressed as:

$$P_{ik}\left(k\middle|\theta_j,\xi_i\right)=\frac{\exp\left[Da_i\left(\theta_j-b_i+d_k\right)\right]}{1+\exp\left[Da_i\left(\theta_j-b_i+d_k\right)\right]}-\frac{\exp\left[Da_i\left(\theta_j-b_i+d_{k+1}\right)\right]}{1+\exp\left[Da_i\left(\theta_j-b_i+d_{k+1}\right)\right]}$$

where, ξ_i represents the set of item parameters for item *i*.

Finally, the ICC for polytomous items is computed as a weighted sum of ICCCs, where each ICCC is weighted by a score assigned to a corresponding category.

$$P_i(1\,|\,\theta_j)=\sum_{k}^{m+1}w_{ik}P_{ik}(1\,|\,\theta_j)$$

For more information about item calibration and determination, the reader is referred to Lord and Novick (1968), Hambleton and Swaminathan (1985), or Baker and Kim (2004).

6.2 Item Response Theory Analyses

The previous section introduced IRT and gave a thorough description of the topic. It was discussed there that all 2008–09 NECAP items were calibrated using IRT and that the calibrated item parameters were ultimately used to scale both the items and students onto a common framework. The results of those analyses are presented in Appendix J.

The tables in Appendix J give the IRT item parameters of all common items on the 2007–08 NECAP tests by grade and content area. Accompanying the parameter tables are graphs of the corresponding Test Characteristic Curves (TCCs) and Test Information Functions (TIFs), which are defined below.

TCCs display the expected (average) raw score associated with each θ_j value between –4.0 and 4.0. Mathematically, the TCC is computed by summing the ICCs of all items that contribute to the raw score. Using the notation introduced in Section 6.1, the expected raw score at a given value of θ_j is

$$E(X \mid \theta_j) = \sum_{i=1}^{n} P_i\left(1 \middle| \theta_j\right),$$

where

i indexes the items (and n is the number of items contributing to the raw score),

j indexes students (here, θ_j runs from –4 to 4)

$E(X \mid \theta_j)$ is the expected raw score for a student of ability θ_j.

The expected raw score monotonically increases with θ_j, consistent with the notion that students of high ability tend to earn higher raw scores than do students of low ability. Most TCCs are "S-shaped" flatter at the ends of the distribution and steeper in the middle.

The TIF displays the amount of statistical information that the test provides at each value of θ_j. Information functions depict test precision across the entire latent trait continuum. There is an inverse relationship between the information of a test and its standard error of measurement (SEM). For long tests, the SEM at a given θ_j is approximately equal to the inverse of the square root of the statistical information at θ_j (Hambleton, Swaminathan, & Rogers, 1991), as follows:

$$SEM(\theta_j) = \frac{1}{\sqrt{I(\theta_j)}}$$

Compared to the tails, TIFs are often higher near the middle of the θ distribution where most students are located and where most items are sensitive by design.

6.3 Equating

The purpose of equating is to ensure that scores obtained from different forms of a test are equivalent to each other. Equating may be used if multiple test forms are administered in the same year, as well as to equate one year's forms to those given in the previous year. Equating ensures that students are not given an unfair advantage or disadvantage because the test form they took is easier or harder than those taken by other students.

The 2008–09 administration of NECAP used a raw score-to-theta equating procedure in which test forms are equated every year to the theta scale of the reference test forms. (In the case of NECAP, the reference forms are those from the 2005–06 administration for grades 3 through 8 and 2007–08 for grade 11.) This is accomplished through the chained linking design, in which every new form is equated back to the theta scale of the previous year's test form. It can therefore be assumed that the theta scale of every new test form is the same as the theta scale of the reference form, since this is where the chain originated.

Students who took the equating items on the 2008–09 and 2007–08 NECAP tests are not equivalent groups. IRT is particularly useful for equating scenarios that involve nonequivalent groups (Allen and Yen, 1979). Equating for NECAP uses the *anchor-test-nonequivalent-groups* design described by Petersen, Kolen, and Hoover (1989). In this equating design, no assumption is made about the equivalence of the examinee groups taking different test forms (that is, naturally occurring groups are assumed). Comparability is instead evaluated through utilizing a set of anchor items (also called equating items). The NECAP uses an *external* anchor test design, which means that the equating items are not counted toward students' test scores. However, the equating items are designed to mirror the common test in terms of item types and distribution of emphasis. Subsets of the equating items are matrixed across forms.

Item parameter estimates for 2008–09 were placed on the 2007–08 scale by using the method of Stocking and Lord (1983), which is based on the IRT principle of item parameter invariance. According to this principle, the equating items for both the 2007–08 and 2008–09 NECAP tests should have the same item parameters. After the item parameters for each 2007–08 NECAP mathematics and reading test were estimated using PARSCALE, as described earlier, he Stocking and Lord method was employed to find the linear transformation (slope and intercept) that adjusted the equating items' parameter estimates such that the 2008–09 TCC was as close as possible to that of 2007–08. The transformation constants are presented in Table 6-1. It should be noted that grades 5 and 8 writing were excepted from the equating process; writing test forms are equated through the scoring rubric.

Table 6-1. 2008-09 NECAP: Stocking & Lord Transformation Constants

Grade	*Content*	*A*	*B*
3	Mathematics	1.005479	0.087349
4	Mathematics	1.052071	0.069017
5	Mathematics	0.981083	0.180927
6	Mathematics	1.067933	0.110943
7	Mathematics	1.038854	0.179566
8	Mathematics	0.948991	0.176795
11	Mathematics	1.079684	0.094488
3	Reading	0.950923	0.098595
4	Reading	1.013818	0.253671
5	Reading	0.970414	0.125344
6	Reading	1.039932	0.001675
7	Reading	1.044422	0.135001
8	Reading	1.098949	-0.035390
11	Reading	1.001144	0.155915

A = Slope, B = Intercept

The next administration of NECAP (2009–10) will be scaled to the 2008–09 administration by the same equating method described above.

6.4 Equating Results

An Equating Report was submitted to the NECAP state testing directors for their approval prior to production of student reports. Various elements from the Equating Report are presented throughout this technical report and its appendices.

In addition to the equating and scaling activities described in the previous section (IRT calibrations, identification of equating items, and execution of the Stocking and Lord equating procedure) various quality control procedures were implemented within the Psychometrics Department at Measured Progress and reviewed with the NECAP state testing directors and NECAP Technical Advisory Committee. A variety of quality control activities were undertaken during the IRT calibration, equating, and scaling, and various results are presented throughout this report.

The number of Newton cycles required for convergence for each grade and content area during the IRT analysis can be found in Table 6-2. The number of cycles required in order for the solution to converge fell within acceptable ranges (e.g. below 150 cycles)

Table 6-2. 2008-09 NECAP: Number of Newton Cycles required for Convergence

Grade	*Content*	*Cycles*
3	Mathematics	32
4		61
5		38
6		66
7		55
8		64
11		84
3	Reading	45
4		46
5		54
6		49
7		49
8		49
11		47

The number of items that required intervention during the IRT analysis are presented in Table 6-3. The number of items that required intervention was very typical across the various grades and content areas. Appendix K presents the results from the Delta analysis. This procedure was used to evaluate adequacy of equating items, and the discard status presented in the appendix indicates whether or not the item was used in equating. Also presented in Appendix K are the results from the rescore analysis. With this analysis, 200 random papers from the previous year were interspersed with this year's papers to evaluate scorer consistency from on year to the next. Only items with effect sizes greater than 0.80 were automatically excluded as equating items.

6.5 Achievement Standards

NECAP standards to establish achievement-level cut scores in reading, mathematics, and writing for grades 3 through 8 were set in January 2006 and for grade 11 in January 2008. The standard-setting meetings and results were discussed in the technical reports of those years. As alluded to in the discussion of equating above, the respective NECAP reporting scales were established during those base years, and the forms serve as the reference for subsequent equating. The θ-metric cut scores that emerged from the standard-setting meetings will remain fixed throughout the assessment program unless standards are reset for any reason.

6.6 Reported Scale Scores

6.6.1 Description of Scale

Because the θ scale used in IRT calibrations is not readily understood by most stakeholders, reporting scales were developed for the NECAP tests. The reporting scales are simple linear transformations of the underlying θ scale. The scales were developed such that they ranged from X00 through X80, where X is grade level. In other words, grade 3 scaled scores ranged from 300 to 380, grade 4 from 400 through 480, and so forth through grade 11, where scores ranged from 1100 through 1180. The lowest scaled score in the *Proficient* range was set at "X40" for each grade level. For example, to be classified in the *Proficient* achievement level or above, a minimum scaled score of 340 was required at grade 3, 440 at grade 4, and so forth.

By providing information that is more specific about the position of a student's results, scaled scores supplement achievement-level. School- and district-level scaled scores are calculated by computing the average of student-level scaled scores. Students' raw scores (i.e., total number of points) on the 2008–09 NECAP tests were translated to scaled scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another. In the same way that a given temperature can be expressed on either Fahrenheit or Celsius scales, or the same distance can be expressed in either miles or kilometers, student scores on the 2008–09 NECAP tests can be expressed in raw or scaled scores.

In Figure 6-1, two-way arrows depict how raw scores (the vertical axis) map through the S-shaped TCC to corresponding scores on the θ scale (horizontal axis), which in turn map directly to scaled scores.

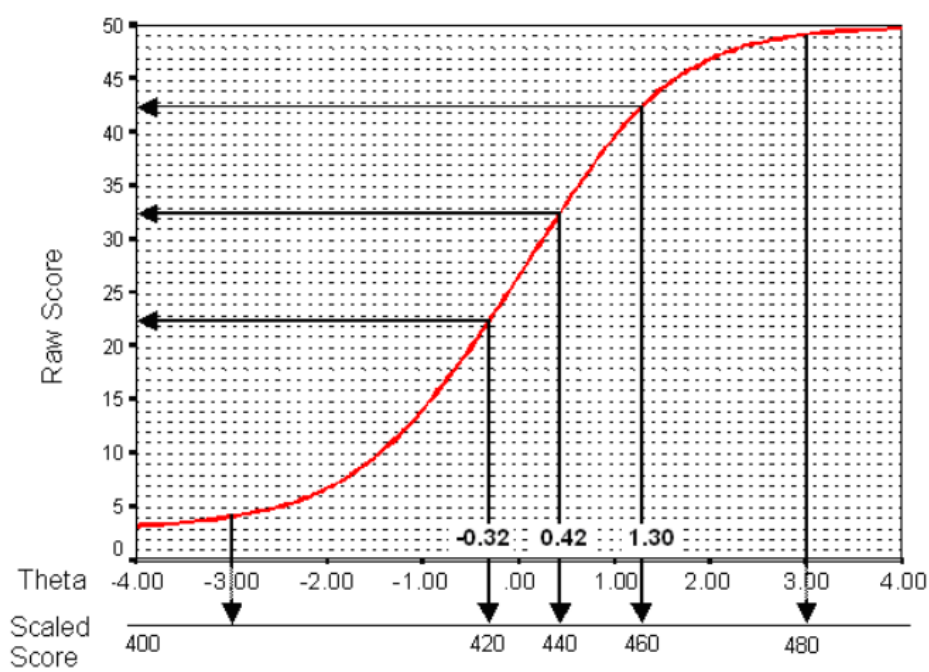

Figure 6-1. Conceptual Illustration of Raw Score to Theta Scaled Score Transformation Using a Test Characteristic Curve

It is important to note that converting from raw scores to scaled scores does not change students' achievement-level classifications. Given the relative simplicity of raw scores, it is fair to question why scaled scores for NECAP are reported instead of raw scores. Scaled scores make consistent the reporting of results. To illustrate, standard setting typically results in different *raw* cut scores across content areas. The raw cut score between *Partially Proficient* and *Proficient* could be, for example, 35 in mathematics but 33 in reading, yet both of these raw scores would be transformed to scaled scores of X40. It is this uniformity across scale scores that facilitates the understanding of student performance. The psychometric advantage of scaled scores over raw scores comes from their being *linear* transformations of θ. Since the θ scale is used for equating, scaled scores are comparable from one year to the next. Raw scores are not

6.6.2 Calculations

The scaled scores are obtained by a simple translation of ability estimates ($\hat{\theta}$) using the linear relationship between threshold values on the θ metric and their equivalent values on the scaled score metric. Students' ability estimates are based on their raw scores and are found by mapping through the TCC. Scaled scores are calculated using the linear equation

$$SS = m\hat{\theta} + b$$

where
m is the slope and
b is the intercept

A separate linear transformation is used for each grade content combination. For NECAP tests, each line is determined by fixing both the *Partially Proficient/Proficient* cut score and the bottom of the scale; that is, the X40 value and the X00 value (e.g., 340 and 300, respectively, for grade 3). The lowest scaled score is at a location on the θ scale beyond the scaling of all the items across the various grade content combinations. To determine its location, a chance raw score (approximately equal to a student's expected performance by guessing), and a raw score of 0, are both mapped to a θ value of –4.0. At the other extreme, the maximum possible raw score is assigned the scaled score of X80 (e.g., 380 in the case of grade 3).

Because only the *Partially Proficient/Proficient* cut score is fixed within the θ scaled-score space, the cut scores between *Substantially Below Proficient* and *Partially Proficient* (SBP/PP) and between *Proficient* and *Proficient with Distinction* (P/PWD) vary across the grade/content combinations.

Table 6-3 presents the scaled score cuts for each grade content combination (i.e., the minimum scaled score for getting into the next achievement level). It is important to repeat that the values in Table 6-3 do not change from year to year, because the cut scores along the θ scale do not change. Also, in a given year it may not be possible to attain a particular scaled score, but the scaled score cuts will remain the same.

Table 6-3.2008–09 NECAP: Scaled Score Cuts and Minimum and Maximum Scores by Grade and Content

Grade	*Content*	*Min*	*Scaled Score Cuts*			*Max*
			SBP/PP	PP/P	P/PWD	
3	Mathematics	300	332	340	353	380
4		400	431	440	455	480
5		500	533	540	554	580
6		600	633	640	653	680
7		700	734	740	752	780
8		800	834	840	852	880
11		1100	1134	1140	1152	1180
3	Reading	300	331	340	357	380
4		400	431	440	456	480
5		500	530	540	556	580
6		600	629	640	659	680
7		700	729	740	760	780
8		800	828	840	859	880
11		1100	1130	1140	1154	1180
5	Writing*	500	528	540	555	580
8		800	829	840	857	880

SBP = *Substantially Below Proficient*; PP = *Partially Proficient* ; P = *Proficient*; PWD = *Proficient with Distinction* *Scaled scores are not produced for grade 11 writing

Table 6-4 shows the cut scores on θ and the slope and intercept terms used to calculate the scaled scores. Note that the values in Table 6-4 will not change unless the standards are reset.

Table 6-4. 2008–09 NECAP: Cut scores (on θ Metric), Intercept, and Slope by Grade and Content

Grade	*Content*	*θ Cuts*			*Intercept*	*Slope*
		SBP/PP	PP/P	P/PWD		
3	Mathematics	–1.0381	–0.2685	0.9704	342.8782	10.7195
4		–1.1504	–0.3779	0.9493	444.1727	11.0432
5		–0.9279	–0.2846	1.0313	543.0634	10.7659
6		–0.8743	–0.2237	1.0343	642.3690	10.5922
7		–0.7080	–0.0787	1.0995	740.8028	10.2007
8		–0.6444	–0.0286	1.1178	840.2881	10.0720
11		–0.1169	0.6190	2.0586	1134.640	8.6600
3	Reading	–1.3229	–0.4970	1.0307	345.6751	11.4188
4		–1.1730	–0.3142	1.1473	443.4098	10.8525
5		–1.3355	–0.4276	1.0404	544.7878	11.1970
6		–1.4780	–0.5180	1.1255	645.9499	11.4875
7		–1.4833	–0.5223	1.2058	746.0074	11.5019
8		–1.5251	–0.5224	1.1344	846.0087	11.5022
11		–1.2071	–0.3099	1.0038	1143.3600	10.8399
5	Writing	–1.2008	–0.0232	1.5163	540.2334	10.0583
8		–1.0674	–0.0914	1.8230	839.1064	9.7766

SBP = Substantially Below Proficient; PP = Partially Proficient ; P = Proficient; PWD = Proficient with Distinction

Table 6-5 shows the raw scores associated with the cut scores for each performance level by grade and content area. In order to evaluated changes in test difficulty, the results are compared to the previous year's results which are also reflected in Table 6-5.

Table 6-5. 2008-09 NECAP: Raw Scores on θ for each Performance Level, by Grade and Content

Grade & Content		*Year 1*				*Year 2*			
		SbP/PP	PP/P	P/PwD	Max	SbP/PP	PP/P	P/PwD	Max
3	Mathematics	29	40	56	65	28	39	55	65
4		26	38	55	65	28	39	55	65
5		21	28	47	66	23	31	52	66
6		20	28	47	66	19	27	47	66
7		18	27	45	66	19	27	47	66
8		17	27	48	66	20	29	49	66
11		18	29	53	64	21	33	56	64
3	Reading	22	31	44	52	21	30	44	52
4		22	31	43	52	21	30	42	52
5		18	27	38	52	17	26	38	52
6		20	28	39	52	20	30	42	52
7		19	28	42	52	21	30	42	52
8		23	33	44	52	22	31	42	52
11		19	28	39	52	22	31	42	52

SBP = Substantially Below Proficient; PP = Partially Proficient ; P = Proficient; PWD = Proficient with Distinction
Note 1: The values presented are not the cut scores per se. The cut scores are defined on the θ metric and do not change from year to year. The values in this table represent the raw scores associated with the cut scores, and these values are found via a TCC mapping.

Appendix L contains raw score-to-scaled score Look-up tables. These are the actual tables that were used to determine student scaled scores, error bands, and achievement levels.

6.6.3 Distributions

The State specific performance level distribution results for each grade and content area are presented in Tables 6-6 through 6-8. As with Table 6-5, we have included historical results for purposes of comparison.

Appendix M contains percentages and cumulative percentages across the scale score ranges. These distributions were calculated using the sparse data matrix files that were used in the IRT calibrations.

Table 6-6.2008-09 NECAP: Preliminary Performance Level Distribution Results—New Hampshire

Math	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	22	21	20	24	18	15	23	20	18	23	22	21	22	19	17	18	15	14	2	2
P	49	52	49	49	50	51	50	49	49	46	46	46	44	44	45	47	43	43	30	26
BP	18	17	20	16	21	21	14	16	18	16	17	16	18	21	18	19	22	19	30	28
SBP	10	10	10	10	11	13	13	16	15	15	16	16	17	16	20	16	19	24	38	45
Ave. SS	345	345	345	446	445	444	546	545	544	645	644	644	744	743	742	843	841	841	1135	1134
P+ (delta)	71 (-2)	73 (4)	69	73 (5)	68 (2)	66	73 (4)	69 (2)	67	69 (1)	68 (1)	67	66 (3)	63 (1)	62	65 (7)	58 (1)	57	32 (4)	28

Reading	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	23	14	18	23	19	17	16	18	16	14	15	14	17	15	9	14	13	10	18	17
P	55	63	57	52	55	55	60	53	56	60	58	58	60	60	58	56	54	56	54	50
BP	14	14	15	17	17	19	18	21	18	19	19	19	17	18	25	22	23	24	18	22
SBP	8	9	10	8	9	9	6	8	10	8	8	9	6	7	8	8	10	10	10	12
Ave. SS	348	347	347	447	446	446	547	546	545	647	646	646	748	747	745	846	845	844	1145	1144
P+ (delta)	78 (1)	77 (2)	75	75 (1)	74 (2)	72	76 (5)	71 (-1)	72	74 (1)	73 (1)	72	77 (2)	75 (8)	67	70 (3)	67 (1)	66	72 (5)	67

Writing	Grade 5			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	30	16	13	9	7	8	3	3
P	30	36	37	42	36	34	36	30
BP	26	30	29	35	40	38	50	49
SBP	14	18	21	14	17	20	11	18
Ave. SS	546	541	540	841	839	839	6	5.5
P+ (delta)	60 (8)	52 (2)	50	51 (8)	43 (1)	42	39 (6)	33

Percentages may not sum to 100 due to rounding.

Table 6-7.2008-09 NECAP: Preliminary Performance Level Distribution Results—Rhode Island

Math	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	16	15	16	19	12	10	15	13	13	15	14	13	14	12	12	13	11	11	1	1
P	44	45	40	44	42	44	45	44	44	40	40	41	38	37	39	40	37	37	25	21
BP	21	21	24	20	25	24	17	18	21	20	20	21	20	24	20	22	25	19	28	27
SBP	19	19	20	17	21	22	23	25	21	25	26	25	28	27	28	25	27	33	45	51
Ave. SS	342	342	341	443	441	440	542	541	541	641	640	640	740	739	739	840	839	838	1134	1132
P+ (delta)	60 (0)	60 (4)	56	63 (9)	54 (0)	54	60 (3)	57 (0)	57	55 (1)	54 (0)	54	52 (3)	49 (-2)	51	53 (5)	48 (0)	48	26 (4)	22

Reading	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	18	11	13	20	16	14	15	18	15	13	11	11	15	14	10	14	13	10	18	16
P	51	57	52	48	48	49	53	48	50	55	51	53	56	53	49	51	48	49	51	45
BP	18	18	20	20	21	23	21	23	21	21	24	23	20	21	28	24	24	26	20	24
SBP	13	14	15	12	15	13	11	11	14	11	14	13	9	12	13	12	15	16	11	14
Ave. SS	345	344	344	445	444	443	545	545	544	645	643	644	747	745	743	845	844	842	1145	1143
P+ (delta)	69 (1)	68 (3)	65	68 (4)	64 (1)	63	68 (2)	66 (1)	65	68 (6)	62 (-2)	64	71 (4)	67 (8)	59	65 (4)	61 (2)	59	69 (8)	61

Writing	Grade 5			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	28	18	15	8	7	9	4	3
P	30	35	36	40	34	33	38	34
BP	26	28	27	35	40	35	50	49
SBP	16	18	22	16	19	22	9	14
Ave. SS	545	542	540	840	838	838	6.1	5.7
P+ (delta)	58 (5)	53 (2)	51	48 (7)	41 (-1)	42	42 (5)	37

Percentages may not sum to 100 due to rounding.

Table 6-8.2008-09 NECAP: Preliminary Performance Level Distribution Results—Vermont

Math	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	19	19	24	23	17	16	20	18	19	22	22	20	21	20	20	21	18	17	2	2
P	47	49	44	46	45	48	48	46	46	44	44	44	42	40	42	42	41	42	33	28
BP	19	18	20	18	22	21	15	17	18	17	17	16	18	20	17	19	21	18	28	28
SBP	16	14	13	14	16	16	17	20	17	18	18	19	20	20	20	19	20	23	37	42
Ave. SS	343	344	344	445	443	443	544	543	543	644	644	643	743	742	742	842	841	841	1135	1134
P+ (delta)	66 (-2)	68 (0)	68	69 (7)	62 (-2)	64	68 (4)	64 (-1)	65	66 (0)	66 (2)	64	63 (3)	60 (-2)	62	63 (4)	59 (0)	59	35 (5)	30

Reading	Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	21	14	17	21	19	17	18	19	18	15	17	14	20	16	13	17	18	12	22	20
P	50	57	54	48	49	51	52	49	51	56	55	55	56	55	54	52	51	53	50	48
BP	16	15	17	19	19	20	21	21	18	19	19	20	17	18	24	21	21	24	18	20
SBP	13	13	13	12	12	12	9	11	12	10	10	11	8	10	10	10	10	11	10	12
Ave. SS	346	345	346	445	445	444	546	545	545	646	646	645	749	747	746	846	847	844	1146	1144
P+ (delta)	71 (0)	71 (0)	71	69 (1)	68 (0)	68	70 (2)	68 (-1)	69	71 (-1)	72 (3)	69	76 (5)	71 (4)	67	69 (0)	69 (4)	65	72 (4)	68

Writing	Grade 5			Grade 8			Grade 11	
	08-09	07-08	06-07	08-09	07-08	06-07	08-09	07-08
PwD	27	16	14	12	10	11	4	4
P	28	32	36	42	37	36	38	35
BP	26	28	26	31	36	33	46	46
SBP	19	23	25	15	17	20	11	15
Ave. SS	544	540	540	842	840	840	6.1	5.8
P+ (delta)	55 (7)	48 (-2)	50	54 (7)	47 (0)	47	42 (3)	39

Chapter 7. RELIABILITY

Although an individual item's performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way items function together and complement one another. Tests that function well provide a dependable assessment of the student's level of ability. Unfortunately, no test can do this perfectly. A variety of factors can contribute to a given student's score being either higher or lower than his or her true ability. For example, a student may mis-read an item, or mistakenly fill in the wrong bubble when he or she knew the answer. Collectively, extraneous factors that impact a student's score are referred to as measurement error. Any assessment includes some amount of measurement error; that is, no measurement is perfect. This is true of all academic assessments—some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. When tests have a high amount of measurement error, student scores are very unstable. Students with high ability may get low scores or vice versa. Consequently, one cannot reliably measure a student's true level of ability with such a test. Assessments that have less measurement error (i.e., errors made are small on average and student scores on such a test will consistently represent their ability) are described as reliable.

There are a number of ways to estimate an assessment's reliability. One possible approach is to give the same test to the same students at two different points in time. If students receive the same scores on each test, then the extraneous factors affecting performance are small and the test is reliable. (This is referred to as test-retest reliability.) A potential problem with this approach is that students may remember items from the first administration or may have gained (or lost) knowledge or skills in the interim between the two administrations. A solution to the "remembering items" problem is to give a different, but parallel test at the second administration. If student scores on each test correlate highly the test is considered reliable. (This is known as alternate forms reliability, because an alternate form of the test is used in each administration.) This approach, however, does not address the problem that students may have gained (or lost) knowledge or skills in the interim between the two administrations. In addition, the practical challenges of developing and administering parallel forms generally preclude the use of parallel forms reliability indices. One way to address the latter problems is to split the test in half and then correlate students' scores on the two half-tests; this in effect treats each half-test as a complete test. By doing this, the problems associated with an intervening time interval, and of creating and administering two parallel forms of the test, are alleviated. This is known as a split-half estimate of reliability. If the two half-test scores correlate highly, items on the two half-tests must be measuring very similar knowledge or skills. This is evidence that the items complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test halves will result in a different correlation. Another problem with the split-half method of calculating reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter

test is less reliable than a longer test. Cronbach (1951) provided a statistic, alpha (α), which eliminates the problem of the split-half method by comparing individual item variances to total test variance. Cronbach's α was used to assess the reliability of the 2008–09 NECAP tests:

$$\alpha \equiv \frac{n}{n-1}\left[1-\frac{\sum_{i=1}^{n}\sigma^{2}{}_{(Y_i)}}{\sigma_x^2}\right]$$

Where

i indexes the item

n is the total number of items,

$\sigma^{2}{}_{(Y_i)}$ represents individual item variance, and

σ_x^2 represents the total test variance

7.1 Reliability and Standard Errors of Measurement

Table 7-1 presents descriptive statistics, Cronbach's α coefficient, and raw score standard errors of measurement (SEMs) for each content area and grade (statistics are based on common items only).

Table 7-1.2008–09 NECAP: Common Item Raw Score Descriptives, Reliability, and SEM by Grade and Content

Grade	*Content*	*N*	*Possible Score*	*Min Score*	*Max Score*	*Mean Score*	*Score SD*	*Reliability (α)*	*SEM*
3	Mathematics	30449	65	0	65	42.725	12.669	0.933	3.285
	Reading	30385	52	0	52	34.916	9.982	0.887	3.357
4	Mathematics	30499	65	0	65	43.563	12.925	0.933	3.337
	Reading	30435	52	0	52	33.772	9.406	0.855	3.576
5	Mathematics	32200	66	0	66	37.734	14.072	0.913	4.149
	Reading	32130	52	0	51	29.433	8.353	0.880	2.891
	Writing	32065	37	0	37	23.126	5.507	0.720	2.914
6	Mathematics	32434	66	0	66	32.802	14.140	0.925	3.873
	Reading	32379	52	0	52	32.790	8.845	0.897	2.840
7	Mathematics	33011	66	0	66	31.835	14.343	0.923	3.989
	Reading	32915	52	0	52	33.926	8.258	0.880	2.867
8	Mathematics	33771	66	0	66	33.214	14.160	0.928	3.802
	Reading	33698	52	0	52	33.319	8.258	0.881	2.848
	Writing	33563	37	0	37	23.935	5.481	0.760	2.685
11	Mathematics	33771	64	0	64	24.684	13.730	0.926	3.743
	Reading	33702	52	0	52	33.909	8.894	0.890	2.945
	Writing	33653	12	0	12	6.023	1.994		

For mathematics, the reliability coefficient ranged from 0.91 to 0.93, for reading 0.86 to 0.90. For the grade 5 and grade 8 writing tests, the values were 0.72 and 0.76, respectively. Because different grades and content areas have different test designs (e.g., the number of items varies by test), it is inappropriate to make inferences about the quality of one test by comparing its reliability to that of another test from a different grade and/or content area.

The Alpha coefficients, broken down by subgroup, item type, and reporting category, are presented in Appendix N. These detailed α coefficient results are discussed in sections 7.2 through 7.4.

7.2 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2008–09 NECAP tests. Subgroup Cronbach's α's were calculated using the formula defined above using only the members of the subgroup in question in the computations. For mathematics, subgroup reliabilities ranged from 0.54 to 0.95, for reading from 0.69 to 0.92, and for writing from 0.62 to 0.80.

For several reasons, the results of this section should be interpreted with caution. First, inherent differences between grades and content areas preclude making valid inferences about the quality of a test based on statistical comparisons with other tests. Second, reliabilities are dependent not only on the measurement properties of a test but on the statistical distribution of the studied subgroup. For example, it can be readily seen in Appendix N that subgroup sample sizes may vary considerably, which results in natural variation in reliability coefficients. Or α, which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper and Smith, 1998). Third, there is no industry standard to interpret the strength of a reliability coefficient, and this is particularly true when the population of interest is a single subgroup.

7.3 Item Type Reliability

Another approach to estimating the reliability for a test with differing item types (i.e., multiple-choice and constructed-response) is to assume that at least a small, but important, degree of unique variance is associated with item type (Feldt and Brennan, 1989), in contrast to Cronbach's α, which assumes that there are no such local or clustered dependencies. A stratified version of coefficient α corrects for this problem by using the following formula:

$$\alpha_{strat} = 1 - \frac{\sum_{j=1}^{k} \sigma_{x_j}^2 (1-\alpha)}{\sigma_x^2}$$

where j indexes the subtests or categories,

$\sigma_{x_j}^2$ represents the variance of each of the k individual subtests or categories,

α_j is the unstratified Cronbach's α coefficient for each subtest, and

σ_x^2 represents the total test variance.

Stratified α was calculated separately for all grade contents but grade 11 writing. The results of stratification based on item type (MC versus CR) are presented in Appendix N. Not surprisingly, reliabilities were higher on the full test than on subsets of items.

Table 7-2.2008–09 NECAP: Common Item Cronbach's Alpha Reliability by Grade and Content—Overall, by Item Type, and Stratified by Item Type

		All	*MC*		*CR*		
Grade	*Content*	α	α	*N*	α	*N (poss.)*	*Stratified* α
3	Mathematics	0.93	0.88	35	0.87	20 (30)	0.93
	Reading	0.89	0.87	28	0.76	6 (24)	0.90
4	Mathematics	0.93	0.89	35	0.87	20 (30)	0.94
	Reading	0.86	0.87	28	0.66	6 (24)	0.87
5	Mathematics	0.91	0.85	32	0.86	16 (34)	0.92
	Reading	0.88	0.85	28	0.80	6 (24)	0.89
	Writing	0.72	0.70	10	0.59	3 (12)	0.73
6	Mathematics	0.92	0.87	32	0.87	16 (34)	0.93
	Reading	0.90	0.88	28	0.86	6 (24)	0.92
7	Mathematics	0.92	0.85	32	0.88	16 (34)	0.93
	Reading	0.88	0.84	28	0.86	6 (24)	0.90
8	Mathematics	0.93	0.88	32	0.87	16 (34)	0.93
	Reading	0.88	0.83	28	0.87	6 (24)	0.91
	Writing	0.76	0.63	10	0.70	3 (12)	0.78
11	Mathematics	0.93	0.83	24	0.89	22 (40)	0.93
	Reading	0.89	0.85	28	0.89	6 (24)	0.92

All = MC and CR; MC = multiple-choice; CR = constructed response= number of items; poss = total possible constructed-response points

7.4 Reporting Categories Reliability

In Section 7.3, the reliability coefficients were calculated based on form and item type. Item type represents just one way of breaking an overall test into subtests. Of even more interest are reliabilities for the reporting categories within NECAP content areas described in Chapter 2. Cronbach's α coefficients for reporting categories were calculated via the same alpha formula defined at the beginning of Chapter 7 using just the items of a given reporting category in the computations. These results are presented in Appendix N.

Once again as expected, because they are based on a subset of items rather than the full test, computed reporting category reliabilities were lower (sometimes substantially so) than were overall test reliabilities, and interpretations should take this into account.

For mathematics, reporting category reliabilities ranged from 0.61 to 0.74, for reading from 0.56 to 0.72, and for writing from 0.06 to 0.76. The reliabilities for the Extended Response writing categories were substantially lower than those of other categories because 12 of the 15 points for the category came from a single 12-point writing prompt item. In general, the reporting category reliabilities were lower than those based on the total test and approximately to the degree one would expect based on CTT. Qualitative differences between grades and content areas once again preclude valid inferences about the quality of the full test based on statistical comparisons among subtest.

7.5 Reliability of Achievement Level Categorization

All test scores contain measurement error; thus, classifications based on test scores are also subject to measurement error. After the 2008–09 NECAP achievement levels were specified and students classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. For every 2008–09 NECAP grade and content area, each student was classified into one of the following achievement levels: *Substantially Below Proficient* (SBP), *Partially Proficient* (PP), *Proficient* (P), or *Proficient With Distinction* (PWD). This section of the report explains the methodologies used to assess the reliability of classification decisions and presents the results.

7.5.1 Accuracy and Consistency

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated, because errorless test scores do not exist.

Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are given to the same group of students. In operational test programs, however, such a design is usually impractical. Instead, techniques, such as one due to Livingston and Lewis (1995), have been developed to estimate both the accuracy and consistency of classification decisions based on a single administration of a test. The Livingston and Lewis technique was used for the 2008–09 NECAP because it is easily adaptable to all types of testing formats, including mixed-format tests.

7.5.2 Calculating Accuracy

The accuracy and consistency estimates reported below make use of "true scores" in the CTT sense. A true score is the score that would be obtained if a test had no measurement error. Of course, true scores cannot be observed and so must be estimated. In the Livingston and Lewis method, estimated true scores are used to classify students into their "true" achievement level.

For the 2007–08 NECAP, after various technical adjustments were made (described in Livingston and Lewis, 1995), a 4 x 4 contingency table of accuracy was created for each content area and grade, where cell [i,j] represented the estimated proportion of students whose true score fell into achievement level i (where i = 1 – 4) and observed score into achievement level j (where j = 1 – 4). The sum of the diagonal entries, i.e., the proportion of students whose true and observed achievement levels matched one another, signified overall accuracy.

7.5.3 Calculating Consistency

To estimate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following statistical adjustments (per Livingston and Lewis, 1995), a new 4×4 contingency table was created for each content area and grade and populated by the proportion of students who would be classified into each combination of achievement levels according to the two (hypothetical) parallel test forms. Cell [i,j] of this table represented the estimated proportion of students whose observed score on the first form would fall into achievement level i (where i = 1 – 4), and whose observed score on the second form would fall into achievement level j(where j = 1 – 4). The sum of the diagonal entries, i.e., the proportion of students classified by the two forms into exactly the same achievement level, signified overall consistency.

7.5.4 Calculating Kappa

Another way to measure consistency is to use Cohen's (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. It is calculated using the following formula:

$$\kappa = \frac{\text{(Observed agreement) - (Chance agreement)}}{\text{1 - (Chance agreement)}} = \frac{\sum_i C_{ii} - \sum_i C_{i.}C_{.i}}{1 - \sum_i C_{i.}C_{.i}},$$

where:

$C_{i.}$ is the proportion of students whose observed achievement level would be *Level i* (where *i*=1 – 4) on the first hypothetical parallel form of the test;

$C_{.i}$ is the proportion of students whose observed achievement level would be *Level i* (where *i*=1 – 4) on the second hypothetical parallel form of the test;

C_{ii} is the proportion of students whose observed achievement level would be *Level i* (where *i*=1 – 4) on both hypothetical parallel forms of the test.

Because κ is corrected for chance, its values are lower than are other consistency estimates.

7.5.5 Results of Accuracy, Consistency, and Kappa Analyses

The accuracy and consistency analyses described above are tabulated in Appendix O. The appendix includes the accuracy and consistency contingency tables described above and the overall accuracy and consistency indices, including kappa.

Accuracy and consistency values conditional upon achievement level are also given in Appendix O. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.713 for the PP achievement level for mathematics grade 3. This figure indicates that among the students whose true scores placed them in the PP achievement level, 71.3% of them would be expected to be in the PP achievement level when categorized according to their observed score. Similarly, the corresponding consistency value of 0.616 indicates that 61.6% of students with observed scores in PP would be expected to score in the PP achievement level again if a second, parallel test form were used.

For some testing situations, the greatest concern may be decisions around level thresholds. For example, if a college gave credit to students who achieved an Advanced Placement test score of 4 or 5, but not to students with scores of 1, 2, or 3, one might be interested in the accuracy of the dichotomous decision below -4 versus 4-or- above. For the 2008–09 NECAP, Appendix O provides accuracy and consistency estimates at each cut point as well as false positive and false negative decision rates. (False positives are the proportion of students whose observed scores were above the cut and whose true scores were below the cut. False negatives are the proportion of students whose observed scores were below the cut and whose true scores were above the cut.)

The above indices are derived from Livingston and Lewis's (1995) method of estimating the accuracy and consistency of classifications. It should be noted that Livingston and Lewis discuss two versions of the accuracy and consistency tables. A standard version performs calculations for forms parallel to the form taken. An "adjusted" version adjusts the results of one form to match the observed score distribution obtained

in the data. The tables reported in Appendix O use the standard version for two reasons: 1) this "unadjusted" version can be considered a smoothing of the data, thereby decreasing the variability of the results; and 2) for results dealing with the consistency of two parallel forms, the unadjusted tables are symmetric, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms that are parallel, i.e., it is more intuitive and interpretable for two parallel forms to have the same statistical distribution as one another.

Descriptive statistics relating to the decision accuracy and consistency of the 2008–09 NECAP tests can be derived from Appendix O. For mathematics, overall accuracy ranged from 0.803 to 0.842; overall consistency ranged from 0.732 to 0.780; the kappa statistic ranged from 0.608 to 0.674. For reading, overall accuracy ranged from 0.759 to 0.823; overall consistency ranged from 0.672 to 0.753; the kappa statistic ranged from 0.519 to 0.627. Finally, for writing grades 5 and 8, overall accuracy was 0.595 and 0.659 in the two grades tested; overall consistency was 0.504 and 0.552; the kappa statistic was 0.321 and 0.362. For writing grade 11, DAC analysis was not possible considering the assessment comprised only one writing prompt.

Table 7-3 summarizes most of the results of Appendix O at a glance. As with other types of reliability, it is inappropriate when analyzing the decision accuracy and consistency of a given test to compare results between grades and content areas.

Table 7-3. 2008–09 NECAP: Summary of Decision Accuracy (and Consistency) Results

			Conditional on Level				*At Cut Point*		
Grade	*Content*	*Overall*	*SBP*	*PP*	*P*	*PWD*	*SBP:PP*	*PP:P*	*P:PWD*
3	Mathematics	.81 (.74)	.85 (.79)	.71 (.62)	.82 (.77)	.86 (.75)	.96 (.94)	.93 (.90)	.92 (.89)
4	Mathematics	.83 (.76)	.85 (.78)	.72 (.63)	.83 (.79)	.90 (.81)	.96 (.95)	.93 (.91)	.93 (.91)
5	Mathematics	.80 (.73)	.83 (.76)	.58 (.46)	.84 (.79)	.88 (.77)	.94 (.92)	.93 (.90)	.93 (.90)
6	Mathematics	.81 (.74)	.82 (.76)	.61 (.50)	.84 (.78)	.89 (.81)	.94 (.91)	.93 (.90)	.94 (.92)
7	Mathematics	.80 (.73)	.83 (.77)	.61 (.50)	.84 (.78)	.89 (.80)	.93 (.91)	.92 (.89)	.94 (.92)
8	Mathematics	.81 (.74)	.83 (.78)	.65 (.54)	.84 (.79)	.88 (.79)	.94 (.91)	.92 (.89)	.94 (.92)
11	Mathematics	.84 (.78)	.88 (.85)	.75 (.66)	.88 (.82)	.78 (.54)	.92 (.89)	.93 (.91)	.99 (.98)
3	Reading	.77 (.69)	.82 (.74)	.67 (.56)	.79 (.74)	.78 (.64)	.96 (.95)	.93 (.90)	.88 (.84)
4	Reading	.76 (.67)	.76 (.64)	.65 (.55)	.76 (.70)	.86 (.73)	.96 (.94)	.90 (.86)	.90 (.87)
5	Reading	.80 (.72)	.77 (.63)	.74 (.65)	.82 (.77)	.86 (.73)	.97 (.95)	.91 (.87)	.93 (.90)
6	Reading	.82 (.75)	.81 (.70)	.77 (.69)	.84 (.79)	.87 (.76)	.97 (.95)	.92 (.89)	.94 (.91)
7	Reading	.81 (.74)	.78 (.66)	.73 (.64)	.83 (.78)	.87 (.75)	.97 (.96)	.92 (.88)	.92 (.89)
8	Reading	.81 (.73)	.80 (.70)	.75 (.67)	.82 (.77)	.87 (.75)	.96 (.95)	.91 (.88)	.93 (.91)
11	Reading	.81 (.74)	.81 (.72)	.74 (.66)	.82 (.76)	.88 (.78)	.96 (.95)	.92 (.89)	.93 (.90)
5	Writing	.60 (.50)	.69 (.53)	.51 (.42)	.46 (.38)	.82 (.66)	.91 (.87)	.83 (.77)	.84 (.78)
8	Writing	.66 (.55)	.72 (.58)	.61 (.52)	.66 (.57)	.80 (.54)	.91 (.87)	.83 (.78)	.91 (.88)
11	Writing*	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

SBP = *Substantially Below Proficient*; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
* Because grade 11 writing comprised only one writing prompt, DAC analysis was not possible.

Chapter 8. VALIDITY

Because interpretations of test scores, and not a test itself, are evaluated for validity, the purpose of the 2008–09 NECAP Technical Report is to describe several technical aspects of the NECAP tests in support of score interpretations (AERA, 1999). Each chapter contributes an important component in the investigation of score validation: test development and design; test administration; scoring, scaling, and equating; item analyses; reliability; and score reporting.

The NECAP tests are based on and aligned with the content standards and performance indicators in the GLEs for mathematics, reading, and writing. Inferences about student achievement on the content standards are intended from NECAP results, which in turn serve the evaluation of school accountability and inform the improvement of programs and instruction.

The *Standards for Educational and Psychological Testing* (AERA, 1999) provides a framework for describing sources of evidence that should be considered when evaluating validity. These sources include evidence on the following five general areas: test content, response processes, internal structure, consequences of testing, and relationship to other variables. Although each of these sources may speak to a different *aspect* of validity, they are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

A measure of test content validity is to determine how well the test's tasks represent the curriculum and standards for each content area and grade level. This is informed by the item development process, including how test blueprints and test items align with the curriculum and standards. Validation through this content lens was extensively described in Chapter 2. In other words, the element's components discussed in the chapter—item alignment with content standards; item bias; sensitivity and content appropriateness review processes; adherence to the test blueprint; use of multiple item types; use of standardized administration procedures, with accommodated options for participation; and appropriate test administration training—are all components of content-based validity evidence. Every NECAP test question or prompt was aligned by educators to specific content standards and underwent several rounds of review for content fidelity and appropriateness. Items of multiple formats (MC, SA, and CR) were presented to students. Finally, tests were administered according to mandated standardized procedures, with allowable accommodations, and all test coordinators and test administrators were required to familiarize themselves with and adhere to all of the procedures outlined in the *NECAP Test Coordinator Manual* and *Test Administrator* M*anual.*

The scoring information in Chapter 4 describes the steps taken to train and monitor hand-scorers as well as the quality control procedures related to machine-scanning and scoring. Additional studies might be informative on student response processes. For example, think-aloud protocols could be used to investigate students' cognitive processes when confronting test items.

Evidence on internal structure is extensively detailed in the chapters on item analyses, scaling and equating, and reliability (Chapters 5–7). Technical characteristics of the internal structure of the tests were

presented in terms of classical item statistics (p-values and discriminations), DIF analyses, several reliability coefficients, SEMs, multidimensionality hypothesis-testing and effect size estimation, and IRT analyses. In general, item difficulty indices were within acceptable and expected ranges. Chapter 6 also describes the procedures used to equate the 2008–09 test to the 2007–08 scales.

Evidence on the consequences of testing is addressed in information on scaled score and reporting in Chapters 6 and 9 and in the *Guide to Using the 2008 NECAP Reports*, which is a separate document referenced in the discussion of reporting. Each of these speak to efforts undertaken for providing the public with accurate and clear test score information. Scaled scores simplify results reporting across content areas, grade levels, and successive years. Achievement levels give reference points for mastery at each grade level, another useful and simple way to interpret scores. Several different standard reports were provided to stakeholders. Evidence on the consequences of testing could be supplemented with broader research on the impact on student learning of NECAP testing.

8.1 Questionnaire Data

A measure of external validity was provided by comparing student performance with answers to a questionnaire administered at the end of test. The grades 3 through 8 questionnaire contained 32 questions (9 concerned reading, 11 mathematics, and 12 writing). The grade 11 questionnaire contained 36 questions (12 concerned writing, 11 reading, and 13 mathematics). Most of the questions were designed to gather information about students and their study habits; however, a subset could be utilized in the evaluation of external validity. One question from each content area was most expected to correlate with student performance on NECAP tests. To the extent that the answers to those questions did correlate with student performance in the anticipated manner, the external validity of score interpretations was confirmed. The three questions are now discussed one at a time.

Question 8 (grades 3–8)/Question 21 (grade 11) concerning reading, read as follows

> How often do you choose to read in your free time?
>
> A. almost every day
> B. a few times a week
> C. a few times a month
> D. I almost never read.

It was anticipated that students who read more in their free time would have higher average scaled scores than students who did not read as much. In particular, it was expected that the reading performance of students who chose “A” would exceed that of students who chose “B,” whose performance would exceed that of students who chose “C,” whose performance would exceed that of students who chose “D.” This pattern may be observed in the scaled scores in Table 8-1.

Table 8-1. 2008–09 NECAP: Average Scaled Score, Counts and Percentages, within Performance Levels, of Responses to Spare-Time Reading Item[1] on Student Questionnaire—Reading

Grade	Resp	Number Resp	% Resp	Avg SS	N SBP	N PP	N P	N PWD	% SBP	% PP	% P	% PWD
3	(blank)	3953	13	344	694	702	1863	694	18	18	47	18
	A	15094	50	348	1288	2095	8113	3598	9	14	54	24
	B	7339	24	348	619	1045	4097	1578	8	14	56	22
	C	1648	5	344	255	284	847	262	15	17	51	16
	D	2353	8	342	436	515	1179	223	19	22	50	9
4	(blank)	3487	11	443	525	710	1618	634	15	20	46	18
	A	14432	47	448	1137	2316	7206	3773	8	16	50	26
	B	8740	29	446	812	1609	4613	1706	9	18	53	20
	C	1733	6	444	222	352	911	248	13	20	53	14
	D	2043	7	440	423	508	929	183	21	25	45	9
5	(blank)	3535	11	544	485	765	1780	505	14	22	50	14
	A	14373	45	548	876	2366	8195	2936	6	16	57	20
	B	9999	31	546	755	2138	5762	1344	8	21	58	13
	C	2054	6	544	221	466	1159	208	11	23	56	10
	D	2169	7	540	384	645	1038	102	18	30	48	5
6	(blank)	3876	12	643	591	903	1915	467	15	23	49	12
	A	11228	35	650	685	1582	6701	2260	6	14	60	20
	B	10968	34	646	826	2253	6575	1314	8	21	60	12
	C	3208	10	644	311	733	1905	259	10	23	59	8
	D	3099	10	640	584	921	1465	129	19	30	47	4
7	(blank)	3808	12	744	538	831	1943	496	14	22	51	13
	A	9464	29	752	411	1014	5500	2539	4	11	58	27
	B	10850	33	748	628	1828	6595	1799	6	17	61	17
	C	4173	13	746	295	901	2532	445	7	22	61	11
	D	4620	14	742	591	1256	2491	282	13	27	54	6
8	(blank)	3442	10	841	683	864	1480	415	20	25	43	12
	A	8865	26	851	378	1187	5092	2208	4	13	57	25
	B	10271	30	847	739	2128	5901	1503	7	21	57	15
	C	5039	15	844	460	1297	2774	508	9	26	55	10
	D	6081	18	840	876	2062	2864	279	14	34	47	5
11	(blank)	7779	23	1144	1152	1533	3643	1451	15	20	47	19
	A	5805	17	1147	406	830	3137	1432	7	14	54	25
	B	6807	20	1147	566	1132	3572	1537	8	17	52	23
	C	6011	18	1146	486	1045	3326	1154	8	17	55	19
	D	7183	21	1143	791	1695	3909	788	11	24	54	11

[1]Question: How often do you choose to read in your free time? A = almost every day; B = a few times a week; C = a few times a month; D = I almost never read.
SBP = *Substantially Below Proficient*; PP = *Partially Proficient*; P = *Proficient*; PWD = *Proficient with Distinction*

Question 16 (grades 3–8)/Question 32 (grade 11) concerning mathematics read as follows

How often do you have mathematics homework [assignments]?

A. almost every day

B. a few times a week

C. a few times a month

D. I usually don't have homework in mathematics.

The relationship between the question and student performance in mathematics (see Table 8-2) mirrored the scaled-score pattern observed for reading as, on average, mathematics performance appeared related to the frequency of mathematics homework assignments."

Table 8-2.2008–09 NECAP: Average Scaled Score, Counts and Percentages within Performance Levels, of Responses to Frequency of Mathematics-Homework Item[1] of Student Questionnaire—Mathematics

Grade	*Resp*	*N Resp*	*% Resp*	*Avg SS*	*N SBP*	*N PP*	*N P*	*N PWD*	*% SBP*	*% PP*	*% P*	*% PWD*
3	(blank)	3982	13	341	848	812	1674	648	21	20	42	16
	A	14307	47	345	1803	2643	6728	3133	13	18	47	22
	B	8838	29	344	1044	1711	4459	1624	12	19	50	18
	C	1685	6	344	258	318	797	312	15	19	47	19
	D	1639	5	341	351	357	735	196	21	22	45	12
4	(blank)	3514	12	442	676	685	1503	650	19	19	43	18
	A	16261	53	446	1861	2746	7832	3822	11	17	48	24
	B	8279	27	445	975	1509	3955	1840	12	18	48	22
	C	1411	5	445	213	243	617	338	15	17	44	24
	D	1034	3	439	275	226	397	136	27	22	38	13
5	(blank)	3573	11	542	856	541	1541	635	24	15	43	18
	A	18176	56	545	2662	2612	8995	3907	15	14	49	21
	B	8187	25	544	1473	1258	4019	1437	18	15	49	18
	C	1336	4	543	275	225	614	222	21	17	46	17
	D	928	3	539	285	169	363	111	31	18	39	12
6	(blank)	3873	12	641	1007	680	1515	671	26	18	39	17
	A	18095	56	644	2847	3127	8152	3969	16	17	45	22
	B	8690	27	643	1644	1653	3775	1618	19	19	43	19
	C	996	3	641	242	180	406	168	24	18	41	17
	D	780	2	636	319	135	241	85	41	17	31	11
7	(blank)	3862	12	739	1221	744	1350	547	32	19	35	14
	A	19007	58	744	3095	3370	8358	4184	16	18	44	22
	B	8355	25	741	1940	1622	3465	1328	23	19	41	16
	C	982	3	738	331	199	326	126	34	20	33	13
	D	805	2	732	427	139	182	57	53	17	23	7
8	(blank)	3509	10	838	1144	691	1231	443	33	20	35	13
	A	20343	60	844	2837	3793	9497	4216	14	19	47	21
	B	8013	24	840	1838	1923	3349	903	23	24	42	11
	C	967	3	837	334	227	331	75	35	23	34	8
	D	939	3	834	433	182	244	80	46	19	26	9
11	(blank)	7924	24	1133	3492	2074	2249	109	44	26	28	1
	A	17724	53	1137	5465	5617	6286	356	31	32	35	2
	B	5082	15	1133	2534	1437	1069	42	50	28	21	1
	C	1336	4	1129	865	309	159	3	65	23	12	0
	D	1481	4	1128	1056	304	115	6	71	21	8	0

[1]Question: How often do you have mathematics homework [assignments]? A = almost every day; B = a few times a week; C = a few times a month; D = I usually don't have homework in mathematics.
SBP = *Substantially Below Proficient*; PP = *Partially Proficient*; P = *Proficient*; PWD = *Proficient with Distinction*

Question 32 (grades 3–8)/Question 12 (grade 11), concerning writing, read as follows

What kinds of writing do you do most in school?

A. I mostly write stories.

B. I mostly write reports.

C. I mostly write about things I've read.

D. I do all kinds of writing.

For this question, the only anticipated outcome was that students who selected choice "D," i.e., those who ostensibly had experience in many different kinds of writing, would tend to outperform students who selected any other answer choice. The expected outcome was realized in all three grades, though report writing in grade 11 was also associated with higher performance (see Table 8-3).

Table 8-3.2008–09 NECAP: Average Scaled Score, Counts and Percentages, within Performance Levels, of Responses to Kinds of School Writing Item[1] of Student Questionnaire—Writing

Grade	*Resp*	*N Resp*	*% Resp*	*Avg SS*	*N SBP*	*N PP*	*N P*	*N PWD*	*% SBP*	*% PP*	*% P*	*% PWD*
5	(blank)	4255	13	542	954	1146	1148	1007	22	27	27	24
	A	6627	21	543	1122	1954	1963	1588	17	29	30	24
	B	2654	8	542	535	818	778	523	20	31	29	20
	C	2922	9	544	503	823	843	753	17	28	29	26
	D	15607	49	548	1859	3589	4804	5355	12	23	31	34
8	(blank)	4283	13	837	1094	1446	1438	305	26	34	34	7
	A	3764	11	837	834	1575	1195	160	22	42	32	4
	B	4973	15	839	878	1901	1829	365	18	38	37	7
	C	3784	11	839	648	1406	1454	276	17	37	38	7
	D	16759	50	843	1594	5132	8034	1999	10	31	48	12
11	(blank)	7770	23	5.8	1119	3692	2668	291	14	48	34	4
	A	1343	4	5.2	246	758	317	22	18	56	24	2
	B	7816	23	6.2	578	3923	3026	289	7	50	39	4
	C	4128	12	5.9	474	2047	1458	149	11	50	35	4
	D	12461	37	6.2	950	6055	4994	462	8	49	40	4
	E	7770	23	5.8	1119	3692	2668	291	14	48	34	4

[1]Question: What kinds of writing do you do most in school? A = I mostly write stories; B = I mostly write reports; C = I mostly write about things I've read; D = I do all kinds of writing. SBP = *Substantially Below Proficient*; PP = *Partially Proficient*; P = *Proficient*; PWD = *Proficient with Distinction*

Based on the foregoing analysis, the relationship between questionnaire data and performance on the NECAP was consistent with expectations of the three questions selected for the investigation of external validity. See Appendix P for a copy of the questionnaire and complete data comparing questionnaire items and test performance.

8.2 Validity Studies Agenda

The remaining part of this chapter describes further studies of validity that could enhance the investigations of validity that have already been performed. The proposed areas of validity to be examined fall into four categories: *external validity, convergent and discriminant validity, structural validity, and procedural validity*. These will be discussed in turn.

8.2.1 External Validity

In the future, investigations of external validity could involve targeted examination of variables which one might expect correlate with NECAP results, like classroom grades or classroom test scores in the same content areas as the NECAP test in question, for example.

Further evidence of external validity might come from correlating NECAP scores with scores on another standardized test, such as the Iowa Test of Basic Skills (ITBS). As with the study of concordance between NECAP scores and grades, this investigation would compare scores in analogous content areas (e.g., NECAP reading and ITBS reading comprehension). All tests taken by each student would be appropriate to the student's grade level.

8.2.2 Convergent and Discriminant Validity

The concepts of convergent and discriminant validity were defined by Campbell and Fiske (1959) as specific types of validity that fall under the umbrella of *construct validity*. The notion of convergent validity states that measures or variables that are intended to align with one another should actually be aligned in practice. Discriminant validity, on the other hand, is the idea that measures or variables that are intended to differ from one another should not be too highly correlated. Evidence for validity comes from examining whether the correlations among variables are as expected in direction and magnitude.

Campbell and Fiske (1959) introduced the study of different *traits* and *methods* as the means of assessing convergent and discriminant validity. Traits refer to the constructs that are being measured (e.g., mathematical ability), and methods are the instruments of measuring them (e.g., a mathematics test or grade). To utilize the framework of Campbell and Fiske, it is necessary that more than one trait and more than one method be examined. Analysis is performed through the multi-trait/multi-method matrix, which gives all possible correlations of the different combinations of traits and methods. Campbell and Fiske defined four properties of the multi-trait/multi-method matrix that serve as evidence of convergent and discriminant validity:

- The correlation among different methods of measuring the same trait should be sufficiently different from zero. For example, scores on a mathematics test and grades in a mathematics class should be positively correlated.

- The correlation among different methods of measuring the same trait should be higher than that of different methods of measuring different traits. For example, scores on a mathematics test and grades in a mathematics class should be more highly correlated than are scores on a mathematics test and grades in a reading class.
- The correlation among different methods of measuring the same trait should be higher than the same method of measuring different traits. For example, scores on a mathematics test and grades in a mathematics class should be more highly correlated than scores on a mathematics test and scores on an analogous reading test.
- The pattern of correlations should be similar across comparisons of different traits and methods. For example, if the correlation between test scores in reading and writing is higher than the correlation between test scores in reading and mathematics, it is expected that the correlation between grades in reading and writing would also be higher than the correlation between grades in reading and mathematics.

For NECAP, convergent and discriminant validity could be examined by constructing a multi-trait/multi-method matrix and analyzing the four pieces of evidence described above. The traits examined would be mathematics, reading, and writing; different methods could include NECAP score and such variables as grades, teacher judgments, scores on another standardized test, etc.

8.2.3 Structural Validity

Though the previous types of validity examine the concurrence between different measures of the same content area, structural validity focuses on the relation between strands *within* a content area, thus supporting *content validity*. Standardized tests are carefully designed to ensure that all appropriate strands of a content area are adequately covered in a test, and structural validity is the degree to which related elements of a test are correlated in the intended manner. For instance, it is desired that performance on different strands of a content area be positively correlated; however, as these strands are designed to measure distinct components of the content area, it is reasonable to expect that each strand would contribute a unique component to the test. Additionally, it is desired that the correlation between different item types (MC, SA, and CR) of the same content area be positive.

As an example, an analysis of NECAP structural validity would investigate the correlation between performance in Geometry and Measurement and performance in Functions and Algebra. Additionally, the concordance between performance on MC items and CR items would be examined. Such a study would address the consistency of NECAP tests within each grade and content area. In particular, the dimensionality analyses of Chapter 5 could be expanded to include confirmatory analyses addressing these concerns.

8.2.4 Procedural Validity

As mentioned earlier, the *NECAP Test Coordinator* and *Test Administrator* manuals delineated the procedures to which all NECAP test coordinators and test administrators were required to adhere. A study of procedural validity would provide a comprehensive documentation of the procedures that were followed throughout the NECAP administration. The results of the documentation would then be compared to the manuals, and procedural validity would be confirmed to the extent that the two were in alignment. Evidence of procedural validity is important because it verifies that the actual administration practices were in accord with the intentions of the design.

Possible instances where discrepancies can exist between design and implementation include the following: A teacher spirals test forms incorrectly within a classroom; cheating among students occurs; answer documents are scanned incorrectly. These are examples of *administration error*. A study of procedural validity involves capturing any administration errors and presenting them within a cohesive document for review.

All potential tests of validity that have been introduced in this chapter should be considered by the NECAP Technical Advisory Committee (NECAP TAC) during 2009–10. With the advice of the NECAP TAC, the states will develop short-term and longer term (e.g., 2-year to 5-year) plans for validity studies

SECTION III—NECAP REPORTING

Chapter 9. SCORE REPORTING

9.1 Teaching Year vs. Testing Year Reporting

The data used for the NECAP reports are the results of the fall 2008 administration of the NECAP test. It is important to note that the NECAP tests are based on the GLEs from the previous year. For example, the grade 7 NECAP test administered in the fall of seventh grade is based on the grade 6 GLEs. Because many students receive instruction at a different school from where they were tested, the state Departments of Education determined that access to results information would be valuable to both the school where the student was tested and the school where the student received instruction. To achieve this goal, separate Item Analysis, School and District Results, and School and District Summary reports were created for the "testing" school and the "teaching" school. Every student who participated in the NECAP test was represented in testing reports, and most students were represented in teaching reports. In some cases (e.g. a student who recently moved to the state), it is not possible to provide information for a student in a "teaching" report.

9.2 Primary Reports

There were four primary reports for the 2008–09 NECAP:

- Student Report
- Item Analysis Report
- School and District Results Report
- School and District Summary Report

With the exception of the Student Report, all reports were available for schools and districts to view or download on a password-secure Website hosted by Measured Progress. Student-level data files were also available for districts to download from the secure Web site. Each of these reports is described in the following sections. Sample reports are provided in Appendix Q.

9.3 Student Report

The *NECAP Student Report* is a single-page double-sided report printed on 8.5" by 14" paper. The front of the report includes informational text about the design and uses of the assessment. The front of the report also contains text describing the three corresponding sections on the reverse side of the student report and the achievement level definitions. The reverse side of the student report provides a complete picture of an individual student's performance on the NECAP, divided into three sections. The first section provides the student's overall performance for each content area. The student's achievement levels are provided, and

scaled scores are presented numerically as well as in a graphic that depicts the scaled score with the standard error of measurement bar constructed about it, set within the full range of possible scaled scores demarcated into the four achievement levels.

The second section displays the student's achievement level in each content area relative to the percentage of students at each achievement level within the school, district, and state.

The third section shows the student's raw-score performance in content area reporting categories relative to possible points, gives the average points earned for the school, district, and state; and gives the average points earned by students at the *Proficient* level on the overall content area test. For reading, with the exception of Word ID/Vocabulary items, items are reported by Type of Text (Literary, Informational) and Level of Comprehension (Initial Understanding, Analysis and Interpretation). For mathematics, the reporting subcategories are Numbers and Operations; Geometry and Measurement; Functions and Algebra; and Data, Statistics, and Probability. The content area categories for writing at grades 5 and 8 are Structures of Language & Writing Conventions, Short Responses and Extended Response. Grade 11 writing only reports Extended Response as a category.

During scoring of the writing prompt, each scorer selects up to three comments about the student's writing performance. The comments are selected from a predetermined list produced by the writing representatives from each state's Department of Education. These scorers' comments are presented in a box next to the writing results.

The *NECAP Student Report* is confidential and should be kept secure within the school and district. The Family Educational Rights and Privacy Act (FERPA) requires that access to individual student results be restricted to the student, the student's parents/guardians, and authorized school personnel.

9.4 Item Analysis Reports

The *NECAP Item Analysis Report* provides a roster of all students in a school and provides their performances on the common items that are released to the public, one report per content area. For all grades and content areas, the student names and identification numbers are listed as row headers down the left side of the report. For grades 3 through 8 and 11 in reading and mathematics, and grades 5 and 8 writing, the items are listed as column headers in the same order they appeared in the released item documents (not the position in which they appeared on the test).

For each item, seven pieces of information are shown: the released item number, the content strand for the item, the GLE/GSE code for the item, the DOK code for the item, the item type, the correct response key for MC items, and the total possible points.

For each student, MC items are marked either with a plus sign (+), indicating that the student chose the correct MC response, or a letter (from A to D), indicating the incorrect response chosen by the student. For SA and CR items, the number of points earned is shown. All responses to released items are shown in the report, regardless of the student's participation status.

The columns on the right side of the report show the Total Test Results, broken into several categories. A Subcategory Points Earned columns show points earned by the student in each content area relative to total points possible. A Total Points Earned column is a summary of all points earned and total possible points in the content area. The last two columns show the student's Scaled Score and Achievement Level. Students reported as Not Tested are given a code in the Achievement Level column to indicate the reason why the student did not test. Descriptions of these codes can be found on the legend, after the last page of data on the report. It is important to note that not all items used to compute student scores are included in this report, only released items. At the bottom of the report, the average percentage correct for each MC item and average scores for the SA and CR items and writing prompts are shown for the school, district, and state.

For grade 11 writing, the top portion of the *NECAP Item Analysis Report* consists of a single row of item information containing: the content stand, GSE codes, DOK code, item type/writing prompt, and total possible points. The student names and identification numbers are listed as row headers down the left side of the report. The Total Test Results section to the right includes Total Points Earned and Achievement Level for each student. At the bottom, the average points earned on the writing prompt are provided for the school, district, and state.

The *NECAP Item Analysis Report* is confidential and should be kept secure within the school and district. The FERPA requires that access to individual student results be restricted to the student, the student's parents/guardians, and authorized school personnel.

9.5 School and District Results Reports

The *NECAP School Results Report* and the *NECAP District Results Report* consist of three parts: the grade level summary report (page 2), the results for the content areas (pages 3, 5, and 7), and the disaggregated content area results (pages 4, 6, and 8).

The grade level summary report provides a summary of participation in the NECAP and a summary of NECAP results. The participation section on the top half of the page shows the number and percentage of students who were enrolled on or after October 1, 2008. The total number of students enrolled is defined as the number of students tested plus the number of students not tested.

Because students who were not tested did not participate, average school scores were not affected by non-tested students. These students were included in the calculation of the percentage of students participating, but not in the calculation of scores. For students who participated in some but not all sessions of the NECAP test, actual scores were reported for the content areas in which they participated. These reporting decisions were made to support the requirement that all students participate in the NECAP testing program.

Data are provided for the following groups of students who may not have completed the entire battery of NECAP tests:

- **Alternate Assessment**: Students in this category completed an alternate test for the 2007–08 school year.
- **First-Year LEP**: Students in this category are defined as being new to the United States after October 1, 2007 and were not required to take the NECAP tests in reading and writing. Students in this category were expected to take the mathematics portion of the NECAP.
- **Withdrew After October 1**: Students withdrawing from a school after October 1, 2008 may have taken some sessions of the NECAP tests prior to their withdrawal from the school.
- **Enrolled After October 1**: Students enrolling in a school after October 1, 2008 may not have had adequate time to participate fully in all sessions of NECAP testing.
- **Special Consideration**: Schools received state approval for special consideration for an exemption on all or part of the NECAP tests for any student whose circumstances are not described by the previous categories but for whom the school determined that taking the NECAP tests would not be possible.
- **Other**: Occasionally students will not have completed the NECAP tests for reasons other than those listed above. These "other" categories were considered not state approved.

The results section in the bottom half of the page shows the number and percentage of students performing at each achievement level in each of the three content areas across the school, district, and state. In addition, a mean scaled score is provided for each content area across school, district, and state levels except for grade 11 writing where the mean raw score is provided across the school, district, and state. School information is blank for the district version of this report..

The content area results pages provide information on performance in specific content categories of the tested content areas (for example, Geometry, and Measurement within mathematics). The purpose of these sections is to help schools to determine the extent to which their curricula are effective in helping students to achieve the particular standards and benchmarks contained in the *Grade-Level and Grade-Span Expectations*. Information about each content area (reading, mathematics and writing) for school, district, and state includes

- the total number of students enrolled, not tested (state-approved reason), not tested (other reason), and tested;
- the total number and percentage of students at each achievement level (based on the number in the tested column); and
- the mean scaled score.

Information about each content area reporting category for reading, mathematics and writing includes the following

- The total possible points for that category. In order to provide as much information as possible for each category, the total number of points includes both the common items used to calculate scores and additional items in each category used for equating the test from year to year.
- A graphic display of the percent of total possible points for the school, state, and district. In this graphic display, there are symbols representing school, district, and state performance. In addition, there is a line representing the standard error of measurement. This statistic indicates how much a student's score could vary if the student were examined repeatedly with the same test (assuming that no learning were to occur between test administrations).
- For grade 11 writing only, a column showing the number of prompts for each subtopic (strand) is also provided, as well as the distribution of score points across prompts within each strand in terms of percentages for the school, district, and state.

The disaggregated content area results pages present the relationship between performance and student reporting variables (see list below) in each content area across school, district, and state levels. Each content area page shows the number of students categorized as enrolled, not tested (state-approved reason), not tested (other reason), and tested. The tables also provide the number and percentage of students within each of the four achievement levels and the mean scaled score by each reporting category.

The list of student reporting categories is as follows:

- All Students
- Gender
- Primary Race/Ethnicity
- LEP Status (limited English proficiency)
- IEP
- SES (socioeconomic status)
- Migrant
- Title I
- 504 Plan

The data for achievement levels and mean scaled score are based on the number shown in the tested column. The data for the reporting categories were provided by information coded on the students' answer booklets by teachers and/or data linked to the student label. Because performance is being reported by categories that can contain relatively low numbers of students, school personnel are advised, under FERPA guidelines, to treat these pages confidentially.

It should be noted that for NH and VT, no data were reported for the 504 Plan in any of the content areas. In addition, for VT, no data were reported for Title I in any of the content areas.

9.6 School and District Summary Reports

The *NECAP School Summary Report* and the *NECAP District Summary Report* provide details, broken down by content area, on student performance by grade level tested in the school. The purpose of the summary is to help schools determine the extent to which their students achieve the particular standards and benchmarks contained in the *Grade-Level Expectations* and *Grad- Span Expectations*.

Information about each content area and grade level for school, district, and state includes

- the total number of students enrolled, not tested (state-approved reason), not tested (other reason), and tested;
- the total number and percentage of students at each achievement level (based on the number in the tested column); and
- the mean scaled score (mean raw score for grade 11 writing)

The data reported, the report format, and the guidelines for using the reported data are identical for both the school and district reports. The only difference between the reports is that the *NECAP District Summary Report* includes no individual school data. Separate school report and district reports were produced for each grade level tested.

9.7 Decision Rules

To ensure that reported results for the 2008–09 NECAP are accurate relative to collected data and other pertinent information, a document that delineates analysis and reporting rules was created. These decision rules were observed in the analyses of NECAP test data and in reporting the test results. Moreover, these rules are the main reference for quality assurance checks.

The decision rules document used for reporting results of the October 2008 administration of the NECAP is found in Appendix R.

The first set of rules pertains to general issues in reporting scores. Each issue is described, and pertinent variables are identified. The actual rules applied are described by the way they impact analyses and aggregations and their specific impact on each of the reports. The general rules are further grouped into issues pertaining to test items, school type, student exclusions, and number of students for aggregations.

The second set of rules pertains to reporting student participation. These rules describe which students were counted and reported for each subgroup in the student participation report.

9.8 Quality Assurance

Quality assurance measures are embedded throughout the entire process of analysis and reporting. The data processor, data analyst, and psychometrician assigned to work on NECAP implement quality control checks of their respective computer programs and intermediate products. Moreover, when data are handed off to different functions within the Research and Analysis division, the sending function verifies that the data are accurate before handoff. Additionally, when a function receives a data set, the first step is to verify the data for accuracy.

Another type of quality assurance measure is parallel processing. Students' scaled scores for each content area are assigned by a psychometrician through a process of equating and scaling. The scaled scores are also computed by a data analyst to verify that scaled scores and corresponding achievement levels are assigned accurately. Respective scaled scores and assigned achievement levels are compared across all students for 100% agreement. Different exclusions that determine whether each student receives scaled scores and/or is included in different levels of aggregation are also parallel-processed. Using the decision rules document, two data analysts independently write a computer program that assigns students' exclusions. For each content area and grade combination, the exclusions assigned by each data analyst are compared across all students. Only when 100% agreement is achieved can the rest of data analysis be completed.

The third aspect of quality control involves the procedures implemented by the quality assurance group to check the accuracy of reported data. Using a sample of schools and districts, the quality assurance group verifies that reported information is correct. The step is conducted in two parts: (1) verify that the computed information was obtained correctly through appropriate application of different decision rules and (2) verify that the correct data points populate each cell in the NECAP reports. The selection of sample schools and districts for this purpose is very specific and can affect the success of the quality control efforts. There are two sets of samples selected that may not be mutually exclusive.

The first set includes those that satisfy the following criteria:

- One-school district
- Two-school district
- Multi-school district

The second set of samples includes districts or schools that have unique reporting situations as indicated by decision rules. This second set is necessary to ensure that each rule is applied correctly. The second set includes the following criteria:

- Private school
- Small school that receives no school report
- Small district that receives no district report
- District that receives a report but all schools are too small to receive a school report
- School with excluded (not tested) students
- School with home-schooled students

The quality assurance group uses a checklist to implement its procedures. After the checklist is completed, sample reports are circulated for psychometric checks and program management review. The appropriate sample reports are then presented to the client for review and sign-off.

SECTION IV—REFERENCES

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

Brown, F. G. (1983). Principles of educational and psychological testing (3rd ed.). Fort Worth: Holt, Rinehart and Winston.

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

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37–46.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297–334.

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

Dorans, N. J., and Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. Journal of Educational Measurement, 23, 355–368.

Draper, N. R., and Smith, H. (1998). Applied Regression Analysis (3rd ed.). New York: John Wiley and Sons, Inc.

Feldt, L. S., and Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), Educational measurement (3rd ed.) (pp. 105–146). New York: Macmillan Publishing Co.

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

Hambleton, R. K., Swaminathan, H., and Rogers, H. J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage Publications.

Hambleton, R. K., and van der Linden, W. J. (1997). Handbook of modern item response theory. New York: Springer-Verlag.

Joint Committee on Testing Practices (1988). Code of Fair Testing Practices in Education. Washington, D.C.: National Council on Measurement in Education.

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

Lord, F. M., and Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

Muraki, E. and R. D. Bock (2003). PARSCALE 4.1. Lincolnwood, IL: Scientific Software International.

Subkoviak, M.J. (1976). Estimating reliability from a single administration of a mastery test. Journal of Educational Measurement, 13, 265 –276.

Stout, W. F. (1987). A nonparametric approach for assessing latent trait dimensionality. Psychometrika, 52, 589 –617.

Stout, W. F., Froelich, A. G., and Gao, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. In A. Boomsma, M. A. J. van Duign, & T. A. B. Snijders (Eds.), Essays on Item Response Theory, (pp. 357 –375). New York: Springer-Verlag.

Zhang, J., and Stout, W. F. (1999). The theoretical DETECT index of dimensionality and its application to approximate simple structure. Psychometrika, 64, 213 –249.

APPENDICES

Appendix A—COMMITTEE MEMBERSHIP

TECHNICAL ADVISORY COMMITTEE

NEW HAMPSHIRE

Name	*Association/Affiliation*
Richard Hill	Center for Assessment, Board of Trustees Chair
Scott Marion	Center for Assessment, Associate Director
Charles Pugh	Moultonborough District Assessment Coordinator
Rachel Quenemoen	University of Minnesota
Stanley Rabinowitz	WestEd, Assessment & Standards Development Services Director
Christine Rath	Concord, Superintendent
Steve Sireci	University of Massachusetts Professor
Carina Wong	Consultant

RHODE ISLAND

Name	*Association/Affiliation*
Sylvia Blanda	Westerly School Department
Bill Erpenbach	WJE Consulting, Ltd.
Richard Hill	Center for Assessment, Board of Trustees Chair
Jon Mickelson	Providence School Department
Joe Ryan	Consultant
Lauress Wise	HumRRO, President

VERMONT

Name	*Association/Affiliation*
Dale Carlson	NAEP Coach, NAEO-Westat
Lizanne DeStefano	Bureau of Educational Research
Jonathan Dings	Boulder, Co. School District
Brian Gong	Center for Assessment, Executive Director
Bill Mathis	Rutland Northeast Supervisory Union, Superintendent of Schools
Bob McNamara	Washington West Supervisory Union, Superintendent of Schools
Bob Stanton	Lamoille South Supervisory Union, Assistant Superintendent of Schools
Phoebe Winter	Consultant

NEW HAMPSHIRE ITEM REVIEW COMMITTEE MARCH 18 & 19, 2008

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Kenneth	Boisselle	Souhegan HS	English Teacher - World Lit
Barbara	Boschmans	Plymouth State University	Mathematics Professor
Gail	Bourn	Elem Street School	Reading/Writing Teacher
Nancy	Bozek	Mt. Pleasant School	Elem Ed, ELA, M.ED. Administration
Susan	Broadhurst	Timberlane Middle School	Mathematics Teacher
Heather	Caldwell	Piermont Village School	Mathematics Teacher
Marina	Capen	Souhegan HS	Mathematics Teacher
Susan	Dean-Olson	Kingswood Regional HS	Reading Teacher
Donna	Dubey	Winnisquam Reg HS	Mathematics Teacher
Lisa	Dwyer	Merrimack Valley Middle	Reading Teacher
Judy	Filkins	Lebanon District	District Mathematics Coordinator
Jack	Finley	Franklin HS	English Language Arts Teacher
Kathy	Fowler	Sandown Central	Mathematics Teacher
Ann	Gehring	New Durham School	K-8 Reading Specialist
Amy	Kramer	Southside Middle School	Elem Ed Teacher
Nancy	Maquire	Deerfield Community School	Elem ED and Reading Specialist
Laura	Maroney	Wilkins Elementary	Elem Reading Specialist
Nancy	Monks	Amherst Middle School	Mathematics Teacher
Denise	Pazdon	North Hampton School	Spec Ed Teacher - Language Specialist
John	Potucek	Southside Middle School	Mathematics Teacher
Thomas	Power	Spaulding HS	English Teacher
Bonnie	Robinson	Newport Middle HS	English Language Arts Teacher
Mary-Ellen	Russell	Manchester HS West	Mathematics Teacher
Sara	Scheuch	New London Elementary	Elem ED and Reading Specialist
Marilyn	St. George	Wilkins Elementary	Elem ED and Reading Specialist
Trish	Underwood	Maple Street School	K-12 SpecED, Elem ed, Middle/High English
Lynda	Walsh	Pennichuck Middle School	Mathematics Teacher
Stephanie	Wheeler	Wilson Elementary School	Title 1 Math Supervisor/Coach
Tracy	Bricchi	Kearsarge	Mathematics Teacher
Matt	Cygan	Manchester Memorial HS	Mathematics Teacher
Jill	Duffield	New Durham Elementary	Core subjects Teacher, literacy coordinator/leader
Martha	Hardiman	Jefferson/Whitefield Schools	English Language Arts Teacher

RHODE ISLAND ITEM REVIEW COMMITTEE MARCH 18 & 19, 2008

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Kara	Alling	Woonsocket Middle School	English Language Arts Teacher
Michele	Bassett	Chariho HS	Classroom Teacher
Marie	Bernier	Ella Risk School	Classroom Teacher
Michaeline	Bouchard	Metcalf School	Classroom Teacher
Jill	Burke	Chariho HS	English Language Arts Teacher
Sarah	Croteau	Leo A. Savoie Elementary School	Classroom Teacher
Nicole	Dantas	Pawtucket School Department	Elementary Mathematics Coach
Corinne	Ferri	North Smithfield Elementary	Classroom Teacher
Susan	Friendson	Central HS	English Language Arts Teacher
Arlene	Hall	Central Falls School Department	Mathematics Coach
Bonney	Henschel	Chariho Middle School	Classroom Teacher
Kathy	Kennhagan	Pawtucket School Department	Mathematics Coach
Melissa	Kerins	J.H. Gaudet Middles School	Title 1 Teacher, Mathematics Coach
Karen	Kment	North Smithfield Elementary School	Classroom Teacher
Justene	Loiselle	Cumberland HS	Classroom Teacher
Karen	Luth	West Glocester Elementary School	Mathematics Coach
Kim	Marchwicki	Alan Shawn Feinstein Middle School	Special Education Teacher
Stephen	Marnik	Dr. E.A. Ricci Middle School	Classroom Teacher
Cheryl	McElroy	Alan Shawn Feinstein Middle School	Classroom Teacher
Jeff	Miner	Toll Gate HS	Department Chair
Laurie	Mokaba	Fogarty Memorial School	Mathematic Coach
Renee	Palazzo	Smithfield HS	Classroom Teacher
Lora	Palys	Feinstein Middle School	Classroom Teacher
Frank	Piccirilli	West Broadway Elementary	Principal
Claire	Pollard	Central Falls School Department	Mathematics Consultant
Patricia	Pora	Leo A. Savoie Elementary School	Classroom Teacher
Maria	Rollin	Aldrich Junior High	Classroom Teacher
Monique	Rousselle-Condon	West Warwick HS	Mathematics Teacher
Helene	Scola	A. S. Feinstein Middle School	Classroom Teacher
Amy	Simoes	Halliwell Memorial School	Classroom Teacher
Sharon	Solway	Mt. Hope HS	English Language Arts Teacher
Jeff	Taylor	Aldrich Jr. HS	Classroom Teacher

VERMONT ITEM REVIEW COMMITTEE MARCH 18 & 19, 2008

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Carol	Amos	Twinfield Union	Teacher and Mathematics Coordinator
Julie	Bacon	Deerfield Elem.	Grade 4 Teacher
Ellen	Cameron	Cavendish School	Grade 4 Teacher
Beth Ann	Drinker	Grafton School	District Reading Coach
Julie	Graham	Chamberlain School	Literacy Coach K-6
Jennifer	Harper	Cavendish Elem.	Netwk. Leader, VT Teacher/Year, Gr. 4 Teacher
Colleen	Healey	North Country UHS	Teacher 9-12
Karen	Herbert Mayne	Manchester Elem.	Math Coordinator
Stephannie	Hockenbury	Chamberlain School	Math Coach K-6
Linda	Horn	Leicester School	Teacher of grades 4 & ^
Emily	Knisley	Blue Mt. UHS	Language Arts, Grades 7-10
Perry	Lessing	Middlebury UHS	HS Math Teacher
Suzanne	McDevitt	Browns River Middle School	Mathematics Teacher
Beth	Miller	Hartwick Elem.	Grade 4-5 Teacher
Kris	Muleller	Addison NW SU	District Math Coach
Lynn	Murphy	Waits River MS	Science Teacher 7-8, Native American
Martina	O'Donnell	Randolph UHS	Eng Dept. Chair, Teacher Gr. 10
Nancy	Philips	Warren Elementary	Literacy Coach
Travis	Redman	Rutland Town School	Grades 6 & 8 Teacher
Carol	Royce	Orange East SU	District Literacy Teacher Leader
Amy	S. Thivierge	Warren Elem.	Netwk. Leader, VT Teacher/Year, Gr. 4 Teacher
Katie	Sullivan	Mt. Anthony UHS	Teacher 11-12
Cherrie	Torrey	Dothan Brook Elementary	Reading Teacher
Susan	Weed	Sutton School	Lang. Arts Teacher, Grades 7 & 8
Loretta	Whitehead	Lyndon Elem.	5-6 Teacher
Kathleen	Wright	Manchester Elem.	Literacy Coordinator, K-8
Jennifer	Barone	Lyndon Town School	Grade 5/ School Literacy Coach
Ron	Murdoch	Hartford MS	MS Math Teacher, grades 7 & 8
John	Pandolfo	Stowe HS	Mathematics Teacher & Department Head
John	Tague	BFA Fairfax HS	HS Math Teacher
Mariann	Thompson	Dothan Brook Elementary	Teacher 4-5
Robert	Wing	North Country UHS	Teacher 10-11

BIAS AND SENSITIVITY COMMITTEE
MARCH 18 & 19, 2009

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Gina	Bell	Hillside Middles School	Mathematics/Special Education/ESL
Suzanne	Bergman	Winnisquam Regional Middle School	Enrichment Coordinator
Diane	Bush	Jaffrey-Rindge Middle School	Counselor
Enchi	Chen	Farmington High School	ESL Teacher
Ashley	Meehan	James Mastrocola Upper Elementary School	Teacher
Mary	Sohm	Londonderry HS	Special Educator

RHODE ISLAND

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Lisa	Bianco	E.W. Flynn Elementary School	Teacher
Adam	Flynn	Davies Career and Tech. Center	Classroom Teacher
Marice Ann	Piquette	Thompson Middle School	English Language Literacy Teacher
Kathleen	Pora	Harris School	Reading Specialist
Renay	Sawyer	Citizen's Memorial School	Teacher
Carolyn	Taylor	Beacon Charter High School	Teacher

VERMONT

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Cathy	Knight	Albert Bridge School	Principal/ Teacher
Cathy	Newton	Dothan Brook School	Special Educator
Pam	Parro	Hartwick Elementary	Special Educator
Darlene	Petke	Central Elem. School	Intensive Special Education
Rebekah	Thomas	John J. Flynn Elementary	ELL Teacher
Linda	Waisleben	Burlington School District	ELL Leader for District

BIAS AND SENSITIVITY COMMITTEE
NOVEMBER 5 & 6, 2008

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Diane	Bush	Jaffrey Rindge Middle School	School Counselor
Emilie	Carter	James Mastricola Elementary School	Assistant Principal
Enchi	Chen	Farmington Senior HS	ESL Teacher
Karen	Dow	Southwick Intermediate School	Reading Specialist
Mary	Sohm	Londonderry High School	Special Educator

RHODE ISLAND

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Christine	Cipolla	Wm. M. Davies Jr. Career-Technical High School	Reading Specialist
Heather	Forman	Exeter-West Greenwich Junior HS	English Lanuguae Arts Teacher
Holly	Gray	Citizens Memorial School	4th grade inclusion
Linda	Guarino	Northern Lincoln Elementary	Intensive Reading Specialist
Jeff	Miner	Toll Gate High School	English Department Chair
Marice Ann	Piquette	Thompson Middle School	Special Education Teacher

VERMONT

First Name	*Last Name*	*School/Association Affiliation*	*Position*
Barbara	Dall	Barre City Elementary and Middle School	ELL Teacher
Sharon	Hunt	Gilman Middle School	Special Educator
Maria	Lamson	Chelsea School	Librarian
Lynn	Murphy	Waits River Valley School	Science Teacher
Darlene	Petke	Central Elementary School	Intensive Special Education
Rebekah	Thomas	John J Flynn Elementary School	ELL Teacher

Appendix B—Accommodation Frequencies by Content Area

Table B-1. 2008–09 NECAP: Accommodation Frequencies by Content—Grades 3–5

Accommodation	*Grade 3*		*Grade 4*		*Grade 5*		
	Mathematics	*Reading*	*Mathematics*	*Reading*	*Mathematics*	*Reading*	*Writing*
A01	746	796	757	766	705	713	663
A02	4014	3972	4203	4111	4324	4254	4218
A03	1453	1470	1482	1496	1354	1371	1322
A04	254	264	261	278	256	254	257
A05	12	14	17	17	11	12	12
A06	21	17	10	10	14	12	14
A07	1386	1379	1435	1444	1580	1587	1573
A08	1576	1580	1579	1541	1388	1354	1325
A09	7	9	5	4	18	16	16
B01	257	263	252	258	236	243	230
B02	2357	2415	2356	2393	2361	2364	2290
B03	2194	2193	2298	2259	2852	2719	2567
C01	1	1	2	2	3	2	2
C02	32	36	46	48	40	39	41
C03	12	13	14	15	8	4	5
C04	3656	0	3400	0	3225	0	3020
C05	616	869	645	803	506	659	470
C06	66	35	53	35	32	23	21
C07	576	611	630	663	540	596	511
C08	10	16	9	12	8	10	9
C09	211	275	209	267	146	208	132
C10	16	26	18	25	15	23	12
C11	65	82	45	55	57	63	61
C12	13	0	29	0	44	0	24
C13	3	0	0	0	1	0	0
D01	13	20	17	27	60	94	155
D02	43	61	34	51	73	87	104
D03	4	8	2	3	3	3	3
D04	127	135	70	68	119	118	97
D05	983	1079	949	1077	871	1012	0
D06	20	26	26	30	32	36	0
E01	5	5	4	4	2	1	2
E02	0	0	0	0	0	0	32
F01	20	0	12	0	5	0	0
F02	0	15	0	14	0	5	0
F03	2	3	1	1	3	2	8

Table B-2. 2008–09 NECAP: Accommodation Frequencies by Content area—Grades 6–11

Accommodation	*Grade 6*		*Grade 7*		*Grade 8*			Grade 11		
	Mathematics	*Reading*	*Mathematics*	*Reading*	*Mathematics*	*Reading*	*Writing*	*Mathematics*	*Reading*	*Writing*
A01	552	556	426	445	399	391	365	293	282	288
A02	3960	3966	3682	3727	3654	3659	3527	2427	2442	2484
A03	1020	1041	790	808	569	580	539	352	356	374
A04	198	207	305	333	238	246	232	85	86	87
A05	8	8	7	12	8	9	8	4	3	3
A06	7	7	4	8	5	5	5	4	2	2
A07	1505	1528	1477	1500	1339	1379	1303	1198	1209	1209
A08	755	766	616	605	521	503	510	266	263	261
A09	14	13	16	5	8	7	7	13	11	13
B01	195	196	188	198	126	136	127	46	46	45
B02	1810	1872	1652	1679	1401	1408	1319	664	681	671
B03	2280	2270	2361	2266	2151	2030	1843	1687	1306	1222
C01	3	3	3	3	0	0	0	3	3	3
C02	34	34	28	29	15	17	15	13	12	14
C03	18	18	12	12	17	16	16	10	10	11
C04	2261	0	1650	0	1504	0	1464	578	0	645
C05	333	415	188	202	99	114	96	7	16	12
C06	32	24	43	25	31	21	20	19	15	19
C07	481	496	416	445	396	409	396	75	81	90
C08	10	20	5	5	4	4	4	1	1	2
C09	48	54	33	46	24	28	24	7	8	8
C10	4	5	0	1	1	1	1	0	0	1
C11	35	35	16	15	17	17	18	2	2	3
C12	55	0	63	0	38	0	39	39	0	42
C13	1	0	0	0	0	0	0	0	0	0
D01	65	127	62	145	110	201	254	23	56	92
D02	59	79	21	37	35	49	53	15	23	42
D03	2	2	4	3	1	1	2	1	0	1
D04	64	63	47	58	49	55	44	15	15	13
D05	548	713	302	406	239	286	0	41	61	0
D06	24	31	9	14	6	10	0	3	4	0
E01	10	10	5	5	0	0	1	0	0	2
E02	0	0	0	0	0	0	18	0	0	20
F01	39	0	33	0	33	0	0	177	0	0
F02	0	12	0	7	0	9	0	0	23	0
F03	7	8	0	1	0	1	2	1	1	6

Appendix C—Table of Standard Test Accommodations

Table of Standard Test Accommodations

Any accommodation(s) utilized for the assessment of individual students shall be the result of a formal or informal team decision made at the local level. Accommodations are available to all students on the basis of individual need, regardless of disability status.

A. Alternative Settings

- A-1 Administer the test individually in a separate location
- A-2 Administer the test to a small group in a separate location
- A-3 Administer the test in locations with minimal distractions (e.g., study carrel or different room from rest of class)
- A-4 Preferential seating (e.g., front of room)
- A-5 Provide special acoustics
- A-6 Provide special lighting or furniture
- A-7 Administer the test with special education personnel
- A-8 Administer the test with other school personnel known to the student
- A-9 Administer the test with school personnel at a non-school setting

B. Scheduling and Timing

- B-1 Administer the test at the time of day that takes into account the student's medical needs or learning style
- B-2 Allow short supervised breaks during testing
- B-3 Allow extended time, beyond recommended until in the administrator's judgment the student can no longer sustain the activity

C. Presentation Formats

- C-1 Braille
- C-2 Large-print version
- C-3 Sign directions to student
- C-4 Test and directions read aloud to student (Math, Science, and Writing only) [1]
- C-5 Student reads test and directions aloud to self
- C-6 Translate directions into other language
- C-7 Underlining key information in directions
- C-8 Visual magnification devices
- C-9 Reduction of visual print by blocking or other techniques
- C-10 Acetate shield
- C-11 Auditory amplification device or noise buffers
- C-12 Word-to-word translation dictionary, non-electronic with no definitions (For ELL students in Math, Science, and Writing only)
- C-13 Abacus use for student with severe visual impairment or blindness (Mathematics and Science – any session)

D. Response Formats

- D-1 Student writes using word processor, typewriter, computer [2] (School personnel transcribes student responses exactly as written into the Student Answer Booklet.)
- D-2 Student hand writes responses on separate paper. (School personnel transcribes student responses exactly as written into the Student Answer Booklet.)
- D-3 Student writes using brailler (School personnel transcribes student responses exactly as written into the Student Answer Booklet.)
- D-4 Student indicates response to multiple-choice items. (School personnel records student responses into the Student Answer Booklet.)
- D-5 Student dictates constructed responses (Reading, Math, and Science only) or observations (during the Science Inquiry Task) to school personnel. (School personnel scribes student responses exactly as dictated into the Student Answer Booklet.)
- D-6 Student dictates constructed responses (Reading, Math, and Science only) or observations (during the Science Inquiry Task) using assistive technology. (School personnel transcribes student response exactly as written into the Student Answer Booklet.)

If an accommodation is needed for a student that is not listed above, please contact the state personnel for accommodations to discuss it.

E. Other Accommodations [3]

- E-1 Accommodations team requested other accommodation not on list and DOE approved as comparable
- E-2 Scribing the Writing Test (only for students requiring special consideration)

F. Modifications [4]

- F-1 Using calculator and/or manipulatives on Session 1 of the Mathematics test or using a scientific or graphing calculator on Session 3 of the Science test.
- F-2 Reading the Reading test
- F-3 Other

1. Reading the reading test to the student invalidates all reading sessions.
2. Spell and grammar checks must be turned off. This accommodation is intended for unique individual needs, not an entire class
3. Test coordinators must obtain approval for the accommodation from the Department of Education prior to test administration.
4. All affected sessions using these modifications are counted as incorrect.

Appendix D—APPROPRIATENESS OF THE ACCOMMODATIONS ALLOWED IN NECAP GENERAL ASSESSMENT AND THEIR IMPACT ON STUDENT RESULTS

The New England Common Assessment Program
New Hampshire + Rhode Island + Vermont

Appropriateness of the Accommodations Allowed in NECAP General Assessment and Their Impact on Student Results

Overview & Purpose:

To meet federal peer review requirements for approval of state assessment systems, in the spring of 2006 New Hampshire, Rhode Island and Vermont submitted extensive documentation to the United States Department of Education on the design, implementation and technical adequacy of the New England Common Assessment Program (NECAP), a state level achievement assessment program developed through the collaborative effort of the three states. In response to peer review findings, the states were required to submit additional documentation for a second round of peer review, including information on the use, appropriateness, and impact of NECAP accommodations. This report was prepared in response to the questions posed by the peer reviewers, and has been included in the 2008 NECAP Technical Report for other groups or individuals who may be interested in NECAP accommodation policies and procedures, and how well they have been working.

Report on the Appropriateness and Comparability of Accommodations allowed in statewide NECAP General Assessment

A. Who may use accommodations in NECAP assessment?

NECAP test accommodations are available to *all* students, regardless of whether or not a disability has been identified. Allowable sccommodations are not group specific. For example, students in Title I reading programs, though not formally identified as disabled, may have an additional concern such as a broken arm and would therefore need to dictate multiple choice responses. Other students may need low vision accommodations even though they are not considered to be "blind." Before they are members of any subgroup, each student is first an individual with unique learning needs. NECAP assessment accommodations policy treats students in this way. The decision to allow *all* students to use accommodations, as needed, is consistent with prior research on best practice in the provision of accommodations (c.f., Elbaum, Aguelles, Campbell, & Saleh, 2004):

"…the challenge of assigning the most effective and appropriate testing accommodations for students with disabilities, like that of designing the most effective and appropriate instructional programs for these

students, is unlikely to be successfully addressed by disability. Instead, much more attention will need to be paid to individual student's characteristics and responses to accommodations in relation to particular types of testing and testing situations."(pp. 71-87)

The NECAP management team believes strongly that a fair and valid path of access to a universally designed test should not require that a student carry a label of disability. Rather, much like differentiated instruction, accommodated conditions of test participation that *preserve the essential construct of the standard being assessed* should be supported for *any* student who has been shown to need these differentiated test conditions. This philosophy is consistent with the NECAP management team's commitment to building a universally accessible test that provides an accurate measure of what each student knows in reading and mathematics content.

The following critical variables drive the process of providing NECAP accommodations:

1. The decision to use an accommodation for an individual student must be made using a valid and carefully structured team process consistent with daily instructional practice, and
2. The accommodated test condition *must preserve the essential construct being assessed*, resulting in a criterion-referenced measure of competency considered to be comparable to that produced under standard test conditions.

B. Are NECAP Accommodations Consistent with Accepted Best Practice?

NECAP provides a Table of Standard Test Accommodations that was assembled from the experience and long assessment histories of the three partner states. The NECAP Table of Standard Accommodations was created by establishing a three state cross-disciplinary consensus with key expert groups: special educators, ELL specialists, and reading, writing and mathematics content specialists from each of the partner states.

In addition, the work of various stakeholder and research groups with special instructional expertise was also considered. These sources included:

- Meetings with state advocacy groups for students with severe visual impairment or blindness,
- Meetings with state advocacy groups for students with deafness or hearing impairment, and consultations with other research-based groups like:
- The American Printing House for the Blind, Accessible Tests Division,
- The National Center on Educational Outcomes (NCEO), and
- The New England Compact Group, who conducted federally-funded enhanced assessment research on accommodations, in partnership with Boston College (inTASC group) and the Center for Applied Special Technologies (CAST).

The NECAP cross-disciplinary team, consulting with these other specialists, chose accommodations that were commonly accepted as standard, well established on a national basis, and that were consistent with assessment practice across all the NECAP states. Each identified standard accommodation was chosen to support best educational practice as it is currently understood.

Examples of the impact on accommodations design resulting from consultation with the American Printing House for the Blind experts in accessible test development included the addition of the use of an abacus in place of scrap paper to the for students with severe visual impairment. Recent research from the American Printing House for the Blind also indicated that 20 pt. font was producing better outcomes for students using large print accommodations (Personal communication, October, 2004). Based on this input, the NECAP team decided to provide a minimum of 20 pt. instead of 18 point font for large print editions of the NECAP assessment. This, in turn, led to improved production and type setting for large print NECAP tests. Consultation with advocacy groups for the deaf and hard of hearing led to improved item design, in particular helping item developers avoid the unnecessary use of rhyming words and homophones, supporting a decreased need for sign language accommodations with this group.

Impact of WIDA Partnership on development of Accommodations for LEP students. An important relationship exists between NECAP assessment and the NECAP partner states' active membership in WIDA/ACCESS for ELL's Assessment Consortium. New understandings in the area of accommodations policy and practice are beginning to emerge. For example, we have learned that word-to-word dictionary accommodations are most effective when used by LEP students at an intermediate level of proficiency while they are not advised for beginning LEP students. Learning opportunities created through the WIDA partnership have established a strong and supportive context for long term benefit and mutual growth potential.

During the last two years, assessment leaders from the NECAP states, as active partners in the WIDA consortium developing the new ACCESS for ELLs Test of English Language Proficiency, have collaborated in a cross-disciplinary team process to establish accommodations policy for this English language proficiency assessment. The ACCESS for ELLs accommodations team was composed of ESOL teachers, special educators, measurement specialists, and SEA assessment leaders. All three NECAP states took an active role and learned much from this process. This joint development effort opened dialog across ELL and special education accommodation groups and continues to support the ongoing review and improvement of both ACCESS and NECAP accommodations. The states are learning from each other, and with each new development cycle they are improving the accommodations system. The community of professional practice in this area is growing. Best practice understandings are expanding with the increasing experience and additional communication about the needs of LEP student groups. Specifically, the states have learned about the importance of academic language to English Language Learners who are attempting to take the state-level general content assessments. Accommodations specific to the academic language support issue are being explored and considered. The states have found that vocabulary lists, practice tests, computer-based read-alouds and other supports and accommodations elicit positive responses from LEP students who take the state content assessments. This will be addressed in greater detail in a later section.

C. How are NECAP Accommodations Structured?

Standard Accommodations: NECAP divides standard accommodations into four categories (labeled A-D), which include: A) Alternative Settings, B) Scheduling and Timing, C) Presentation Formats, and D) Response Formats. School teams may choose any combination of standard (A-D) accommodations to use with any student so long as proper accommodation selection and usage procedure is followed and documented (see following subsection). Students who use standard accommodations on NECAP tests receive full performance credit as earned for the test items taken under these standard conditions. NECAP standard accommodations are treated as fully comparable to test conditions where no accommodation is used.

In addition, NECAP lists two additional categories of altered test conditions which require formal state level review and approval on a student by student basis. These special test conditions are: E) Other Accommodations and F) Modifications. (See: NECAP Accommodations, Guidelines and Procedures Training Manual, (2005), p 5, Available on state websites, listed following references.)

Non-Standard Test Conditions – Review, Monitoring and Documentation of Preservation of the Intended Construct: "Other (E type) Accommodations" are accommodations without long or wide history of use that are not listed under the standard (A-D) categories. If schools wish to use accommodations that are not listed in A-D as standard, then they must send a formal written *Request for Use of Other Accommodations* to the state department of education for review and approval of usage with an individual student. This request documents the team's decision and describes fully the procedure to be used. Upon receipt by the SEA, these requests are thoroughly reviewed by state assessment content specialists together with special educators to determine if the accommodation proposed will allow performance of the essential constructs intended by the impacted test items. If the requested "other" accommodation is found to allow performance that will *not alter* the intended construct or criterion referenced standard to be assessed, then the school is issued a written receipt giving permission for use of this other accommodation as a standard accommodation for one test cycle. Schools are instructed on how to document the use of this approved "E) Other Accommodation" and the SEA monitors the process, ensuring that both school test booklets and state records accurately reflect the final test data. All "E) Other Accommodations" are approved in this way by the Department and, *if approved*, are treated as standard accommodations. Item responses completed under approved "E) Other" test conditions receive full credit as earned by the student.

If a requested "other" accommodation is found by the state review team to NOT preserve the intended construct, then the review team sends the school a receipt and notice that the requested change in test condition will be considered to be a test modification "F) Modification". All items completed under these test conditions will NOT receive performance credit. An example of a non-credited "F) Modification" would be any test condition where reading test passages, items, or response options are read to a student. State reading content specialists have determined that this change in a reading test condition does, in fact, alter the decoding construct being tested in all reading items. Therefore, reading items completed under this test condition would not be credited.

Use and approval of "E) Other Accommodations" are carefully monitored by the state. If any school claims use of an "E) Other Accommodation" that has not received prior state review and documented approval, then the test data documentation is similarly flagged to reflect that an F) Modification was instead provided. This flagged situation is treated as a non-credited test modification and the items impacted are invalidated. Further, any sections of the test completed under "F) Modification" conditions are later documented in student reports as not credited due to the non-standard and non-comparable test administration conditions used.

D. How does the NECAP Structure Guide Appropriate Use of Accommodations by Schools?

In 2005, New Hampshire, Rhode Island, and Vermont collaborated on the *NECAP Accommodations Guidelines and Procedures Training Manual*. The guide was disseminated through a series of regional test coordinator's workshops, as well as additional professional development opportunities provided by the individual states, and was also posted on each states website. This tool was designed to provide schools with a structured and valid process for decision making regarding the selection and use of accommodations for students on statewide assessment. Prior studies have outlined assessment guidelines that maximize the participation of students with disabilities in large-scale assessment. The National Center on Educational Outcomes (NCEO), in Synthesis Report 25 (1996), presented a set of criteria that states should meet in providing guidelines to schools for using accommodations (pp. 13-14, and 25). The NCEO recommendations figured prominently in preparation of the NECAP accommodations guide.

The *NECAP Accommodations Guidelines and Procedures Training Manual* (2005) meets all seven of the criteria established by NCEO as follows:

1. The decision about accommodations is made by a team of educators who know the student's instructional needs. NECAP goes beyond this recommendation and requires that the student's parent or guardian also be part of this decision team, (NECAP Accommodations Manual, pp. 2-3, and 20-22).
2. The decision about accommodations is based on the student's current level of functioning and learning characteristics. (Manual, pp20-22).
3. A form is used that lists the variables to consider in making the accommodations decisions, and that documents for each student the decision and reasons for it. (Manual, pp. 20-22).
4. Accommodation guidelines require alignment of instructional accommodations and assessment accommodations. (Manual, pp2 and 20-22).
5. Decisions about accommodations are not based on program setting, category of disability, percent time in the mainstream classroom (Manual, p.15, p.20-22).
6. Decisions about accommodations are documented on the student's IEP or on an additional form that is attached to the IEP. (Manual, pp.2, 15, and 20-22).
7. Parents are informed about accommodation options and about the implications for their child (1) not being allowed to use the needed accommodations, or (2) being excluded from the accountability system when certain accommodations are used, (Manual pp 3 and 20-22).

As described above**,** NECAP states use a highly structured process for the review, approval, and monitoring of requests by schools for the use of other (non-standard) accommodations for individual students. As described in section B, above, the NECAP Accommodations Manual provides a Table of Standard Accommodations each year. The manual provides two structured decision making worksheets (pp. 20-22) to guide the decision process of educational teams. One worksheet guides the selection of standard accommodations; the second provides guidance on the selection of other accommodations. The manual contains information on the entire decision making process. In addition, the manual provides detailed descriptions and research-based information on many specific accommodations.

Ongoing Teacher Training and Support: Throughout each academic year, several teacher workshops on planning and implementing accommodations are offered at multiple locations regionally in each of the three states to teams of educators. In the spring of 2005, prior to the launch of the first NECAP assessment, a series of introductory statewide 2-hour workshops in accommodations administration was offered in multiple locations. Each year thereafter, in late summer prior to the administration of the NECAP tests, a series of accommodations usage updates is offered as part of the NECAP Test Administration Workshop series; five regional workshops are offered in each state. Additionally, each state's Department of Education has consultants who are available to provide individualized support and problem solving, as well as small and large group in-service for schools. Finally, the DOE assessment consultants work directly with a variety of statewide groups and organizations to promote the use of effective accommodations, and to gather feedback on the efficacy of the NECAP accommodation policies and procedures. These include university-based Disability Centers, statewide parent advocacy organizations, organizations representing individuals with vision impairments and hearing disabilities. Finally, each state has systems in place to provide schools with individualized support and consultation: New Hampshire employs two distinguished special field educators who, by appointment and free of charge, provide onsite training and support in alternate assessment and accommodations strategies. Rhode Island has an IEP Network that provides on-site consultation with schools on a variety of special services topics including planning and implementing assessment accommodations. Vermont has a cadre of district-level alternate assessment mentors who provide a point of contact for disseminating information, and who are also available in schools and school districts for intensive consultation related to the assessment needs of individual students.

Monitoring of the Use of Accommodations in the Field: Each year during the NECAP test window, the DOE content specialists schedule a limited number of on-site visitations to observe test administration as it is occurring in the schools. State capacity to provide such direct monitoring during the test window is limited, but such monitoring is conducted during each test window and observers report observations directly to the state assessment team. Additional on-site accommodations monitoring is provided by district special education directors and the NECAP test coordinators. Both of these groups also receive training each year. Throughout each school year, program review teams from the DOEs' special education divisions conduct on-site focused monitoring of all special education programs. These comprehensive visits

include on-site monitoring of the use of accommodations for students who have Individualized Educational Programs (IEPs).

E. Are NECAP Accommodations Consistent with Recent Research Findings?

The NECAP development team has attempted to learn from the research on accommodations, but this has not been a simple matter. In 2002, Thompson, Johnstone, and Thurlow concluded in their report on universal design in large scale assessments that research validating the use of standard and non-standard accommodations has yet to provide conclusive evidence about the influence of many accommodations on test scores. In 2006, Johnstone, Altman, Thurlow, & Thompson published an updated review of 49 research studies conducted between 2002 and 2004 on the use of accommodations and again found accommodations research to be inconclusive. They noted the similarity to past findings from NCEO summaries of research (Thompson, Blount & Thurlow, 2002). The authors of the 2006 review state:

> "Although accommodations research has been part of educational research for decades, it appears that it is still in its nascence. There is still much scientific disagreement on the effects, validity, and decision-making surrounding accommodations." (p 12)

However, a frequently cited research review by Sireci, Li, & Scarpati, (2005) documented evidence of support for the accommodation of providing extended time. This accommodation is one of the most frequently used standard NECAP accommodations. Extended time accommodations appeared to hold up best under the interaction hypothesis for judging the validity of an accommodation. In a 2006 presentation addressing lessons learned from the research on assessment accommodations to date, Sireci and Pitoniak, (2006), concluded that, in general, "accommodations being used are sensible and defensible." They replicated their prior finding that the extended time accommodation seems to be a valid accommodation and noted that many other accommodations have produced less convincing results. They noted that oral or read-aloud accommodation for math appears to be valid, but that a similar read-aloud accommodation for *reading* involves consideration of specific construct changes which threaten score comparability. These findings are also consistent with and support the NECAP accommodation policy of allowing the read-aloud accommodation for mathematics, but not allowing this accommodation for reading tests. Despite the inconclusive and conflicting current state of accommodations research, findings seem to be emerging that do, in fact, provide validation for some of the most frequently used NECAP accommodations: the extended time and mathematics read-aloud accommodations.

Accommodations for English language learners. In a presentation on the validity and effectiveness of accommodations for English language learners with disabilities, Abedi (2006) reported that students who use an English or bilingual dictionary accommodation (word meanings allowed) may be advantaged over those without access to dictionaries and that this may jeopardize the validity of the assessment. Abedi argues persuasively that linguistic accommodations for English language learners should *not* be allowed to alter the construct being tested. He also argues that the language of assessment should be the same language as that

used in instruction in the classroom – otherwise student performance is hindered. NECAP assessment policy is consistent with both of these findings: ELL students may use word-to-word translations as linguistic accommodation support, but may not use dictionaries with definitions provided. Abedi's research supports this decision. Also NECAP assessment items are not translated into primary languages for ELL students. This, too, is consistent with classroom practice in the NECAP states and is supported by the current literature.

At the same conference referenced just above, Frances (2006), presented findings from a meta-analysis in which he compared the results of eleven studies of the use of linguistic accommodations provided for ELL students in large scale assessments. In his presentation, given at the LEP Partnership Meeting in Washington, DC, he noted that *no significant differences in student performance were observed for 7 of the 8 most commonly provided linguistic accommodations.* Although Frances was not recommending its use, the *only* linguistic accommodation that showed any significant positive effect on the performance of ELL students was an accommodation allowing the use of an English dictionary or glossary during statewide assessment. This is the very same accommodation that Abedi (2006) recommends *against* using because it violates intended test constructs. As noted above, in NECAP assessment, the use of word-to-word translations is an allowed standard linguistic accommodation. However, the use of an English dictionary with glossary meanings *is not* an allowable standard accommodation. It is the position of the NECAP reading content team that allowing *any* student to use a dictionary with definitions or a glossary of meanings violates the vocabulary and comprehension constructs intended in the NECAP reading test and would invalidate test results. For this reason, NECAP does not allow this linguistic accommodation.

As reported by Frances, analysis of the remaining 7 linguistic accommodations typically allowed for ELL students showed *no significant positive effect* on test performance. These included: bilingual dictionary use, dual language booklets, dual language questions and read-aloud in Spanish, extra time to test, simplified English, and offering a Spanish version of a test. Despite the lack of positive effects observed for these other linguistic accommodations to date, NECAP does provide a number of linguistic supports for ELL students. One of these linguistic supports includes: employing the universal design technique of simplifying the English in *all* test items. Review and editing of test items for language simplicity and clarity has been a formal part of the annual process of test item development and review since the inception of the NECAP. In addition to word-to-word translations, a number of other standard linguistic accommodations are allowed in NECAP testing to provide a path of access for ELL students to show what they know and can do in reading and mathematics. Standard linguistic accommodations permitted by NECAP include: allowing mathematics test items to be read aloud to the student, allowing students to read aloud to themselves (if bundled with an individual test setting), translation of test directions into primary language, underlining key information in written directions and dictation/ scribing of reading and math test responses. NECAP assessments provide linguistic access for students who are English language learners.

As noted earlier, a number of studies have shown some positive effect of the use of the extended time and read-aloud accommodations for students in general. As ELL students continue to gain proficiency in

English, they may also increasingly benefit from these accommodations. More research is needed to clarify how states can most appropriately support ELL students to show us what they know and can do.

NECAP Supported Research Studies: Through the New England Compact Enhanced Assessment Project (2007), the NECAP states have completed a number of accommodations and universal design research studies. These studies have shed additional light on the appropriateness of existing standard accommodations and have helped to inform the development of new accommodations and improved universal design of assessment. Under the Enhanced Assessment Grant, in joint partnership with: the inTASC group of Boston College, the Center for Applied Special Technologies (CAST), the state of Maine, and the Educational Development Center, Inc., the NECAP states supported research studies on accommodations and universal design in four distinct areas. These studies, summarized below, are described more fully in the appendix to this report:

Use of computer-based read-aloud tools. NECAP supported a study of 274 students in New Hampshire high schools. This study, Miranda, H., Russell, M., Seeley, K., Hoffman, T., (2004), provided evidence that computer–based read aloud accommodations led to improved content access and performance of students with disabilities when taking mathematics tests.

As direct result of this study, New Hampshire was able to build and pilot a new computer-based read aloud tool that is now under development for use with NECAP assessments for all three NECAP states. Following this New Hampshire pilot of the new computer-based read aloud tool on the state high school assessment, the New Hampshire Department of Education conducted a focus group study with participating students from Nashua North High School. The results of this focus group (May 17, 2006) are available from the New Hampshire Department of Education. One of the primary findings from this focus group was the strong impact of having experienced the read-aloud in practice test format prior to actual testing. Experience with this tool *prior to testing* appeared to be very important for student performance. High school students indicated a *very strong* preference for computer-based read aloud over the same accommodation provided by a person. Both groups of students, those with limited English proficiency and those with disabilities consistently reported that they were able to focus much more clearly on the math content (not just the words) than in prior math tests they had taken without this accommodation. Based on student reports, use of this read-aloud seemed to improve content access for these students. The ability to benefit from the individual work of each of the three NECAP states is a major benefit of the tri-state partnership.

- **Use of computers to improve student writing performance on tests.** Another research study conducted by Higgins, J., Russell, M., & Hoffmann, T., (2004), studied 1000 students from the three states to examine how the use of computers for writing tests affected student performance. The study found that minority girls tended to perform about the same whether using a computer or pencil-and-paper to provide written responses. However, *all other groups*, on average, tended to perform better when using a computer to produce written responses. A minimum degree of

keyboarding skill correlated with improved performance. Lack of keyboarding skill produced results that did not significantly differ from pencil-and–paper responding and therefore, appeared to 'do no harm'. As a result, NECAP states entered into talks to determine how a computer based response might be more fully supported in future versions of the assessment. The study suggested that a minimum number of words (18-20) typed accurately per minute was the recommended threshold to obtain benefit from this accommodation. This finding has been incorporated into NECAP training and support activities. At the present time, NECAP allows use of a word processor to produce written test responses as a standard accommodation on all NECAP content tests. The research supports this practice.

- **Use of Computers for Reading Tests**. A third study conducted by Miranda, H., Russell, M., & Hoffmann, T., (2004), examined how the presentation of reading passages via computer screen impacted the test performance of 219 fourth grade students from eight schools in Vermont. This study found no significant differences in reading comprehension scores across the 3 (silent) presentation modes studied: 1. Standard presentation on paper, 2. On computer screen with use of a scrolling feature, and 3. On computer with passages divided into sections presented as whole pages without the scrolling feature. Results from this study were not conclusive, but some trend data suggested that the scrolling presentation feature may disadvantage many students, especially those with weaker computer skills. The majority of students indicated an overall preference for computer-based presentation over pencil-and-paper. As other research studies, previously cited, continue to show that read-aloud accommodations are generally effective, it can be expected that pressure to offer computer-based read-alouds involving text presentation will increase. Additional research in this area may help shed important light on the most effective ways to provide this useful accommodation. (See also: Higgins, J., Russell, M., & Hoffmann, T., (2004).)
- **Use of Computer-Based Speak-Aloud Responses to Short Answer Items**. The states' enhanced assessment grant also supported a study by Miranda, H., Russell, M., Seeley, K., Hoffman, T., (2004) that looked at the feasibility and effectiveness of using a computer to transcribe spoken responses into written text in response to short answer test items. This was considered as a possible linguistic accommodation for use with English language learners in reading and mathematics tests. Unfortunately, this study found that it is not yet feasible to use computers to record student's verbal responses to short-answer items. A variety of technical problems occurred and students were not comfortable speaking to the computer. The researchers concluded that, with existing technology limitations, use of this kind of computer based accommodation may not be feasible for some years.
- What evidence has the state gathered on the impact and comparability of accommodations allowed on NECAP test scores?

Direct and Immediate Score Impact. First, as a matter of policy, there is a direct and immediate impact on NECAP test scores for students when standard accommodations (accepted *and credited* as comparable) vs. non-standard accommodations (not accepted *and not credited* as comparable) are used during test administration. The student performance score is significantly reduced for each subtest where test items and the constructs they were designed to measure have been modified by use of a non-standard accommodation. Sessions with modified items receive no credit in the student total score for that content area. If the entire reading test is read to a student, the student will earn 0 points in that content area. If only certain sessions of the reading test are read to the student, then only the score of those sessions will be impacted, but this will result in a lower overall reading content score.

Empirical bases for Comparability of NECAP Test Scores Obtained from Accommodated vs. Non-Accommodated Test Conditions: During the NECAP Pilot Test in 2004, differential item functioning (DIF) analyses were conducted on the use of accommodations by various student subgroups. In December 2006, the NECAP Technical Advisory Committee (TAC) reviewed the use of these DIF analyses and discussed long range planning for ongoing review of the use of accommodations in NECAP assessment. There was consensus among TAC members that the current use of DIF analyses for evaluation of accommodation use allows very limited inferences to be made therefore is of minimal practical value to the states. Other general methods of organizing and reviewing accommodations data and performance outcomes should be developed for states to employ.

A NECAP TAC subgroup was formed to consider and respond to the following question: What should NECAP states be doing at this stage in our development to review use, appropriateness, design, etc, of the NECAP Accommodations and related policy & guidelines? What information and processes will help us learn, clarify & communicate how, why, and when to use what accommodations? The results of this December 2006 TAC accommodations workgroup are available on each of the three states' websites. In summary, the TAC workgroup recommended 5 categories of activity for the NECAP states:

1. Given what states have learned from initial implementation and recent research, they should review, revise, describe and more fully document NECAP Accommodations Policies and Guidelines. This should be part of an ongoing review process.
2. Explore available research on questionable or controversial accommodations. Document this review and revise where indicated.
3. Transparency of reporting should be examined. There was group consensus that the use of accommodations during assessment should be fully disclosed, and thereby made transparent in the reporting process. NECAP states should work to sort out this aspect of reporting policy and determine where and how to report what aspects of accommodation usage to parents and to the public at large.
4. States need to further address monitoring of accommodation usage. Find ways to improve the quality of district/school choices in the selection and use of accommodations for students. Strategies that take limited state resource capacity into account must be considered. The issue is fundamentally one of putting improved quality control processes in place in the most efficient, cost effective ways. Several resources currently under development may assist the states in this effort. One of these resources is already being

developed in the OSEP funded General Supervision Grant to one of the NECAP states. This grant will develop digitized video clips illustrating proper ways to provide certain accommodations, especially for students with severe disabilities. Creation of this video tool may enhance state capacity to provide and distribute effective training to districts and improved local monitoring of day to day use of accommodations for both instruction and assessment.

5. Available data needs to be mined and organized on the current use of accommodations in NECAP testing. Usage and outcomes for various subgroups should be examined. DIF analyses may not be as useful in this regard as other types of carefully planned descriptive comparisons.

Some research concerns were also identified. How do states differentiate between an access issue for a student where the student has skills they cannot show as opposed to a lack of opportunity to learn or lack of skill development? This issue appears repeatedly in a number of research studies reviewed. It is not a simple matter to differentiate between these situations. One indicates a need for an assessment design change. The other indicates a need for instructional change. Research to help sort this out should be supported.

F. Test Access Fairness as One Kind of Evidence for Comparability:

NECAP states have made a commitment to work with stakeholders representing various groups of students who typically use accommodations or who may benefit from improved universal assessment design. The feedback received from these stakeholder groups is a valuable source of information and ideas for continued improvement of our assessment program.

NECAP consults regularly with experts in accessible test design at the American Printing House for the Blind in Lexington, KY (Allman (2004), and Personal Communications: (October 2004), (September 2006)). This group has informed NECAP management about the recent research in the use of larger print fonts and the abacus as standard accommodations for students with severe visual impairments. This consultation has directly impacted test development and has resulted in positive feedback from the stakeholders who represent students with visual impairment in our states.

In addition, all three states work closely with stakeholders representing students with hearing impairment and deafness to help inform test item development and improved access to test items for students with vision or hearing impairments. An example of this commitment is contained in two focus group reports prepared by the New Hampshire Department of Education; a February 2006 focus group report from NH Teachers of the Visually Impaired (TVI) on NECAP Test Accessibility for Students with Severe Visual Impairment and a May 2006 report on the performance of English language learners and students with disabilities for the Grade 10 New Hampshire Educational Improvement & Assessment Program (NHEIAP). The latter of these two reports addressed computer-based read aloud accommodation for mathematics assessment. *(Both Focus Group Reports are available from the New Hampshire Department of Education).*

NECAP states are also pursuing other grant–funded research to support and explore development of new comparable accommodations that might provide meaningful access to general assessment at grade level for students who currently take only alternate assessments based on alternate achievement standards.

G. Summary of the Evidence - Are NECAP Accommodations Appropriate and Do They Yield Reasonably Comparable Results?

- Yes, it is clear from the evidence cited in sections 2 A, B, C and D above, that NECAP accommodations are highly consistent with established best practice.
- For accommodations with a consistent research basis available, research evidence suggests that continued use of the following accommodations in NECAP testing is valid:
    - Extended time accommodation
    - Mathematics Read-Aloud Accommodation
    - Word-to-word translation for ELL students
    - Use of Computer-Based Read-Aloud Tools (for mathematics)
    - Use of Computers to write extended test item responses (NECAP accommodation - D1)
- Preliminary research evidence from The New England Compact Enhanced Assessment Project, presented above (2004), does not appear to support improved student performance with NECAP accommodation D6- Using assistive technology (specifically speech-to-text technology) to dictate open responses via computer. However, if consistently used in classroom settings for students with severe access limitations, sufficient familiarity may be gained to make this a viable accommodation for certain students. Further review of this accommodation by the NECAP management team is recommended.
- Early focus group results (NHDOE, May 17, 2006) and trial experience with computer-based read aloud testing is very promising and merits further research.
- NECAP Focus group responses (NHDOE, February 22, 2006) from Teachers of the Visually Impaired support existing NECAP accommodations and are helping inform improvement in other aspects of universal design of items, test booklets and materials.
- Structured DIF analysis of the performance of NECAP accommodations is in an early and inconclusive phase. Currently, development of other increasingly useful accommodations data analysis designs is going forward and is supported by all NECAP states. The NECAP Technical Advisory Committee (TAC) will continue to explore this line of inquiry in the future.
- As each yearly cycle of large scale NECAP DIF item analysis allows the group to gain insight and to clarify questions, the design of future DIF data collection may be refined to more fully inform item selection to improve the fairness and accessibility of NECAP assessment items. This exploration is highly valued by the NECAP management group and will continue to be supported. Limitations in this kind of statistical analysis will continue to occur when sample sizes are too small to draw reliable or useful conclusions.
- NECAP states are developing an ongoing review and improvement process for the NECAP accommodations policy and procedures.

H. Concluding Comment:

NECAP Commitment to Universal Design and Continuous Improvement. The NECAP management group has made a solid commitment to continuously improve and strengthen the universal design of our assessment instruments. As the quality of universal design elements of the NECAP assessment continues to improve, it is conceivable that the number of students who need to use accommodations may decline. In fact, this is a worthy goal. Although this would cause diminishing sample sizes and challenges for accommodations analysis, declining use of accommodations due to improved universal accessibility in overall test design would be viewed as a very positive outcome.

Since its inception in 2003, the NECAP group has supported and funded research and development in accommodations policy and procedures. This is evidenced by the many research activities generated through the multiple Enhanced Assessment Grants of the three participating states referenced earlier in this report.

The NECAP group has shown leadership in obtaining funding and actively supporting accommodations and related research in a number of areas:

1. Describing the performance of students in the assessment gap and exploring alternate ways of assessing students performing below proficient levels (see: *New England Compact Enhanced Assessment Project: Task Module Assessment System- Closing the Gap in Assessments*),
2. Research in the design and use of accommodations (New England Compact Enhanced Assessment Project: Using Computers to Improve Test Design and Support Students with Disabilities and English-Language Learners),
3. The relationships among and between elements of English language proficiency test scores, academic language competency scores, and performance on NECAP academic content tests *(Parker, C. (2007)),*
4. Defining and developing technical adequacy in alternate assessments *(NHEAI Grant),*
5. Developing improved accommodations that will foster increased participation in general assessment for students currently alternately assessed *(Jorgensen & McSheehan, (2006)), and*
6. All three NECAP states are partners in the ongoing development of the new *ACCESS for ELLs™* Test of English Language Proficiency. The Vermont Test Director is a member of the Technical Advisory Committee

The NECAP Development Team has been very busy. These efforts are ongoing and will continue. They are committed to the long-term development of a well validated and highly accessible assessment program that meets the highest possible standards of quality. More importantly, we are committed to the establishment of an assessment system that effectively supports *the growth of each and every one* of our students.

REFERENCES

Abedi, J. (2006) *Validity, effectiveness and feasibility of accommodations for English language learners with disabilities (ELLWD).* Paper presented at the Accommodating Students with Disabilities on State Assessments: What Works Conference, Savannah, GA.

Allman, C.B., (Ed.). (2004) *Test Access: Making Tests Accessible for Students with Visual Impairments*. Louisville, KY: American Printing House for the Blind, Inc.

American Printing House for the Blind, Inc., Accessible Tests Division Staff, (personal communication, October 2004)

American Printing House for the Blind, Inc., Accessible Tests Division Staff, (personal communication, September 2006)

Dolan, R. (2004) Computer Accommodations Must Begin As Classroom Accommodation: The New England Compact Enhanced Assessment Project: Using Computers to Improve Test Design and Support Students with Disabilities and English-Language Learners. ©1994-2007 by Education Development Center, Inc. All Rights Reserved. http://www.necompact.org/research.asp

Elbaum, B., Aguelles, M.E., Campbell, Y., & Saleh, M.B. (2004). Effects of a student-reads-aloud accommodation on the performance of students with and without learning disabilities on a test of reading comprehension. *Exceptionality, 12(2), 71-87.*

Elliott, J., Thurlow, M., & Ysseldyke, J. (1996) Assessment guidelines that maximize the participation of students with disabilities in large-scale assessments: Characteristics and considerations, Synthesis report 25. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Frances, D.J. (2006). *Practical guidelines for the education of English language learners*. Paper presented at the 2006 LEP Partnership Meeting. Washington, DC. Presentation retrieved December 21, 2006, from http:// www.centeroninstruction.org.

Higgins, J., Russell, M., & Hoffmann, T., (2004) Examining the Effect of Computer-Based Passage Presentation on Reading Test Performance: Part of the New England Compact Enhanced Assessment Project. Boston, MA, in Technology Assessment Study Collaborative (inTASC), Boston College (http://www.bc.edu/research/intasc/publications.shtml)

Higgins, J., Russell, M., & Hoffmann, T., (2004) Examining the Effect of Text Editor and Robust Word Processor on Student Writing Test Performance: Part of the New England Compact Enhanced Assessment Project. Boston, MA, in Technology Assessment Study Collaborative (inTASC), Boston College (http://www.bc.edu/research/intasc/publications.shtml)

Johnstone, C.J, Altman, J., Thurlow, M.L., & Thompson, S.J. (2006): *A summary of research on the effects of test accommodations: 2002-2004: Synthesis Report 45.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Jorgensen, C. & McSheehan, M. (2006) *Beyond Access for Assessment Accommodations*, General Supervision Enhancement Grant Research (in progress) supported by the US Education Department, Office of Special Education Research, Washington, DC.

Miranda, H., Russell, M., & Hoffmann, T., (2004) Examining the Feasibility and Effect of a Computer-Based Read-Aloud Accommodation on Mathematics Test Performance: Part of the New England Compact Enhanced Assessment Project. Boston, MA, in Technology Assessment Study Collaborative (inTASC), Boston College (http://www.bc.edu/research/intasc/publications.shtml)

Miranda, H., Russell, M., Seeley, K., Hoffman, T., (2004) Examining the Feasibility and Effect of Computer-Based Verbal Response to Open-Ended Reading Comprehension Test Items: Part of the New England Compact Enhanced Assessment Project. Boston, MA, in Technology Assessment Study Collaborative (inTASC), Boston College (http://www.bc.edu/research/intasc/publications.shtml)

Parker, C. *Deepening Analysis of Large-Scale Assessment Data: Understanding the results for English language learners*, Study in progress (2007). Project funded by the U.S. Department of Education, Office of Educational Research and Improvement. **http://www.relnei.org**

Quenemoen, R. (2007). *New Hampshire Enhanced Assessment Initiative (NHEAI): Knowing What Students with Severe Cognitive Disabilities Know.* Research (in progress) supported by the US Education Department, Office of Elementary and Secondary Education, Washington, DC.

Sireci, S.G., Li, S., & Scarpati, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research*, 75 (4), 457-490.

Sireci, S.G. and Pitoniak, M.J. (2006). Assessment accommodations: What have we learned from research? Paper presented at the Accommodating Students with Disabilities on State Assessments: What Works Conference, Savannah, GA.

The New England Compact Enhanced Assessment Project: Using Computers to Improve Test Design and Support Students with Disabilities and English-Language Learners. ©1994-2007 by Education Development Center, Inc. All Rights Reserved. http://www.necompact.org/research.asp

The New England Compact Enhanced Assessment Project: Task Module Assessment System. ©1994-2007 by Education Development Center, Inc. All Rights Reserved. http://www.necompact.org/research.asp

Thompson, S.J., Blount, A., & Thurlow, M.L. (2002): *A summary of research on the effects of test accommodations 1999-2001, Technical Report 34.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thompson, S.J., Johnstone, C.J., & Thurlow, M.L. (2002): *Universal design applied to large-scale assessments: Synthesis Report 44.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Additional Resources:

Rhode Island Department of Education, NECAP Assessment Website:

http://www.ridoe.net/assessment/NECAP.aspx

Vermont Department of Education, NECAP Assessment Website: http://education.vermont.gov/new/html/pgm_assessment.html

New Hampshire Department of Education, NECAP Assessment Website: http://www.ed.state.nh.us/NECAP

Appendix E—Common Item Double Blind Interrater Agreement Percentage Tables

Table E-1. 2008–09 NECAP: Open Response
Interrater Agreement by Item–Mathematics Common Items

Grade	*Content*	*Item Number*	*% Exact + Adjacent*	*% Exact*	*% Adjacent*	*% >1*
3	Mathematics	13	100.0	99.6	0.4	0.0
		14	100.0	98.4	1.6	0.0
		16	100.0	99.1	0.9	0.0
		17	100.0	99.1	0.9	0.0
		18	99.8	94.0	5.8	0.2
		20	99.9	97.1	2.8	0.1
		21	99.9	97.0	2.8	0.1
		22	99.9	98.8	1.1	0.1
		23	100.0	98.2	1.8	0.0
		40	100.0	99.0	1.0	0.0
		41	100.0	99.7	0.3	0.0
		43	100.0	99.2	0.8	0.0
		44	100.0	96.0	4.0	0.0
		45	99.9	99.1	0.8	0.1
		47	100.0	95.0	5.0	0.0
		64	100.0	99.2	0.8	0.0
		66	100.0	99.8	0.2	0.0
		67	100.0	95.1	4.9	0.0
		68	100.0	98.8	1.2	0.0
		70	100.0	96.9	3.1	0.0
		Overall	100.0	97.9	2.1	0.0
4	Mathematics	13	100.0	98.1	1.9	0.0
		14	100.0	99.6	0.4	0.0
		16	100.0	99.8	0.2	0.0
		17	100.0	99.3	0.7	0.0
		18	99.8	96.5	3.3	0.2
		20	100.0	94.0	6.0	0.0
		21	100.0	98.2	1.8	0.0
		22	98.4	91.7	6.7	1.6
		23	99.8	94.3	5.5	0.2
		40	100.0	99.5	0.5	0.0
		41	100.0	99.6	0.4	0.0
		43	100.0	98.8	1.2	0.0
		44	100.0	98.8	1.2	0.0
		45	99.9	97.0	2.9	0.1
		47	99.8	95.5	4.3	0.2
		64	100.0	99.6	0.4	0.0
		66	100.0	98.7	1.3	0.0
		67	100.0	99.5	0.5	0.0
		68	99.9	97.2	2.7	0.1
		70	99.9	90.0	9.9	0.1
		Overall	99.9	97.3	2.5	0.1
5	Mathematics	14	100.0	97.0	3.0	0.0
		15	100.0	98.7	1.3	0.0
		17	95.5	93.1	2.4	4.5
		18	99.3	87.2	12.1	0.7
		21	99.5	92.3	7.2	0.5
		37	100.0	99.3	0.7	0.0
		40	99.0	93.3	5.8	1.0
		41	99.8	94.4	5.4	0.2
		42	98.9	90.1	8.8	1.1
		43	99.9	98.0	1.9	0.1
		58	100.0	93.8	6.2	0.0
		59	100.0	98.4	1.6	0.0

(cont'd)

Grade	Content	Item Number	% Exact + Adjacent	% Exact	% Adjacent	% >1
5	Mathematics	60	100.0	98.8	1.2	0.0
		62	98.1	87.5	10.6	1.9
		63	99.7	96.3	3.5	0.3
		65	99.7	96.7	2.9	0.3
		Overall	99.3	94.5	4.8	0.7
6	Mathematics	14	100.0	98.3	1.7	0.0
		15	100.0	98.7	1.3	0.0
		17	99.6	96.5	3.2	0.4
		18	98.4	89.9	8.6	1.6
		21	99.9	97.5	2.3	0.1
		37	100.0	97.4	2.6	0.0
		40	97.9	86.3	11.7	2.1
		41	99.6	94.9	4.7	0.4
		42	99.6	90.6	8.9	0.4
		43	100.0	97.2	2.8	0.0
		58	100.0	93.6	6.4	0.0
		59	100.0	98.7	1.3	0.0
		60	100.0	98.6	1.4	0.0
		62	99.4	91.4	8.0	0.6
		63	99.4	96.4	3.0	0.6
		65	99.9	96.3	3.6	0.1
		Overall	99.6	95.1	4.5	0.4
7	Mathematics	14	100.0	98.8	1.2	0.0
		15	100.0	99.5	0.5	0.0
		17	99.7	93.6	6.1	0.3
		18	99.6	91.2	8.4	0.4
		21	99.6	95.2	4.4	0.4
		37	100.0	97.5	2.5	0.0
		40	99.7	96.5	3.1	0.3
		41	99.7	95.6	4.1	0.3
		42	99.7	94.8	4.9	0.3
		43	99.7	94.4	5.2	0.3
		58	100.0	99.7	0.3	0.0
		59	100.0	98.7	1.3	0.0
		60	100.0	94.3	5.7	0.0
		62	99.1	86.9	12.2	0.9
		63	99.5	95.1	4.3	0.5
		65	100.0	97.5	2.5	0.0
		Overall	99.7	95.4	4.3	0.3
8	Mathematics	14	100.0	98.5	1.5	0.0
		15	100.0	98.9	1.1	0.0
		17	99.9	98.0	1.9	0.1
		18	98.5	89.3	9.3	1.5
		21	99.8	95.3	4.5	0.2
		37	100.0	95.4	4.6	0.0
		40	99.3	96.3	3.1	0.7
		41	100.0	97.5	2.5	0.0
		42	99.5	87.7	11.8	0.5
		43	99.8	97.5	2.4	0.2
		58	100.0	99.7	0.3	0.0
		59	100.0	93.9	6.1	0.0
		60	100.0	99.1	0.9	0.0
		62	99.0	96.3	2.7	1.0
		63	99.2	92.0	7.2	0.8
		65	99.8	97.5	2.4	0.2
		Overall	99.7	95.6	4.0	0.3

(cont'd)

Grade	Content	Item Number	% Exact + Adjacent	% Exact	% Adjacent	% >1
		17	100.0	99.7	0.3	0.0
		18	100.0	99.8	0.2	0.0
		19	100.0	99.5	0.5	0.0
		22	100.0	99.8	0.2	0.0
		23	100.0	99.8	0.2	0.0
		24	100.0	100.0	0.0	0.0
		25	99.9	97.5	2.4	0.1
		26	99.8	96.3	3.5	0.2
		29	99.9	99.4	0.5	0.1
		30	99.9	98.8	1.1	0.1
		31	100.0	99.7	0.3	0.0
11	Mathematics	48	100.0	99.7	0.3	0.0
		49	100.0	99.9	0.1	0.0
		50	100.0	99.2	0.8	0.0
		53	100.0	99.9	0.1	0.0
		54	100.0	99.5	0.5	0.0
		55	100.0	99.8	0.2	0.0
		56	99.9	97.4	2.5	0.1
		57	99.5	97.2	2.3	0.5
		60	99.9	99.4	0.5	0.1
		61	99.9	99.2	0.7	0.1
		62	100.0	99.0	0.9	0.0
		Overall	99.9	99.2	0.8	0.1

Table E-2. 2008–09 NECAP: Open Response
Interrater Agreement by Item–Reading Common Items

Grade	*Content*	*Item Number*	*% Exact + Adjacent*	*% Exact*	*% Adjacent*	*% >1*
3	Reading	7	98.3	82.6	15.7	1.7
		12	99.1	94.6	4.5	0.9
		17	99.9	86.1	13.8	0.1
		29	100.0	99.1	0.9	0.0
		34	100.0	96.8	3.2	0.0
		41	98.8	89.4	9.4	1.2
		Overall	99.4	91.2	8.2	0.6
4	Reading	7	96.8	79.3	17.5	3.2
		12	97.9	94.1	3.9	2.1
		17	97.8	78.6	19.1	2.2
		29	99.4	93.5	5.9	0.6
		34	98.9	81.7	17.1	1.1
		41	99.5	92.8	6.7	0.5
		Overall	98.4	86.4	12.0	1.6
5	Reading	7	98.9	79.5	19.4	1.1
		12	97.7	75.0	22.7	2.3
		17	98.5	79.4	19.2	1.5
		29	98.7	79.3	19.5	1.3
		34	99.1	81.1	18.0	0.9
		41	98.8	77.7	21.1	1.2
		Overall	98.6	78.7	20.0	1.4
6	Reading	7	98.0	75.3	22.7	2.0
		12	99.1	78.7	20.4	0.9
		17	99.2	82.1	17.1	0.8
		29	99.1	81.9	17.2	0.9
		34	98.9	80.6	18.3	1.1
		41	98.9	75.6	23.4	1.1
		Overall	98.9	79.1	19.8	1.1
7	Reading	7	99.1	77.3	21.9	0.9
		12	98.8	80.8	18.0	1.2
		17	98.5	78.5	19.9	1.5
		29	99.2	82.6	16.6	0.8
		34	98.9	81.0	17.9	1.1
		41	99.4	80.0	19.4	0.6
		Overall	99.0	80.1	18.9	1.0
8	Reading	7	99.2	82.9	16.3	0.8
		12	99.6	84.5	15.1	0.4
		17	98.8	80.7	18.1	1.2
		29	99.2	86.6	12.5	0.8
		34	99.5	82.7	16.9	0.5
		41	99.5	80.3	19.2	0.5
		Overall	99.3	83.0	16.3	0.7
11	Reading	11	99.8	89.4	10.4	0.2
		16	99.6	90.9	8.6	0.4
		21	99.8	90.9	8.9	0.3
		26	99.8	93.6	6.2	0.2
		31	99.8	90.8	9.0	0.2
		36	99.8	92.6	7.1	0.2
		Overall	99.7	91.5	8.2	0.3

Table E-3. 2008–09 NECAP: Open Response and Shorthand Interrater Agreement by Item–Writing Common Items

Grade	*Content Area*	*Item Number*	*Points*	*% Exact + Adjacent*	*% Exact*	*% Adjacent*	*% >1*
5	Writing	11	4	98.9	79.2	19.7	1.1
		12	4	98.2	79.8	18.4	1.8
		13	4	99.2	82.6	16.5	0.8
		14*	6	93.4	51.0	42.5	6.6
		Overall (4pt. items only)		98.8	80.6	18.1	1.2
* Item 14 is the single 6-point Writing ER item administered for Grade 5; Items 11, 12, and 13 are 4-point Writing CR items.							
8	Writing	11	4	99.7	85.5	14.2	0.3
		12	4	99.5	86.0	13.5	0.5
		13	4	99.0	83.1	15.9	1.0
		14*	6	96.5	56.8	39.7	3.5
		Overall (4pt. items only)		99.4	84.8	14.6	0.6
* Item 14 is the single 6-point Writing ER item administered for Grade 8; Items 11, 12, and 13 are 4-point Writing CR items.							
11	Writing	1*	6	96.0	58.1	38.0	4.0
		2*	6	92.7	51.8	40.9	7.3
		3*	6	93.7	52.2	41.4	6.3
		6*	6	93.6	53.9	39.7	6.4
		7*	6	95.0	52.3	42.7	5.0
		14**	6	92.8	48.8	44.0	7.2
		Overall		94.2	53.7	40.5	5.8

* Items 1, 2, 3, 6, and 7 are the five 6-point Matrix Writing ER items administered for Grade 11.
** Item 14 is the single 6-point Common Writing ER item administered for Grade 11; No 4-point Writing CR items were administered as commons for Grade 11.

Appendix F—ITEM-LEVEL CLASSICAL STATISTICS RESULTS

Table F-1. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 3 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
3	Mathematics	0	1	MC	30,449	0.3	0.82	0.34
		0	2	MC	30,449	0.4	0.75	0.51
		0	3	MC	30,449	1.1	0.61	0.52
		0	4	MC	30,449	1.3	0.85	0.39
		0	8	MC	30,449	1.3	0.53	0.46
		0	9	MC	30,449	0.9	0.78	0.54
		0	10	MC	30,449	1.4	0.50	0.37
		0	11	MC	30,449	1.3	0.83	0.33
		0	12	MC	30,449	1.1	0.50	0.47
		0	13	SA	30,449	1.1	0.66	0.43
		0	14	SA	30,449	0.8	0.44	0.46
		0	16	SA	30,449	0.8	0.65	0.44
		0	17	SA	30,449	0.9	0.32	0.45
		0	18	SA	30,449	1.5	0.36	0.51
		0	20	SA	30,449	0.7	0.72	0.52
		0	21	SA	30,449	1.7	0.38	0.44
		0	22	SA	30,449	1	0.65	0.50
		0	23	SA	30,449	1	0.82	0.59
		0	24	MC	30,449	0.2	0.90	0.29
		0	25	MC	30,449	0.5	0.87	0.27
		0	26	MC	30,449	1	0.80	0.52
		0	27	MC	30,449	0.4	0.57	0.32
		0	28	MC	30,449	0.8	0.39	0.31
		0	29	MC	30,449	0.7	0.89	0.35
		0	33	MC	30,449	1.1	0.73	0.52
		0	34	MC	30,449	1	0.75	0.42
		0	35	MC	30,449	4.2	0.59	0.43
		0	36	MC	30,449	1.9	0.68	0.45
		0	37	MC	30,449	4.3	0.61	0.42
		0	38	MC	30,449	1.1	0.40	0.32
		0	39	MC	30,449	1.1	0.86	0.31
		0	40	SA	30,449	0.7	0.79	0.50
		0	41	SA	30,449	0.4	0.91	0.33
		0	43	SA	30,449	1.3	0.40	0.51
		0	44	SA	30,449	0.8	0.66	0.53
		0	45	SA	30,449	0.8	0.52	0.57
		0	47	SA	30,449	0.7	0.74	0.57
		0	48	MC	30,449	0.2	0.87	0.21
		0	49	MC	30,449	0.6	0.87	0.36
		0	50	MC	30,449	0.8	0.81	0.48
		0	51	MC	30,449	0.5	0.48	0.40
		0	52	MC	30,449	0.7	0.70	0.38
		0	56	MC	30,449	1.1	0.78	0.40
		0	57	MC	30,449	0.8	0.67	0.50
		0	58	MC	30,449	2	0.62	0.48
		0	59	MC	30,449	1.6	0.50	0.28
		0	60	MC	30,449	1.9	0.80	0.33
		0	61	MC	30,449	0.9	0.44	0.46

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
3	Mathematics	0	62	MC	30,449	1	0.78	0.55
		0	63	MC	30,449	1.4	0.88	0.45
		0	64	SA	30,449	1.2	0.71	0.43
		0	66	SA	30,449	1	0.64	0.53
		0	67	SA	30,449	1.7	0.46	0.49
		0	68	SA	30,449	0.6	0.87	0.41
		0	70	SA	30,449	0.9	0.47	0.55
		1	5	MC	3,413	1.1	0.89	0.37
		1	7	MC	3,413	0.5	0.51	0.48
		1	15	SA	3,413	0.5	0.79	0.49
		1	30	MC	3,413	0.7	0.54	0.44
		1	32	MC	3,413	0.7	0.49	0.42
		1	46	SA	3,413	1.6	0.40	0.51
		1	53	MC	3,413	0.7	0.89	0.30
		1	55	MC	3,413	1.2	0.89	0.36
		1	65	SA	3,413	0.6	0.62	0.25
		1	69	SA	3,413	0.8	0.75	0.52
		2	5	MC	3,411	1.5	0.62	0.46
		2	7	MC	3,411	0.6	0.88	0.44
		2	15	SA	3,411	0.5	0.72	0.40
		2	30	MC	3,411	0.5	0.41	0.46
		2	32	MC	3,411	0.5	0.87	0.52
		2	46	SA	3,411	1	0.49	0.61
		2	53	MC	3,411	2.4	0.80	0.33
		2	55	MC	3,411	0.4	0.94	0.26
		2	65	SA	3,411	1.3	0.72	0.50
		2	69	SA	3,411	0.4	0.66	0.52
		3	5	MC	3,386	1.2	0.74	0.55
		3	7	MC	3,386	0.5	0.78	0.50
		3	15	SA	3,386	0.6	0.76	0.42
		3	19	SA	3,386	0.9	0.78	0.44
		3	30	MC	3,386	0.7	0.81	0.51
		3	32	MC	3,386	0.7	0.62	0.46
		3	53	MC	3,386	2.2	0.85	0.44
		3	55	MC	3,386	0.4	0.54	0.39
		3	65	SA	3,386	1.8	0.66	0.57
		3	69	SA	3,386	1.4	0.50	0.59
		4	5	MC	3,376	1.8	0.77	0.49
		4	7	MC	3,376	0.6	0.70	0.51
		4	15	SA	3,376	0.8	0.76	0.30
		4	19	SA	3,376	1.5	0.44	0.60
		4	30	MC	3,376	0.8	0.74	0.39
		4	32	MC	3,376	1	0.82	0.41
		4	53	MC	3,376	0.7	0.82	0.43
		4	55	MC	3,376	0.6	0.80	0.48
		4	65	SA	3,376	1.4	0.71	0.55
		4	69	SA	3,376	1.3	0.42	0.54
		5	5	MC	3,364	1.6	0.82	0.37
		5	7	MC	3,364	0.4	0.91	0.37
		5	15	SA	3,364	0.7	0.42	0.52

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		5	19	SA	3,364	1.4	0.71	0.39
		5	30	MC	3,364	1.1	0.77	0.38
		5	32	MC	3,364	2.6	0.60	0.40
		5	53	MC	3,364	0.5	0.84	0.43
		5	55	MC	3,364	0.5	0.89	0.30
		5	65	SA	3,364	1.5	0.44	0.51
		5	69	SA	3,364	0.9	0.43	0.50
		6	5	MC	3,392	1	0.85	0.34
		6	7	MC	3,392	1.8	0.58	0.35
		6	19	SA	3,392	1.2	0.36	0.51
		6	30	MC	3,392	2.9	0.67	0.47
		6	32	MC	3,392	1.4	0.74	0.47
		6	42	SA	3,392	0.9	0.72	0.37
		6	53	MC	3,392	1.3	0.76	0.38
		6	55	MC	3,392	1	0.73	0.36
		6	65	SA	3,392	0.9	0.87	0.39
		6	69	SA	3,392	1	0.69	0.57
		7	5	MC	3,378	1	0.89	0.41
		7	7	MC	3,378	0.9	0.51	0.48
		7	15	SA	3,378	0.3	0.79	0.51
		7	30	MC	3,378	0.7	0.55	0.44
		7	32	MC	3,378	0.8	0.49	0.38
		7	46	SA	3,378	1.5	0.40	0.50
3	Mathematics	7	53	MC	3,378	0.3	0.90	0.28
		7	55	MC	3,378	0.7	0.90	0.38
		7	65	SA	3,378	1.1	0.60	0.23
		7	69	SA	3,378	0.7	0.75	0.55
		8	5	MC	3,381	1.5	0.62	0.46
		8	7	MC	3,381	0.4	0.88	0.45
		8	15	SA	3,381	1.2	0.71	0.40
		8	30	MC	3,381	1	0.40	0.47
		8	32	MC	3,381	0.6	0.88	0.53
		8	46	SA	3,381	1.1	0.49	0.61
		8	53	MC	3,381	1.8	0.80	0.37
		8	55	MC	3,381	0.7	0.94	0.24
		8	65	SA	3,381	1.2	0.72	0.52
		8	69	SA	3,381	0.7	0.66	0.48
		9	5	MC	3,342	1.1	0.76	0.51
		9	7	MC	3,342	0.5	0.78	0.47
		9	15	SA	3,342	1.4	0.78	0.44
		9	19	SA	3,342	0.9	0.80	0.40
		9	30	MC	3,342	0.7	0.82	0.53
		9	32	MC	3,342	0.8	0.62	0.47
		9	53	MC	3,342	2.2	0.85	0.43
		9	55	MC	3,342	0.6	0.55	0.38
		9	65	SA	3,342	1.4	0.66	0.57
		9	69	SA	3,342	1	0.51	0.59

Table F-2. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 3 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
3	Reading	0	1	MC	30,385	0.3	0.90	0.43
		0	2	MC	30,385	0.4	0.88	0.53
		0	3	MC	30,385	0.2	0.90	0.49
		0	4	MC	30,385	0.4	0.88	0.52
		0	5	MC	30,385	0.4	0.85	0.43
		0	6	MC	30,385	0.4	0.51	0.30
		0	7	CR	30,385	1.1	0.38	0.48
		0	8	MC	30,385	0.4	0.66	0.51
		0	9	MC	30,385	0.7	0.64	0.36
		0	10	MC	30,385	0.6	0.72	0.50
		0	11	MC	30,385	0.7	0.91	0.43
		0	12	CR	30,385	0.5	0.61	0.56
		0	13	MC	30,385	0.5	0.59	0.33
		0	14	MC	30,385	1.1	0.68	0.37
		0	15	MC	30,385	1.0	0.58	0.26
		0	16	MC	30,385	1.1	0.83	0.43
		0	17	CR	30,385	1.7	0.38	0.58
		0	25	MC	30,385	0.5	0.84	0.52
		0	26	MC	30,385	0.8	0.68	0.50
		0	27	MC	30,385	0.7	0.68	0.45
		0	28	MC	30,385	1.5	0.77	0.44
		0	29	CR	30,385	0.7	0.81	0.54
		0	30	MC	30,385	0.6	0.60	0.39
		0	31	MC	30,385	1.1	0.76	0.55
		0	32	MC	30,385	0.9	0.63	0.45
		0	33	MC	30,385	2.1	0.69	0.40
		0	34	CR	30,385	1.7	0.60	0.65
		0	35	MC	30,385	0.4	0.89	0.43
		0	36	MC	30,385	0.7	0.80	0.37
		0	37	MC	30,385	0.4	0.76	0.42
		0	38	MC	30,385	0.8	0.81	0.40
		0	39	MC	30,385	0.5	0.78	0.40
		0	40	MC	30,385	1.0	0.69	0.46
		0	41	CR	30,385	1.0	0.71	0.54
		1	18	MC	3,407	0.4	0.87	0.51
		1	19	MC	3,407	0.7	0.86	0.48
		1	20	MC	3,407	0.3	0.53	0.27
		1	21	MC	3,407	0.7	0.82	0.39
		1	22	MC	3,407	0.6	0.52	0.42
		1	23	MC	3,407	0.9	0.69	0.44
		1	24	CR	3,407	1.0	0.38	0.43
		1	42	MC	3,407	0.5	0.74	0.29
		1	43	MC	3,407	0.9	0.67	0.48
		1	44	MC	3,407	0.7	0.77	0.37
		1	45	MC	3,407	0.8	0.90	0.54
		1	46	CR	3,407	1.0	0.78	0.64
		1	47	MC	3,407	1.0	0.80	0.49
		1	48	MC	3,407	1.1	0.61	0.35

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,407	0.9	0.58	0.43
		1	50	MC	3,407	1.6	0.72	0.43
		1	51	CR	3,407	1.1	0.66	0.57
		2	18	MC	3,405	0.4	0.63	0.40
		2	19	MC	3,405	0.7	0.86	0.39
		2	20	MC	3,405	0.4	0.84	0.39
		2	21	MC	3,405	0.5	0.76	0.54
		2	22	MC	3,405	0.7	0.73	0.46
		2	23	MC	3,405	1.2	0.81	0.41
		2	24	CR	3,405	1.1	0.47	0.57
		2	42	MC	3,405	0.5	0.73	0.32
		2	43	MC	3,405	0.8	0.79	0.59
		2	44	MC	3,405	0.8	0.83	0.50
		2	45	MC	3,405	1.8	0.78	0.46
		2	46	CR	3,405	0.6	0.89	0.57
		2	47	MC	3,405	0.5	0.52	0.32
		2	48	MC	3,405	0.9	0.68	0.41
		2	49	MC	3,405	1.4	0.74	0.50
3	Reading	2	50	MC	3,405	2.1	0.68	0.51
		2	51	CR	3,405	0.6	0.43	0.55
		3	18	MC	3,371	0.2	0.85	0.57
		3	19	MC	3,371	0.6	0.92	0.47
		3	20	MC	3,371	0.3	0.76	0.54
		3	21	MC	3,371	0.6	0.81	0.55
		3	22	MC	3,371	0.4	0.83	0.54
		3	23	MC	3,371	0.5	0.74	0.49
		3	24	CR	3,371	1.1	0.59	0.61
		3	42	MC	3,371	0.4	0.75	0.57
		3	43	MC	3,371	0.7	0.65	0.51
		3	44	MC	3,371	0.7	0.59	0.29
		3	45	MC	3,371	1.0	0.61	0.42
		3	46	CR	3,371	0.6	0.90	0.55
		3	47	MC	3,371	0.6	0.63	0.41
		3	48	MC	3,371	1.0	0.61	0.40
		3	49	MC	3,371	1.2	0.52	0.42
		3	50	MC	3,371	2.2	0.77	0.48
		3	51	CR	3,371	1.5	0.56	0.62

Table F-3. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 4 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
4	Mathematics	0	1	MC	30,499	0.2	0.84	0.53
		0	2	MC	30,499	0.4	0.74	0.45
		0	3	MC	30,499	1.9	0.68	0.44
		0	4	MC	30,499	0.9	0.60	0.36
		0	8	MC	30,499	0.5	0.22	0.26
		0	9	MC	30,499	1.2	0.50	0.42
		0	10	MC	30,499	1.9	0.63	0.49
		0	11	MC	30,499	1.6	0.53	0.46
		0	12	MC	30,499	3.1	0.52	0.45
		0	13	SA	30,499	1.1	0.82	0.34
		0	14	SA	30,499	0.7	0.67	0.46
		0	16	SA	30,499	0.8	0.62	0.45
		0	17	SA	30,499	1.3	0.42	0.50
		0	18	SA	30,499	1.1	0.81	0.50
		0	20	SA	30,499	0.8	0.58	0.54
		0	21	SA	30,499	0.9	0.68	0.52
		0	22	SA	30,499	1.4	0.34	0.45
		0	23	SA	30,499	0.8	0.81	0.37
		0	24	MC	30,499	0.2	0.89	0.45
		0	25	MC	30,499	0.3	0.84	0.26
		0	26	MC	30,499	0.3	0.84	0.50
		0	27	MC	30,499	0.6	0.84	0.44
		0	28	MC	30,499	0.4	0.48	0.48
		0	29	MC	30,499	0.9	0.78	0.45
		0	33	MC	30,499	0.7	0.89	0.39
		0	34	MC	30,499	0.9	0.90	0.43
		0	35	MC	30,499	1.0	0.81	0.31
		0	36	MC	30,499	2.5	0.24	0.28
		0	37	MC	30,499	1.0	0.77	0.45
		0	38	MC	30,499	0.9	0.79	0.41
		0	39	MC	30,499	1.3	0.74	0.40
		0	40	SA	30,499	1.3	0.61	0.46
		0	41	SA	30,499	1.3	0.69	0.51
		0	43	SA	30,499	0.6	0.84	0.34
		0	44	SA	30,499	1.3	0.51	0.54
		0	45	SA	30,499	1.2	0.60	0.63
		0	47	SA	30,499	0.7	0.57	0.62
		0	48	MC	30,499	0.4	0.86	0.38
		0	49	MC	30,499	0.6	0.83	0.44
		0	50	MC	30,499	0.6	0.79	0.42
		0	51	MC	30,499	0.7	0.69	0.54
		0	52	MC	30,499	1.2	0.87	0.48
		0	56	MC	30,499	1.1	0.72	0.51
		0	57	MC	30,499	1.0	0.73	0.39
		0	58	MC	30,499	1.0	0.54	0.40
		0	59	MC	30,499	1.3	0.81	0.40
		0	60	MC	30,499	1.4	0.42	0.25
		0	61	MC	30,499	0.8	0.82	0.50

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
4	Mathematics	0	62	MC	30,499	0.7	0.70	0.46
		0	63	MC	30,499	1.2	0.81	0.46
		0	64	SA	30,499	1.3	0.73	0.38
		0	66	SA	30,499	1.7	0.61	0.34
		0	67	SA	30,499	2.4	0.55	0.52
		0	68	SA	30,499	0.7	0.63	0.56
		0	70	SA	30,499	0.5	0.66	0.53
		1	5	MC	3,447	0.9	0.54	0.42
		1	7	MC	3,447	0.3	0.80	0.41
		1	15	SA	3,447	0.6	0.76	0.49
		1	19	SA	3,447	0.7	0.63	0.57
		1	30	MC	3,447	0.4	0.90	0.32
		1	32	MC	3,447	0.7	0.79	0.45
		1	53	MC	3,447	0.6	0.54	0.35
		1	55	MC	3,447	1.5	0.76	0.36
		1	65	SA	3,447	1.5	0.65	0.39
		1	69	SA	3,447	0.6	0.79	0.44
		2	5	MC	3,389	0.5	0.73	0.42
		2	7	MC	3,389	0.4	0.94	0.32
		2	19	SA	3,389	0.7	0.61	0.55
		2	30	MC	3,389	0.3	0.72	0.29
		2	32	MC	3,389	0.4	0.70	0.54
		2	42	SA	3,389	2.3	0.41	0.27
		2	53	MC	3,389	1.7	0.75	0.42
		2	55	MC	3,389	1.7	0.85	0.43
		2	65	SA	3,389	1.0	0.59	0.53
		2	69	SA	3,389	1.1	0.82	0.44
		3	5	MC	3,365	0.3	0.84	0.38
		3	7	MC	3,365	1.2	0.57	0.30
		3	30	MC	3,365	0.1	0.81	0.47
		3	32	MC	3,365	0.3	0.92	0.34
		3	42	SA	3,365	1.2	0.56	0.42
		3	46	SA	3,365	0.4	0.79	0.44
		3	53	MC	3,365	0.6	0.75	0.43
		3	55	MC	3,365	0.8	0.39	0.46
		3	65	SA	3,365	1.8	0.51	0.40
		3	69	SA	3,365	0.7	0.67	0.56
		4	5	MC	3,355	1.4	0.41	0.44
		4	7	MC	3,355	0.4	0.83	0.44
		4	15	SA	3,355	0.8	0.78	0.41
		4	30	MC	3,355	0.7	0.84	0.47
		4	32	MC	3,355	1.3	0.34	0.48
		4	46	SA	3,355	0.6	0.52	0.60
		4	53	MC	3,355	0.4	0.81	0.45
		4	55	MC	3,355	0.3	0.87	0.40
		4	65	SA	3,355	1.3	0.86	0.35
		4	69	SA	3,355	1.1	0.60	0.55
		5	5	MC	3,359	0.2	0.85	0.51
		5	7	MC	3,359	1.5	0.52	0.34
		5	19	SA	3,359	1.0	0.50	0.49

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		5	30	MC	3,359	2.4	0.87	0.32
		5	32	MC	3,359	0.7	0.87	0.41
		5	42	SA	3,359	0.8	0.92	0.33
		5	53	MC	3,359	1.4	0.37	0.31
		5	55	MC	3,359	1.4	0.53	0.41
		5	65	SA	3,359	1.3	0.86	0.29
		5	69	SA	3,359	0.8	0.55	0.55
		6	5	MC	3,405	0.5	0.71	0.45
		6	7	MC	3,405	0.3	0.90	0.33
		6	15	SA	3,405	2.0	0.28	0.40
		6	19	SA	3,405	1.0	0.66	0.50
		6	30	MC	3,405	0.8	0.93	0.33
		6	32	MC	3,405	0.4	0.90	0.29
		6	53	MC	3,405	0.9	0.35	0.39
		6	55	MC	3,405	1.3	0.67	0.48
		6	65	SA	3,405	0.6	0.73	0.32
		6	69	SA	3,405	0.9	0.54	0.56
		7	5	MC	3,398	0.5	0.53	0.46
		7	7	MC	3,398	0.3	0.78	0.43
		7	15	SA	3,398	0.6	0.75	0.50
		7	19	SA	3,398	1.2	0.62	0.59
		7	30	MC	3,398	0.6	0.90	0.31
		7	32	MC	3,398	0.8	0.78	0.46
4	Mathematics	7	53	MC	3,398	0.4	0.53	0.35
		7	55	MC	3,398	1.1	0.75	0.33
		7	65	SA	3,398	1.3	0.65	0.41
		7	69	SA	3,398	0.5	0.79	0.49
		8	5	MC	3,412	0.3	0.72	0.41
		8	7	MC	3,412	0.4	0.93	0.34
		8	19	SA	3,412	1.0	0.60	0.57
		8	30	MC	3,412	0.6	0.70	0.28
		8	32	MC	3,412	0.7	0.69	0.48
		8	42	SA	3,412	1.3	0.41	0.30
		8	53	MC	3,412	0.8	0.76	0.43
		8	55	MC	3,412	0.8	0.84	0.40
		8	65	SA	3,412	1.4	0.58	0.52
		8	69	SA	3,412	0.8	0.82	0.45
		9	5	MC	3,366	0.4	0.84	0.38
		9	7	MC	3,366	1.5	0.55	0.30
		9	30	MC	3,366	1.2	0.82	0.42
		9	32	MC	3,366	1.2	0.91	0.33
		9	42	SA	3,366	1.1	0.56	0.41
		9	46	SA	3,366	0.7	0.79	0.46
		9	53	MC	3,366	0.7	0.76	0.42
		9	55	MC	3,366	0.8	0.38	0.46
		9	65	SA	3,366	1.4	0.51	0.41
		9	69	SA	3,366	0.8	0.68	0.59

Table F-4. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 4 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
4	Reading	0	1	MC	30,435	0.2	0.85	0.39
		0	2	MC	30,435	0.2	0.69	0.36
		0	3	MC	30,435	0.2	0.86	0.49
		0	4	MC	30,435	0.3	0.74	0.34
		0	5	MC	30,435	0.5	0.65	0.33
		0	6	MC	30,435	0.7	0.81	0.40
		0	7	CR	30,435	0.8	0.71	0.54
		0	8	MC	30,435	0.3	0.85	0.51
		0	9	MC	30,435	0.4	0.87	0.46
		0	10	MC	30,435	0.5	0.83	0.50
		0	11	MC	30,435	0.5	0.78	0.54
		0	12	CR	30,435	0.9	0.56	0.47
		0	13	MC	30,435	0.5	0.76	0.44
		0	14	MC	30,435	1.0	0.82	0.49
		0	15	MC	30,435	0.9	0.75	0.44
		0	16	MC	30,435	1.0	0.89	0.40
		0	17	CR	30,435	1.1	0.46	0.55
		0	25	MC	30,435	0.4	0.66	0.44
		0	26	MC	30,435	0.7	0.58	0.42
		0	27	MC	30,435	0.9	0.59	0.39
		0	28	MC	30,435	1.0	0.51	0.30
		0	29	CR	30,435	1.0	0.55	0.38
		0	30	MC	30,435	0.8	0.47	0.26
		0	31	MC	30,435	1.0	0.63	0.42
		0	32	MC	30,435	1.0	0.86	0.46
		0	33	MC	30,435	1.4	0.75	0.43
		0	34	CR	30,435	1.3	0.37	0.44
		0	35	MC	30,435	0.5	0.79	0.32
		0	36	MC	30,435	0.7	0.60	0.41
		0	37	MC	30,435	0.3	0.81	0.33
		0	38	MC	30,435	0.5	0.68	0.33
		0	39	MC	30,435	0.5	0.78	0.34
		0	40	MC	30,435	0.6	0.90	0.38
		0	41	CR	30,435	0.7	0.62	0.35
		1	18	MC	3,442	0.2	0.89	0.41
		1	19	MC	3,442	0.5	0.68	0.34
		1	20	MC	3,442	0.3	0.90	0.46
		1	21	MC	3,442	0.6	0.74	0.48
		1	22	MC	3,442	0.5	0.69	0.39
		1	23	MC	3,442	1.1	0.67	0.29
		1	24	CR	3,442	0.6	0.45	0.56
		1	42	MC	3,442	0.4	0.82	0.44
		1	43	MC	3,442	0.6	0.81	0.39
		1	44	MC	3,442	0.8	0.67	0.36
		1	45	MC	3,442	1.3	0.63	0.51
		1	46	CR	3,442	0.6	0.60	0.45
		1	47	MC	3,442	0.8	0.67	0.51
		1	48	MC	3,442	0.9	0.47	0.38

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,442	0.6	0.73	0.49
		1	50	MC	3,442	1.4	0.69	0.43
		1	51	CR	3,442	0.6	0.39	0.58
		2	18	MC	3,383	0.5	0.66	0.37
		2	19	MC	3,383	0.4	0.81	0.43
		2	20	MC	3,383	0.3	0.63	0.37
		2	21	MC	3,383	0.6	0.72	0.44
		2	22	MC	3,383	0.4	0.75	0.43
		2	23	MC	3,383	1.3	0.68	0.44
		2	24	CR	3,383	0.9	0.37	0.56
		2	42	MC	3,383	0.5	0.81	0.42
		2	43	MC	3,383	0.8	0.79	0.47
		2	44	MC	3,383	0.7	0.64	0.28
		2	45	MC	3,383	0.9	0.58	0.27
		2	46	CR	3,383	0.8	0.69	0.46
		2	47	MC	3,383	0.6	0.74	0.44
		2	48	MC	3,383	1.0	0.70	0.43
		2	49	MC	3,383	0.9	0.81	0.47
4	Reading	2	50	MC	3,383	2.0	0.66	0.42
		2	51	CR	3,383	0.9	0.43	0.56
		3	18	MC	3,361	1.2	0.68	0.31
		3	19	MC	3,361	0.7	0.81	0.30
		3	20	MC	3,361	0.2	0.73	0.43
		3	21	MC	3,361	0.4	0.83	0.40
		3	22	MC	3,361	0.4	0.71	0.47
		3	23	MC	3,361	0.7	0.92	0.37
		3	24	CR	3,361	1.1	0.57	0.56
		3	42	MC	3,361	0.7	0.49	0.38
		3	43	MC	3,361	0.9	0.51	0.27
		3	44	MC	3,361	1.2	0.88	0.49
		3	45	MC	3,361	1.3	0.71	0.42
		3	46	CR	3,361	1.0	0.77	0.59
		3	47	MC	3,361	0.5	0.73	0.51
		3	48	MC	3,361	1.1	0.50	0.38
		3	49	MC	3,361	0.7	0.76	0.47
		3	50	MC	3,361	2.1	0.71	0.51
		3	51	CR	3,361	1.1	0.46	0.58

Table F-5. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 5 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
5	Mathematics	0	1	MC	32,200	0.1	0.87	0.33
		0	2	MC	32,200	0.1	0.86	0.27
		0	3	MC	32,200	0.7	0.59	0.25
		0	4	MC	32,200	0.3	0.37	0.32
		0	5	MC	32,200	0.3	0.73	0.43
		0	6	MC	32,200	0.4	0.61	0.31
		0	10	MC	32,200	0.4	0.36	0.38
		0	11	MC	32,200	0.4	0.65	0.36
		0	12	MC	32,200	0.7	0.68	0.43
		0	13	MC	32,200	1.2	0.54	0.40
		0	14	SA	32,200	0.7	0.56	0.48
		0	15	SA	32,200	0.8	0.49	0.47
		0	17	SA	32,200	1.0	0.60	0.49
		0	18	CR	32,200	0.8	0.50	0.60
		0	21	SA	32,200	2.3	0.64	0.52
		0	22	MC	32,200	0.3	0.51	0.36
		0	23	MC	32,200	0.3	0.60	0.16
		0	24	MC	32,200	0.3	0.80	0.40
		0	25	MC	32,200	0.3	0.51	0.40
		0	29	MC	32,200	0.4	0.64	0.36
		0	30	MC	32,200	0.3	0.75	0.40
		0	31	MC	32,200	0.3	0.62	0.43
		0	32	MC	32,200	0.3	0.89	0.37
		0	33	MC	32,200	0.6	0.48	0.36
		0	34	MC	32,200	0.6	0.56	0.43
		0	35	MC	32,200	1.9	0.72	0.32
		0	37	SA	32,200	0.7	0.41	0.48
		0	40	CR	32,200	0.7	0.55	0.63
		0	41	SA	32,200	1.2	0.59	0.55
		0	42	CR	32,200	1.4	0.40	0.66
		0	43	SA	32,200	1.4	0.71	0.55
		0	44	MC	32,200	0.2	0.73	0.49
		0	45	MC	32,200	0.2	0.83	0.33
		0	46	MC	32,200	0.5	0.72	0.29
		0	47	MC	32,200	0.3	0.52	0.55
		0	48	MC	32,200	0.3	0.44	0.31
		0	52	MC	32,200	0.3	0.59	0.43
		0	53	MC	32,200	0.6	0.46	0.41
		0	54	MC	32,200	0.7	0.47	0.20
		0	55	MC	32,200	0.6	0.64	0.28
		0	56	MC	32,200	0.6	0.69	0.44
		0	57	MC	32,200	1.3	0.39	0.49
		0	58	SA	32,200	0.7	0.44	0.45
		0	59	SA	32,200	0.5	0.57	0.50
		0	60	SA	32,200	0.7	0.45	0.53
		0	62	CR	32,200	1.1	0.51	0.63
		0	63	SA	32,200	1.1	0.62	0.55
		0	65	SA	32,200	2.3	0.43	0.36

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	7	MC	3,651	0.2	0.74	0.42
		1	9	MC	3,651	0.4	0.46	0.50
		1	16	SA	3,651	1.6	0.23	0.48
		1	20	CR	3,651	1.9	0.31	0.67
		1	26	MC	3,651	0.3	0.84	0.35
		1	28	MC	3,651	0.3	0.91	0.30
		1	36	SA	3,651	0.5	0.39	0.45
		1	39	SA	3,651	0.9	0.51	0.60
		1	49	MC	3,651	0.4	0.50	0.41
		1	51	MC	3,651	0.2	0.78	0.23
		1	61	SA	3,651	1.0	0.47	0.53
		2	7	MC	3,608	0.4	0.42	0.29
		2	9	MC	3,608	0.5	0.42	0.46
		2	16	SA	3,608	0.5	0.40	0.53
		2	20	CR	3,608	1.5	0.47	0.67
		2	26	MC	3,608	0.6	0.49	0.17
		2	28	MC	3,608	0.2	0.65	0.46
		2	36	SA	3,608	0.3	0.67	0.40
		2	39	SA	3,608	0.9	0.38	0.59
		2	49	MC	3,608	0.2	0.80	0.36
		2	51	MC	3,608	0.2	0.88	0.31
		2	61	SA	3,608	0.8	0.46	0.57
		3	7	MC	3,550	0.3	0.65	0.36
		3	9	MC	3,550	0.4	0.61	0.51
5	Mathematics	3	19	SA	3,550	2.2	0.32	0.46
		3	26	MC	3,550	0.2	0.59	0.21
		3	28	MC	3,550	0.3	0.79	0.37
		3	36	SA	3,550	0.5	0.31	0.50
		3	39	SA	3,550	0.6	0.33	0.54
		3	49	MC	3,550	0.3	0.77	0.39
		3	51	MC	3,550	0.2	0.79	0.44
		3	61	SA	3,550	0.5	0.54	0.61
		4	7	MC	3,568	0.2	0.54	0.40
		4	9	MC	3,568	0.4	0.83	0.32
		4	16	SA	3,568	0.7	0.73	0.40
		4	26	MC	3,568	0.3	0.35	0.28
		4	28	MC	3,568	0.3	0.29	0.36
		4	36	SA	3,568	0.5	0.77	0.28
		4	38	SA	3,568	0.7	0.44	0.38
		4	49	MC	3,568	0.3	0.70	0.45
		4	51	MC	3,568	0.2	0.55	0.50
		4	61	SA	3,568	0.6	0.74	0.49
		4	64	CR	3,568	1.4	0.37	0.61
		5	7	MC	3,568	0.2	0.84	0.44
		5	9	MC	3,568	0.6	0.53	0.39
		5	19	SA	3,568	1.0	0.52	0.55
		5	26	MC	3,568	0.3	0.65	0.42
		5	28	MC	3,568	0.3	0.83	0.36
		5	36	SA	3,568	2.4	0.18	0.39
		5	38	SA	3,568	0.5	0.75	0.39

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		5	49	MC	3,568	0.4	0.35	0.27
		5	51	MC	3,568	0.4	0.42	0.41
		5	61	SA	3,568	2.1	0.24	0.34
		5	64	CR	3,568	1.5	0.42	0.63
		6	7	MC	3,535	0.3	0.81	0.30
		6	9	MC	3,535	0.3	0.60	0.46
		6	19	SA	3,535	1.6	0.40	0.56
		6	26	MC	3,535	0.2	0.57	0.36
		6	28	MC	3,535	0.3	0.50	0.49
		6	36	SA	3,535	0.4	0.62	0.33
		6	38	SA	3,535	0.5	0.43	0.52
		6	49	MC	3,535	0.3	0.24	0.26
		6	51	MC	3,535	0.3	0.51	0.38
		6	61	SA	3,535	0.7	0.61	0.61
		6	64	CR	3,535	1.1	0.59	0.62
		7	7	MC	3,610	0.4	0.75	0.39
		7	9	MC	3,610	0.4	0.48	0.51
		7	16	SA	3,610	1.5	0.22	0.46
		7	20	CR	3,610	2.4	0.30	0.65
		7	26	MC	3,610	0.3	0.85	0.33
		7	28	MC	3,610	0.2	0.92	0.30
		7	36	SA	3,610	0.6	0.41	0.46
		7	39	SA	3,610	1.0	0.49	0.57
5	Mathematics	7	49	MC	3,610	0.4	0.51	0.40
		7	51	MC	3,610	0.4	0.78	0.23
		7	61	SA	3,610	0.9	0.48	0.50
		8	7	MC	3,570	0.3	0.41	0.29
		8	9	MC	3,570	0.7	0.43	0.46
		8	16	SA	3,570	0.3	0.41	0.52
		8	20	CR	3,570	1.4	0.47	0.64
		8	26	MC	3,570	0.3	0.49	0.16
		8	28	MC	3,570	0.2	0.64	0.47
		8	36	SA	3,570	0.2	0.67	0.39
		8	39	SA	3,570	0.8	0.39	0.56
		8	49	MC	3,570	0.3	0.80	0.35
		8	51	MC	3,570	0.2	0.87	0.28
		8	61	SA	3,570	0.6	0.48	0.54
		9	7	MC	3,533	0.3	0.63	0.38
		9	9	MC	3,533	0.5	0.63	0.49
		9	19	SA	3,533	1.5	0.32	0.44
		9	26	MC	3,533	0.3	0.56	0.25
		9	28	MC	3,533	0.3	0.80	0.40
		9	36	SA	3,533	0.6	0.32	0.52
		9	39	SA	3,533	0.8	0.32	0.55
		9	49	MC	3,533	0.4	0.78	0.36
		9	51	MC	3,533	0.3	0.79	0.40
		9	61	SA	3,533	0.7	0.56	0.57

Table F-6. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 5 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
5	Reading	0	1	MC	32,130	0.3	0.72	0.37
		0	2	MC	32,130	0.3	0.70	0.40
		0	3	MC	32,130	0.2	0.91	0.44
		0	4	MC	32,130	0.2	0.74	0.47
		0	5	MC	32,130	0.2	0.54	0.17
		0	6	MC	32,130	0.4	0.69	0.38
		0	7	CR	32,130	0.7	0.38	0.50
		0	8	MC	32,130	0.3	0.49	0.28
		0	9	MC	32,130	0.3	0.91	0.37
		0	10	MC	32,130	0.3	0.76	0.40
		0	11	MC	32,130	0.5	0.75	0.33
		0	12	CR	32,130	0.9	0.35	0.46
		0	13	MC	32,130	0.5	0.71	0.42
		0	14	MC	32,130	0.6	0.82	0.46
		0	15	MC	32,130	0.6	0.53	0.37
		0	16	MC	32,130	0.7	0.52	0.23
		0	17	CR	32,130	1.1	0.33	0.52
		0	25	MC	32,130	0.2	0.77	0.43
		0	26	MC	32,130	0.3	0.60	0.41
		0	27	MC	32,130	0.3	0.80	0.44
		0	28	MC	32,130	0.6	0.66	0.41
		0	29	CR	32,130	0.8	0.41	0.53
		0	30	MC	32,130	0.6	0.86	0.49
		0	31	MC	32,130	0.8	0.80	0.44
		0	32	MC	32,130	0.7	0.64	0.36
		0	33	MC	32,130	0.8	0.85	0.48
		0	34	CR	32,130	1.1	0.42	0.56
		0	35	MC	32,130	0.3	0.74	0.43
		0	36	MC	32,130	0.3	0.79	0.39
		0	37	MC	32,130	0.3	0.66	0.31
		0	38	MC	32,130	0.3	0.76	0.49
		0	39	MC	32,130	0.3	0.78	0.22
		0	40	MC	32,130	0.6	0.49	0.32
		0	41	CR	32,130	0.6	0.48	0.61
		1	18	MC	3,647	0.3	0.76	0.31
		1	19	MC	3,647	0.4	0.52	0.38
		1	20	MC	3,647	0.4	0.80	0.37
		1	21	MC	3,647	0.4	0.77	0.47
		1	22	MC	3,647	0.4	0.64	0.48
		1	23	MC	3,647	0.6	0.84	0.44
		1	24	CR	3,647	0.7	0.45	0.61
		1	42	MC	3,647	0.4	0.62	0.33
		1	43	MC	3,647	0.4	0.70	0.35
		1	44	MC	3,647	0.5	0.74	0.43
		1	45	MC	3,647	0.6	0.76	0.53
		1	46	CR	3,647	0.9	0.43	0.64
		1	47	MC	3,647	0.4	0.52	0.44
		1	48	MC	3,647	0.5	0.74	0.46

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,647	0.5	0.62	0.45
		1	50	MC	3,647	0.7	0.58	0.43
		1	51	CR	3,647	1.1	0.47	0.58
		2	18	MC	3,602	0.2	0.79	0.40
		2	19	MC	3,602	0.3	0.83	0.48
		2	20	MC	3,602	0.3	0.92	0.32
		2	21	MC	3,602	0.3	0.47	0.32
		2	22	MC	3,602	0.4	0.48	0.28
		2	23	MC	3,602	0.7	0.63	0.37
		2	24	CR	3,602	0.8	0.46	0.61
		2	42	MC	3,602	0.3	0.69	0.47
		2	43	MC	3,602	0.3	0.87	0.38
		2	44	MC	3,602	0.4	0.59	0.33
		2	45	MC	3,602	0.5	0.89	0.47
		2	46	CR	3,602	0.9	0.43	0.60
		2	47	MC	3,602	0.4	0.72	0.45
		2	48	MC	3,602	0.5	0.62	0.38
		2	49	MC	3,602	0.4	0.78	0.53
5	Reading	2	50	MC	3,602	0.6	0.65	0.39
		2	51	CR	3,602	0.9	0.43	0.58
		3	18	MC	3,544	0.1	0.82	0.49
		3	19	MC	3,544	0.2	0.64	0.40
		3	20	MC	3,544	0.1	0.78	0.48
		3	21	MC	3,544	0.2	0.90	0.43
		3	22	MC	3,544	0.3	0.83	0.43
		3	23	MC	3,544	0.5	0.70	0.40
		3	24	CR	3,544	0.6	0.43	0.50
		3	42	MC	3,544	0.3	0.84	0.42
		3	43	MC	3,544	0.3	0.64	0.41
		3	44	MC	3,544	0.4	0.61	0.16
		3	45	MC	3,544	0.5	0.83	0.34
		3	46	CR	3,544	0.9	0.37	0.66
		3	47	MC	3,544	0.3	0.75	0.45
		3	48	MC	3,544	0.5	0.83	0.50
		3	49	MC	3,544	0.5	0.58	0.26
		3	50	MC	3,544	0.7	0.69	0.33
		3	51	CR	3,544	0.6	0.45	0.58

Table F-7. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 5 Writing

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
		1	1	MC	32,065	0.2	0.87	0.34
		1	2	MC	32,065	0.2	0.77	0.41
		1	3	MC	32,065	0.2	0.83	0.34
		1	4	MC	32,065	0.2	0.86	0.37
		1	5	MC	32,065	0.2	0.87	0.37
		1	6	MC	32,065	0.2	0.79	0.25
		1	7	MC	32,065	0.3	0.80	0.31
		1	8	MC	32,065	0.3	0.79	0.42
5	Writing	1	9	MC	32,065	0.3	0.76	0.36
		1	10	MC	32,065	0.6	0.45	0.29
		1	11	CR	32,065	0.4	0.49	0.56
		1	12	CR	32,065	0.6	0.49	0.59
		1	13	CR	32,065	1.2	0.53	0.54
		1	14	SA	32,065	0.2	0.99	0.19
		1	15	SA	32,065	0.2	0.99	0.20
		1	16	SA	32,065	0.2	0.99	0.20
		1	17	WR	32,065	0.2	0.53	0.62

Table F-8. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 6 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
6	Mathematics	0	1	MC	32,434	0.5	0.54	0.26
		0	2	MC	32,434	0.3	0.82	0.23
		0	3	MC	32,434	0.6	0.48	0.39
		0	4	MC	32,434	0.5	0.58	0.47
		0	5	MC	32,434	0.5	0.80	0.43
		0	6	MC	32,434	0.3	0.49	0.50
		0	10	MC	32,434	0.5	0.52	0.45
		0	11	MC	32,434	0.7	0.47	0.40
		0	12	MC	32,434	1.0	0.41	0.44
		0	13	MC	32,434	0.9	0.83	0.32
		0	14	SA	32,434	1.0	0.58	0.46
		0	15	SA	32,434	0.4	0.72	0.34
		0	17	SA	32,434	1.0	0.73	0.50
		0	18	CR	32,434	1.2	0.33	0.71
		0	21	SA	32,434	2.3	0.57	0.56
		0	22	MC	32,434	0.3	0.62	0.39
		0	23	MC	32,434	0.3	0.67	0.20
		0	24	MC	32,434	0.5	0.56	0.22
		0	25	MC	32,434	0.3	0.58	0.48
		0	29	MC	32,434	0.3	0.55	0.54
		0	30	MC	32,434	0.4	0.30	0.47
		0	31	MC	32,434	0.4	0.50	0.57
		0	32	MC	32,434	0.4	0.35	0.42
		0	33	MC	32,434	0.6	0.65	0.50
		0	34	MC	32,434	1.1	0.59	0.34
		0	35	MC	32,434	1.5	0.86	0.37
		0	37	SA	32,434	0.5	0.73	0.46
		0	40	CR	32,434	1.1	0.34	0.61
		0	41	SA	32,434	1.8	0.29	0.55
		0	42	CR	32,434	1.5	0.21	0.55
		0	43	SA	32,434	1.6	0.51	0.52
		0	44	MC	32,434	0.2	0.76	0.54
		0	45	MC	32,434	0.3	0.65	0.32
		0	46	MC	32,434	0.4	0.65	0.47
		0	47	MC	32,434	0.3	0.31	0.26
		0	48	MC	32,434	0.3	0.59	0.45
		0	52	MC	32,434	0.7	0.62	0.45
		0	53	MC	32,434	0.4	0.63	0.28
		0	54	MC	32,434	0.4	0.48	0.49
		0	55	MC	32,434	0.7	0.58	0.49
		0	56	MC	32,434	0.6	0.50	0.31
		0	57	MC	32,434	1.4	0.41	0.25
		0	58	SA	32,434	0.8	0.36	0.34
		0	59	SA	32,434	0.8	0.70	0.59
		0	60	SA	32,434	1.2	0.48	0.56
		0	62	CR	32,434	1.4	0.35	0.73
		0	63	SA	32,434	2.5	0.38	0.54
		0	65	SA	32,434	2.5	0.51	0.50

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	7	MC	3,631	0.7	0.34	0.23
		1	9	MC	3,631	0.4	0.56	0.55
		1	20	CR	3,631	1.7	0.37	0.74
		1	26	MC	3,631	0.4	0.56	0.48
		1	28	MC	3,631	0.2	0.73	0.49
		1	36	SA	3,631	0.5	0.32	0.47
		1	38	SA	3,631	3.3	0.25	0.56
		1	49	MC	3,631	0.9	0.31	0.32
		1	51	MC	3,631	0.3	0.66	0.43
		1	61	SA	3,631	1.7	0.29	0.58
		2	7	MC	3,570	0.4	0.51	0.36
		2	9	MC	3,570	0.6	0.48	0.43
		2	19	SA	3,570	1.2	0.40	0.57
		2	20	CR	3,570	1.7	0.31	0.58
		2	26	MC	3,570	0.6	0.54	0.38
		2	28	MC	3,570	0.3	0.66	0.37
		2	36	SA	3,570	0.7	0.64	0.54
		2	39	SA	3,570	1.9	0.33	0.55
		2	49	MC	3,570	0.6	0.51	0.50
		2	51	MC	3,570	0.6	0.39	0.33
		2	61	SA	3,570	1.4	0.30	0.56
		3	7	MC	3,609	0.8	0.42	0.47
		3	9	MC	3,609	0.7	0.54	0.47
		3	20	CR	3,609	2.5	0.33	0.57
6	Mathematics	3	26	MC	3,609	0.6	0.51	0.40
		3	28	MC	3,609	0.4	0.58	0.49
		3	36	SA	3,609	0.9	0.70	0.49
		3	38	SA	3,609	1.1	0.42	0.34
		3	49	MC	3,609	0.7	0.64	0.38
		3	51	MC	3,609	0.4	0.46	0.40
		3	61	SA	3,609	1.1	0.33	0.42
		4	7	MC	3,597	0.7	0.62	0.47
		4	9	MC	3,597	0.6	0.72	0.36
		4	16	SA	3,597	0.6	0.58	0.46
		4	19	SA	3,597	0.9	0.66	0.38
		4	26	MC	3,597	0.4	0.32	0.49
		4	28	MC	3,597	0.3	0.75	0.22
		4	36	SA	3,597	1.9	0.42	0.46
		4	39	SA	3,597	0.9	0.50	0.58
		4	49	MC	3,597	0.2	0.74	0.29
		4	51	MC	3,597	0.4	0.81	0.42
		4	61	SA	3,597	1.1	0.17	0.53
		4	64	CR	3,597	1.8	0.26	0.58
		5	7	MC	3,608	0.6	0.80	0.42
		5	9	MC	3,608	0.7	0.65	0.58
		5	16	SA	3,608	0.4	0.83	0.29
		5	26	MC	3,608	0.3	0.32	0.32
		5	28	MC	3,608	0.2	0.43	0.21
		5	36	SA	3,608	0.6	0.29	0.31
		5	38	SA	3,608	0.7	0.66	0.41

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		5	49	MC	3,608	0.3	0.74	0.35
		5	51	MC	3,608	0.1	0.66	0.47
		5	61	SA	3,608	1.2	0.42	0.64
		5	64	CR	3,608	1.5	0.40	0.59
		6	7	MC	3,622	0.5	0.37	0.56
		6	9	MC	3,622	0.7	0.66	0.54
		6	16	SA	3,622	0.2	0.49	0.46
		6	19	SA	3,622	0.9	0.20	0.58
		6	26	MC	3,622	0.3	0.62	0.28
		6	28	MC	3,622	0.4	0.34	0.36
		6	36	SA	3,622	0.6	0.66	0.43
		6	39	SA	3,622	1.5	0.23	0.56
		6	49	MC	3,622	0.4	0.42	0.55
		6	51	MC	3,622	0.4	0.48	0.44
		6	61	SA	3,622	1.7	0.56	0.60
		6	64	CR	3,622	1.8	0.43	0.74
		7	7	MC	3,633	1.2	0.34	0.23
		7	9	MC	3,633	1.0	0.57	0.53
		7	20	CR	3,633	2.4	0.39	0.73
		7	26	MC	3,633	0.5	0.58	0.46
		7	28	MC	3,633	0.3	0.74	0.48
		7	36	SA	3,633	0.7	0.33	0.47
		7	38	SA	3,633	2.8	0.27	0.56
6	Mathematics	7	49	MC	3,633	1.0	0.32	0.33
		7	51	MC	3,633	0.6	0.67	0.46
		7	61	SA	3,633	1.8	0.30	0.58
		8	7	MC	3,580	0.5	0.52	0.35
		8	9	MC	3,580	0.6	0.47	0.41
		8	19	SA	3,580	1.1	0.40	0.57
		8	20	CR	3,580	1.8	0.31	0.60
		8	26	MC	3,580	0.2	0.55	0.38
		8	28	MC	3,580	0.3	0.67	0.34
		8	36	SA	3,580	0.7	0.65	0.54
		8	39	SA	3,580	1.6	0.32	0.52
		8	49	MC	3,580	0.6	0.53	0.50
		8	51	MC	3,580	0.4	0.39	0.33
		8	61	SA	3,580	1.8	0.31	0.57
		9	7	MC	3,576	0.7	0.42	0.49
		9	9	MC	3,576	0.8	0.55	0.47
		9	20	CR	3,576	1.6	0.35	0.57
		9	26	MC	3,576	0.4	0.53	0.39
		9	28	MC	3,576	0.4	0.60	0.51
		9	36	SA	3,576	0.5	0.71	0.48
		9	38	SA	3,576	1.1	0.42	0.36
		9	49	MC	3,576	0.5	0.66	0.40
		9	51	MC	3,576	0.4	0.47	0.40
		9	61	SA	3,576	1.0	0.32	0.44

Table F-9. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 6 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
6	Reading	0	1	MC	32,379	0.3	0.79	0.29
		0	2	MC	32,379	0.3	0.72	0.21
		0	3	MC	32,379	0.3	0.70	0.44
		0	4	MC	32,379	0.3	0.73	0.42
		0	5	MC	32,379	0.3	0.81	0.32
		0	6	MC	32,379	0.5	0.76	0.40
		0	7	CR	32,379	0.9	0.44	0.59
		0	8	MC	32,379	0.4	0.75	0.43
		0	9	MC	32,379	0.4	0.90	0.36
		0	10	MC	32,379	0.5	0.82	0.36
		0	11	MC	32,379	0.5	0.87	0.45
		0	12	CR	32,379	0.8	0.49	0.61
		0	13	MC	32,379	0.6	0.71	0.37
		0	14	MC	32,379	0.7	0.92	0.49
		0	15	MC	32,379	0.6	0.82	0.41
		0	16	MC	32,379	0.7	0.80	0.44
		0	17	CR	32,379	1.1	0.41	0.61
		0	25	MC	32,379	0.4	0.80	0.52
		0	26	MC	32,379	0.6	0.74	0.45
		0	27	MC	32,379	0.6	0.67	0.35
		0	28	MC	32,379	0.8	0.84	0.49
		0	29	CR	32,379	1.1	0.49	0.67
		0	30	MC	32,379	0.8	0.75	0.48
		0	31	MC	32,379	1.0	0.50	0.34
		0	32	MC	32,379	0.9	0.86	0.54
		0	33	MC	32,379	1.0	0.76	0.47
		0	34	CR	32,379	1.4	0.44	0.63
		0	35	MC	32,379	0.3	0.82	0.43
		0	36	MC	32,379	0.4	0.89	0.39
		0	37	MC	32,379	0.3	0.56	0.36
		0	38	MC	32,379	0.5	0.74	0.34
		0	39	MC	32,379	0.5	0.86	0.44
		0	40	MC	32,379	0.6	0.86	0.46
		0	41	CR	32,379	0.7	0.49	0.52
		1	18	MC	3,623	0.4	0.66	0.26
		1	19	MC	3,623	0.5	0.90	0.40
		1	20	MC	3,623	0.4	0.68	0.38
		1	21	MC	3,623	0.4	0.89	0.33
		1	22	MC	3,623	0.4	0.95	0.32
		1	23	MC	3,623	0.5	0.68	0.42
		1	24	CR	3,623	0.6	0.46	0.63
		1	42	MC	3,623	0.4	0.74	0.47
		1	43	MC	3,623	0.4	0.82	0.42
		1	44	MC	3,623	0.3	0.71	0.39
		1	45	MC	3,623	0.6	0.82	0.44
		1	46	CR	3,623	0.9	0.48	0.59
		1	47	MC	3,623	0.4	0.91	0.49
		1	48	MC	3,623	0.6	0.69	0.29

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,623	0.6	0.77	0.43
		1	50	MC	3,623	0.6	0.86	0.53
		1	51	CR	3,623	1.7	0.45	0.60
		2	18	MC	3,567	0.2	0.64	0.21
		2	19	MC	3,567	0.2	0.60	0.39
		2	20	MC	3,567	0.2	0.83	0.41
		2	21	MC	3,567	0.4	0.51	0.31
		2	22	MC	3,567	0.4	0.65	0.36
		2	23	MC	3,567	0.5	0.76	0.42
		2	24	CR	3,567	0.5	0.51	0.63
		2	42	MC	3,567	0.5	0.86	0.52
		2	43	MC	3,567	0.6	0.87	0.51
		2	44	MC	3,567	0.5	0.77	0.48
		2	45	MC	3,567	0.8	0.86	0.49
		2	46	CR	3,567	1.0	0.48	0.64
		2	47	MC	3,567	0.7	0.63	0.38
		2	48	MC	3,567	0.8	0.73	0.45
		2	49	MC	3,567	0.7	0.90	0.49
6	Reading	2	50	MC	3,567	0.8	0.93	0.43
		2	51	CR	3,567	1.0	0.51	0.63
		3	18	MC	3,598	0.3	0.91	0.42
		3	19	MC	3,598	0.3	0.90	0.37
		3	20	MC	3,598	0.3	0.87	0.40
		3	21	MC	3,598	0.3	0.88	0.41
		3	22	MC	3,598	0.4	0.57	0.16
		3	23	MC	3,598	0.7	0.45	0.17
		3	24	CR	3,598	0.8	0.49	0.58
		3	42	MC	3,598	0.7	0.50	0.24
		3	43	MC	3,598	0.6	0.53	0.27
		3	44	MC	3,598	0.6	0.76	0.39
		3	45	MC	3,598	0.9	0.53	0.32
		3	46	CR	3,598	1.3	0.53	0.63
		3	47	MC	3,598	0.7	0.73	0.50
		3	48	MC	3,598	0.8	0.74	0.47
		3	49	MC	3,598	0.8	0.77	0.40
		3	50	MC	3,598	1.0	0.53	0.10
		3	51	CR	3,598	1.6	0.46	0.65

Table F-10. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 7 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
7	Mathematics	0	1	MC	33,011	0.4	0.68	0.40
		0	2	MC	33,011	0.4	0.43	0.50
		0	3	MC	33,011	0.4	0.62	0.35
		0	4	MC	33,011	0.4	0.59	0.44
		0	5	MC	33,011	0.3	0.62	0.49
		0	6	MC	33,011	0.6	0.44	0.37
		0	10	MC	33,011	0.7	0.40	0.40
		0	11	MC	33,011	0.6	0.78	0.23
		0	12	MC	33,011	0.9	0.48	0.51
		0	13	MC	33,011	1.1	0.31	0.34
		0	14	SA	33,011	1.1	0.52	0.56
		0	15	SA	33,011	0.6	0.53	0.49
		0	17	SA	33,011	2.0	0.37	0.52
		0	18	CR	33,011	1.0	0.40	0.66
		0	21	SA	33,011	2.1	0.33	0.64
		0	22	MC	33,011	0.2	0.75	0.43
		0	23	MC	33,011	0.4	0.77	0.36
		0	24	MC	33,011	0.5	0.57	0.44
		0	25	MC	33,011	0.3	0.43	0.30
		0	29	MC	33,011	0.3	0.72	0.33
		0	30	MC	33,011	0.4	0.87	0.40
		0	31	MC	33,011	0.4	0.61	0.38
		0	32	MC	33,011	0.4	0.89	0.25
		0	33	MC	33,011	0.7	0.60	0.38
		0	34	MC	33,011	0.6	0.78	0.44
		0	35	MC	33,011	1.3	0.73	0.43
		0	37	SA	33,011	0.7	0.68	0.39
		0	40	CR	33,011	1.6	0.25	0.60
		0	41	SA	33,011	2.5	0.43	0.57
		0	42	CR	33,011	2.3	0.47	0.69
		0	43	SA	33,011	3.5	0.21	0.52
		0	44	MC	33,011	0.3	0.48	0.48
		0	45	MC	33,011	0.4	0.40	0.54
		0	46	MC	33,011	0.5	0.69	0.39
		0	47	MC	33,011	0.5	0.61	0.36
		0	48	MC	33,011	0.4	0.56	0.40
		0	52	MC	33,011	0.7	0.54	0.28
		0	53	MC	33,011	0.5	0.65	0.26
		0	54	MC	33,011	0.5	0.37	0.27
		0	55	MC	33,011	0.8	0.40	0.23
		0	56	MC	33,011	1.0	0.28	0.26
		0	57	MC	33,011	1.5	0.72	0.46
		0	58	SA	33,011	1.4	0.49	0.36
		0	59	SA	33,011	1.1	0.48	0.60
		0	60	SA	33,011	1.4	0.38	0.49
		0	62	CR	33,011	1.4	0.38	0.64
		0	63	SA	33,011	1.8	0.31	0.54
		0	65	SA	33,011	2.6	0.36	0.64

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	7	MC	3,752	0.9	0.54	0.44
		1	9	MC	3,752	0.7	0.67	0.50
		1	19	SA	3,752	1.7	0.28	0.47
		1	20	CR	3,752	1.2	0.38	0.62
		1	26	MC	3,752	0.4	0.47	0.51
		1	28	MC	3,752	0.3	0.60	0.28
		1	36	SA	3,752	2.4	0.33	0.58
		1	38	SA	3,752	2.2	0.46	0.54
		1	49	MC	3,752	0.6	0.31	0.20
		1	51	MC	3,752	0.4	0.51	0.35
		1	61	SA	3,752	1.6	0.31	0.51
		2	7	MC	3,681	0.6	0.53	0.33
		2	9	MC	3,681	0.7	0.38	0.46
		2	19	SA	3,681	1.7	0.27	0.67
		2	26	MC	3,681	0.2	0.66	0.40
		2	28	MC	3,681	0.8	0.50	0.37
		2	36	SA	3,681	2.5	0.42	0.44
		2	49	MC	3,681	0.7	0.75	0.43
		2	51	MC	3,681	0.3	0.61	0.44
		2	61	SA	3,681	1.0	0.35	0.52
		2	64	CR	3,681	2.0	0.56	0.52
		3	7	MC	3,661	1.1	0.41	0.44
		3	9	MC	3,661	1.1	0.43	0.47
		3	26	MC	3,661	0.5	0.56	0.37
7	Mathematics	3	28	MC	3,661	0.4	0.42	0.24
		3	36	SA	3,661	0.8	0.71	0.44
		3	39	SA	3,661	1.5	0.34	0.65
		3	49	MC	3,661	0.4	0.74	0.43
		3	51	MC	3,661	0.3	0.47	0.37
		3	61	SA	3,661	3.0	0.28	0.63
		3	64	CR	3,661	2.0	0.24	0.67
		4	7	MC	3,676	0.5	0.86	0.39
		4	9	MC	3,676	0.8	0.62	0.47
		4	16	SA	3,676	2.1	0.58	0.56
		4	20	CR	3,676	2.0	0.24	0.69
		4	26	MC	3,676	0.4	0.41	0.45
		4	28	MC	3,676	0.4	0.80	0.41
		4	36	SA	3,676	1.7	0.36	0.50
		4	39	SA	3,676	1.6	0.45	0.45
		4	49	MC	3,676	0.5	0.68	0.26
		4	51	MC	3,676	0.5	0.67	0.50
		4	61	SA	3,676	1.8	0.28	0.41
		5	7	MC	3,664	0.5	0.44	0.25
		5	9	MC	3,664	0.6	0.28	0.39
		5	16	SA	3,664	1.6	0.20	0.49
		5	20	CR	3,664	1.5	0.51	0.63
		5	26	MC	3,664	0.4	0.37	0.25
		5	28	MC	3,664	0.3	0.30	0.20
		5	36	SA	3,664	1.8	0.45	0.57
		5	38	SA	3,664	1.3	0.35	0.44

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		5	39	SA	3,664	1.6	0.27	0.54
		5	49	MC	3,664	0.5	0.60	0.43
		5	51	MC	3,664	0.6	0.45	0.40
		5	61	SA	3,664	1.7	0.44	0.59
		6	7	MC	3,641	0.3	0.88	0.36
		6	9	MC	3,641	0.5	0.56	0.31
		6	16	SA	3,641	0.7	0.34	0.28
		6	19	SA	3,641	1.3	0.34	0.42
		6	26	MC	3,641	0.6	0.51	0.44
		6	28	MC	3,641	0.7	0.75	0.30
		6	36	SA	3,641	1.0	0.67	0.42
		6	38	SA	3,641	0.7	0.47	0.47
		6	49	MC	3,641	0.5	0.57	0.53
		6	51	MC	3,641	0.4	0.66	0.33
		6	61	SA	3,641	2.6	0.51	0.36
		6	64	CR	3,641	3.0	0.21	0.63
		7	7	MC	3,642	0.6	0.53	0.44
		7	9	MC	3,642	0.6	0.69	0.50
		7	19	SA	3,642	1.4	0.26	0.44
		7	20	CR	3,642	1.2	0.38	0.63
		7	26	MC	3,642	0.2	0.46	0.50
		7	28	MC	3,642	0.3	0.59	0.28
		7	36	SA	3,642	1.9	0.34	0.57
7	Mathematics	7	38	SA	3,642	2.3	0.47	0.55
		7	49	MC	3,642	0.4	0.31	0.19
		7	51	MC	3,642	0.3	0.52	0.38
		7	61	SA	3,642	1.2	0.31	0.52
		8	7	MC	3,643	0.9	0.53	0.34
		8	9	MC	3,643	0.9	0.39	0.51
		8	19	SA	3,643	2.3	0.28	0.66
		8	26	MC	3,643	0.4	0.66	0.42
		8	28	MC	3,643	0.7	0.50	0.37
		8	36	SA	3,643	3.2	0.41	0.46
		8	49	MC	3,643	0.5	0.75	0.42
		8	51	MC	3,643	0.4	0.59	0.43
		8	61	SA	3,643	1.7	0.39	0.54
		8	64	CR	3,643	2.5	0.59	0.53
		9	7	MC	3,644	0.9	0.41	0.43
		9	9	MC	3,644	0.7	0.43	0.44
		9	26	MC	3,644	0.3	0.55	0.36
		9	28	MC	3,644	0.2	0.43	0.23
		9	36	SA	3,644	0.7	0.70	0.46
		9	39	SA	3,644	1.8	0.33	0.65
		9	49	MC	3,644	0.3	0.75	0.42
		9	51	MC	3,644	0.3	0.46	0.34
		9	61	SA	3,644	2.8	0.27	0.61
		9	64	CR	3,644	1.9	0.23	0.65

Table F-11. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 7 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
7	Reading	0	1	MC	32,915	0.3	0.88	0.42
		0	2	MC	32,915	0.3	0.79	0.34
		0	3	MC	32,915	0.2	0.72	0.31
		0	4	MC	32,915	0.3	0.80	0.41
		0	5	MC	32,915	0.3	0.92	0.38
		0	6	MC	32,915	0.7	0.62	0.25
		0	7	CR	32,915	0.4	0.58	0.58
		0	8	MC	32,915	0.2	0.89	0.34
		0	9	MC	32,915	0.4	0.61	0.36
		0	10	MC	32,915	0.3	0.88	0.36
		0	11	MC	32,915	0.3	0.85	0.42
		0	12	CR	32,915	0.9	0.50	0.66
		0	13	MC	32,915	0.5	0.65	0.40
		0	14	MC	32,915	0.6	0.85	0.46
		0	15	MC	32,915	0.6	0.68	0.42
		0	16	MC	32,915	0.6	0.79	0.47
		0	17	CR	32,915	1.2	0.52	0.64
		0	25	MC	32,915	0.5	0.48	0.24
		0	26	MC	32,915	0.4	0.89	0.40
		0	27	MC	32,915	0.4	0.89	0.40
		0	28	MC	32,915	0.6	0.62	0.33
		0	29	CR	32,915	1.1	0.48	0.61
		0	30	MC	32,915	0.6	0.58	0.35
		0	31	MC	32,915	0.7	0.84	0.47
		0	32	MC	32,915	0.7	0.80	0.47
		0	33	MC	32,915	0.8	0.90	0.36
		0	34	CR	32,915	1.2	0.51	0.58
		0	35	MC	32,915	0.3	0.74	0.34
		0	36	MC	32,915	0.4	0.52	0.24
		0	37	MC	32,915	0.4	0.78	0.20
		0	38	MC	32,915	0.3	0.89	0.32
		0	39	MC	32,915	0.3	0.78	0.25
		0	40	MC	32,915	0.6	0.71	0.28
		0	41	CR	32,915	0.6	0.57	0.56
		1	18	MC	3,741	0.3	0.90	0.44
		1	19	MC	3,741	0.3	0.73	0.42
		1	20	MC	3,741	0.3	0.71	0.40
		1	21	MC	3,741	0.4	0.65	0.33
		1	22	MC	3,741	0.3	0.81	0.42
		1	23	MC	3,741	0.6	0.50	0.34
		1	24	CR	3,741	0.7	0.52	0.67
		1	42	MC	3,741	0.5	0.77	0.36
		1	43	MC	3,741	0.4	0.70	0.44
		1	44	MC	3,741	0.4	0.73	0.43
		1	45	MC	3,741	0.7	0.81	0.42
		1	46	CR	3,741	0.8	0.53	0.62
		1	47	MC	3,741	0.5	0.50	0.38
		1	48	MC	3,741	0.6	0.75	0.45

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,741	0.5	0.77	0.45
		1	50	MC	3,741	0.6	0.55	0.28
		1	51	CR	3,741	0.8	0.45	0.60
		2	18	MC	3,672	0.2	0.55	0.30
		2	19	MC	3,672	0.3	0.66	0.35
		2	20	MC	3,672	0.2	0.75	0.41
		2	21	MC	3,672	0.2	0.87	0.39
		2	22	MC	3,672	0.3	0.86	0.45
		2	23	MC	3,672	0.6	0.48	0.27
		2	24	CR	3,672	0.8	0.48	0.63
		2	42	MC	3,672	0.4	0.67	0.15
		2	43	MC	3,672	0.2	0.69	0.29
		2	44	MC	3,672	0.3	0.68	0.40
		2	45	MC	3,672	0.2	0.51	0.34
		2	46	CR	3,672	0.9	0.49	0.68
		2	47	MC	3,672	0.5	0.77	0.45
		2	48	MC	3,672	0.4	0.76	0.47
		2	49	MC	3,672	0.5	0.78	0.47
7	Reading	2	50	MC	3,672	0.4	0.88	0.47
		2	51	CR	3,672	1.0	0.51	0.61
		3	18	MC	3,656	0.3	0.86	0.30
		3	19	MC	3,656	0.5	0.78	0.35
		3	20	MC	3,656	0.4	0.80	0.50
		3	21	MC	3,656	0.6	0.72	0.28
		3	22	MC	3,656	0.5	0.76	0.20
		3	23	MC	3,656	0.8	0.67	0.31
		3	24	CR	3,656	0.5	0.49	0.61
		3	42	MC	3,656	0.4	0.85	0.47
		3	43	MC	3,656	0.5	0.84	0.38
		3	44	MC	3,656	0.4	0.87	0.51
		3	45	MC	3,656	0.6	0.63	0.32
		3	46	CR	3,656	0.6	0.53	0.66
		3	47	MC	3,656	0.5	0.77	0.46
		3	48	MC	3,656	0.6	0.92	0.39
		3	49	MC	3,656	0.8	0.77	0.39
		3	50	MC	3,656	0.7	0.92	0.41
		3	51	CR	3,656	0.9	0.54	0.60

Table F-12. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 8 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
8	Mathematics	0	1	MC	33,771	0.2	0.85	0.38
		0	2	MC	33,771	0.4	0.56	0.40
		0	3	MC	33,771	0.4	0.83	0.41
		0	4	MC	33,771	0.5	0.43	0.23
		0	5	MC	33,771	0.2	0.76	0.51
		0	6	MC	33,771	0.4	0.74	0.54
		0	10	MC	33,771	0.6	0.45	0.54
		0	11	MC	33,771	0.5	0.27	0.39
		0	12	MC	33,771	0.6	0.41	0.46
		0	13	MC	33,771	0.6	0.70	0.47
		0	14	SA	33,771	0.6	0.67	0.46
		0	15	SA	33,771	1.5	0.57	0.55
		0	17	SA	33,771	0.7	0.56	0.61
		0	18	CR	33,771	1.3	0.36	0.67
		0	21	SA	33,771	3.0	0.59	0.49
		0	22	MC	33,771	0.4	0.79	0.35
		0	23	MC	33,771	0.5	0.56	0.40
		0	24	MC	33,771	0.7	0.46	0.39
		0	25	MC	33,771	0.6	0.57	0.39
		0	29	MC	33,771	0.4	0.54	0.35
		0	30	MC	33,771	0.5	0.59	0.45
		0	31	MC	33,771	0.5	0.47	0.31
		0	32	MC	33,771	0.6	0.54	0.48
		0	33	MC	33,771	0.7	0.53	0.41
		0	34	MC	33,771	0.7	0.57	0.45
		0	35	MC	33,771	1.6	0.39	0.38
		0	37	SA	33,771	2.0	0.45	0.51
		0	40	CR	33,771	3.0	0.13	0.56
		0	41	SA	33,771	1.9	0.57	0.65
		0	42	CR	33,771	1.6	0.42	0.68
		0	43	SA	33,771	4.4	0.22	0.47
		0	44	MC	33,771	0.5	0.56	0.54
		0	45	MC	33,771	0.7	0.25	0.36
		0	46	MC	33,771	0.6	0.85	0.40
		0	47	MC	33,771	0.5	0.84	0.39
		0	48	MC	33,771	0.6	0.53	0.38
		0	52	MC	33,771	0.8	0.50	0.26
		0	53	MC	33,771	0.7	0.68	0.47
		0	54	MC	33,771	0.8	0.62	0.51
		0	55	MC	33,771	0.9	0.51	0.28
		0	56	MC	33,771	1.1	0.55	0.48
		0	57	MC	33,771	1.4	0.57	0.41
		0	58	SA	33,771	0.9	0.81	0.44
		0	59	SA	33,771	1.5	0.69	0.28
		0	60	SA	33,771	1.9	0.35	0.50
		0	62	CR	33,771	1.7	0.58	0.73
		0	63	SA	33,771	3.3	0.26	0.40
		0	65	SA	33,771	3.1	0.45	0.61

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	7	MC	3,789	0.3	0.59	0.52
		1	9	MC	3,789	0.7	0.36	0.26
		1	19	SA	3,789	1.5	0.68	0.59
		1	26	MC	3,789	0.6	0.61	0.35
		1	28	MC	3,789	0.6	0.57	0.39
		1	36	SA	3,789	1.4	0.56	0.50
		1	39	SA	3,789	3.3	0.54	0.64
		1	49	MC	3,789	1.0	0.71	0.55
		1	51	MC	3,789	0.9	0.74	0.46
		1	61	SA	3,789	2.9	0.32	0.60
		1	64	CR	3,789	2.6	0.31	0.69
		2	7	MC	3,750	0.2	0.81	0.38
		2	9	MC	3,750	0.5	0.40	0.52
		2	16	SA	3,750	1.5	0.28	0.50
		2	26	MC	3,750	0.5	0.34	0.19
		2	28	MC	3,750	0.4	0.53	0.46
		2	36	SA	3,750	2.0	0.28	0.53
		2	39	SA	3,750	1.5	0.49	0.59
		2	49	MC	3,750	0.5	0.66	0.48
		2	51	MC	3,750	0.5	0.59	0.44
		2	61	SA	3,750	1.3	0.36	0.39
		2	64	CR	3,750	3.4	0.32	0.72
		3	7	MC	3,745	0.2	0.59	0.46
		3	9	MC	3,745	0.9	0.43	0.24
8	Mathematics	3	16	SA	3,745	2.1	0.48	0.59
		3	19	SA	3,745	1.7	0.36	0.51
		3	20	CR	3,745	2.0	0.46	0.61
		3	26	MC	3,745	0.5	0.32	0.39
		3	28	MC	3,745	0.4	0.68	0.29
		3	36	SA	3,745	3.0	0.30	0.54
		3	38	SA	3,745	1.5	0.62	0.61
		3	49	MC	3,745	0.8	0.32	0.23
		3	51	MC	3,745	0.6	0.73	0.46
		3	61	SA	3,745	4.0	0.11	0.39
		4	7	MC	3,766	0.2	0.58	0.47
		4	9	MC	3,766	0.3	0.74	0.51
		4	16	SA	3,766	2.4	0.49	0.59
		4	20	CR	3,766	1.6	0.29	0.67
		4	26	MC	3,766	0.4	0.85	0.49
		4	28	MC	3,766	0.6	0.75	0.41
		4	36	SA	3,766	1.0	0.76	0.33
		4	38	SA	3,766	3.2	0.38	0.60
		4	49	MC	3,766	0.9	0.36	0.25
		4	51	MC	3,766	0.8	0.62	0.38
		4	61	SA	3,766	2.2	0.59	0.46
		5	7	MC	3,765	0.5	0.25	0.23
		5	9	MC	3,765	0.7	0.60	0.50
		5	20	CR	3,765	1.4	0.53	0.63
		5	26	MC	3,765	0.5	0.33	0.37
		5	28	MC	3,765	0.3	0.50	0.50

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
8	Mathematics	5	36	SA	3,765	1.3	0.23	0.34
		5	49	MC	3,765	0.8	0.43	0.37
		5	51	MC	3,765	0.4	0.64	0.38
		5	61	SA	3,765	3.6	0.36	0.60
		6	7	MC	3,745	0.4	0.64	0.48
		6	9	MC	3,745	0.6	0.54	0.34
		6	19	SA	3,745	2.3	0.25	0.60
		6	26	MC	3,745	0.7	0.26	0.25
		6	28	MC	3,745	0.5	0.72	0.41
		6	36	SA	3,745	1.6	0.44	0.47
		6	38	SA	3,745	2.7	0.35	0.56
		6	39	SA	3,745	4.3	0.35	0.54
		6	49	MC	3,745	0.9	0.70	0.28
		6	51	MC	3,745	0.9	0.34	0.44
		6	61	SA	3,745	2.6	0.42	0.62
		6	64	CR	3,745	2.4	0.52	0.71
		7	7	MC	3,775	0.3	0.58	0.52
		7	9	MC	3,775	0.5	0.36	0.25
		7	19	SA	3,775	1.3	0.68	0.57
		7	26	MC	3,775	0.5	0.62	0.33
		7	28	MC	3,775	0.5	0.55	0.39
		7	36	SA	3,775	1.2	0.58	0.49
		7	39	SA	3,775	2.5	0.55	0.64
		7	49	MC	3,775	0.6	0.71	0.54
		7	51	MC	3,775	0.5	0.75	0.43
		7	61	SA	3,775	2.7	0.32	0.58
		7	64	CR	3,775	1.8	0.31	0.67
		8	7	MC	3,752	0.4	0.81	0.36
		8	9	MC	3,752	0.4	0.41	0.50
		8	16	SA	3,752	1.6	0.29	0.49
		8	26	MC	3,752	0.6	0.36	0.20
		8	28	MC	3,752	0.5	0.53	0.48
		8	36	SA	3,752	1.7	0.29	0.52
		8	39	SA	3,752	1.9	0.48	0.60
		8	49	MC	3,752	0.6	0.70	0.49
		8	51	MC	3,752	0.6	0.62	0.41
		8	61	SA	3,752	1.1	0.35	0.38
		8	64	CR	3,752	2.6	0.33	0.72
		9	7	MC	3,677	0.2	0.60	0.47
		9	9	MC	3,677	0.8	0.43	0.22
		9	16	SA	3,677	2.5	0.51	0.58
		9	19	SA	3,677	1.4	0.37	0.49
		9	20	CR	3,677	1.8	0.48	0.61
		9	26	MC	3,677	0.4	0.33	0.38
		9	28	MC	3,677	0.4	0.69	0.29
		9	36	SA	3,677	3.3	0.31	0.52
		9	38	SA	3,677	1.6	0.65	0.60
		9	49	MC	3,677	0.7	0.33	0.27
		9	51	MC	3,677	0.7	0.76	0.45
		9	61	SA	3,677	3.7	0.11	0.39

Table F-13. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 8 Reading

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
8	Reading	0	1	MC	33,698	0.3	0.88	0.29
		0	2	MC	33,698	0.4	0.91	0.33
		0	3	MC	33,698	0.2	0.92	0.39
		0	4	MC	33,698	0.3	0.78	0.40
		0	5	MC	33,698	0.3	0.79	0.23
		0	6	MC	33,698	0.5	0.90	0.41
		0	7	CR	33,698	1.1	0.51	0.63
		0	8	MC	33,698	0.4	0.77	0.39
		0	9	MC	33,698	0.3	0.77	0.31
		0	10	MC	33,698	0.4	0.76	0.46
		0	11	MC	33,698	0.6	0.71	0.47
		0	12	CR	33,698	0.8	0.49	0.67
		0	13	MC	33,698	0.5	0.69	0.36
		0	14	MC	33,698	0.6	0.85	0.48
		0	15	MC	33,698	0.6	0.75	0.43
		0	16	MC	33,698	0.7	0.88	0.37
		0	17	CR	33,698	0.9	0.41	0.65
		0	25	MC	33,698	0.5	0.64	0.17
		0	26	MC	33,698	0.5	0.79	0.33
		0	27	MC	33,698	0.5	0.81	0.34
		0	28	MC	33,698	0.7	0.72	0.40
		0	29	CR	33,698	1.7	0.45	0.55
		0	30	MC	33,698	0.6	0.87	0.30
		0	31	MC	33,698	0.9	0.80	0.30
		0	32	MC	33,698	0.7	0.43	0.32
		0	33	MC	33,698	0.8	0.88	0.45
		0	34	CR	33,698	1.4	0.50	0.57
		0	35	MC	33,698	0.4	0.83	0.44
		0	36	MC	33,698	0.6	0.67	0.22
		0	37	MC	33,698	0.4	0.86	0.43
		0	38	MC	33,698	0.6	0.64	0.19
		0	39	MC	33,698	0.6	0.67	0.40
		0	40	MC	33,698	0.7	0.70	0.26
		0	41	CR	33,698	1.0	0.54	0.66
		1	18	MC	3,784	0.4	0.81	0.38
		1	19	MC	3,784	0.5	0.77	0.44
		1	20	MC	3,784	0.2	0.34	0.30
		1	21	MC	3,784	0.5	0.68	0.46
		1	22	MC	3,784	0.3	0.74	0.51
		1	23	MC	3,784	0.9	0.75	0.51
		1	24	CR	3,784	1.6	0.51	0.67
		1	42	MC	3,784	0.6	0.91	0.43
		1	43	MC	3,784	0.7	0.89	0.46
		1	44	MC	3,784	0.6	0.82	0.42
		1	45	MC	3,784	0.9	0.84	0.48
		1	46	CR	3,784	1.8	0.49	0.69
		1	47	MC	3,784	0.8	0.88	0.43
		1	48	MC	3,784	1.4	0.65	0.23

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	49	MC	3,784	1.0	0.67	0.32
		1	50	MC	3,784	1.1	0.83	0.43
		1	51	CR	3,784	1.7	0.48	0.61
		2	18	MC	3,743	0.4	0.66	0.30
		2	19	MC	3,743	0.5	0.68	0.37
		2	20	MC	3,743	0.5	0.72	0.45
		2	21	MC	3,743	0.4	0.79	0.35
		2	22	MC	3,743	0.6	0.75	0.44
		2	23	MC	3,743	0.7	0.70	0.35
		2	24	CR	3,743	1.1	0.52	0.66
		2	42	MC	3,743	0.4	0.79	0.48
		2	43	MC	3,743	0.5	0.52	0.38
		2	44	MC	3,743	0.5	0.60	0.30
		2	45	MC	3,743	0.7	0.70	0.39
		2	46	CR	3,743	1.5	0.50	0.62
		2	47	MC	3,743	0.8	0.67	0.34
		2	48	MC	3,743	0.9	0.89	0.46
		2	49	MC	3,743	0.7	0.84	0.40
8	Reading	2	50	MC	3,743	1.0	0.73	0.41
		2	51	CR	3,743	1.5	0.58	0.70
		3	18	MC	3,735	0.5	0.54	0.30
		3	19	MC	3,735	0.6	0.60	0.31
		3	20	MC	3,735	0.5	0.70	0.47
		3	21	MC	3,735	0.5	0.86	0.38
		3	22	MC	3,735	0.5	0.79	0.40
		3	23	MC	3,735	0.6	0.85	0.49
		3	24	CR	3,735	1.1	0.53	0.67
		3	42	MC	3,735	0.7	0.71	0.37
		3	43	MC	3,735	0.8	0.79	0.42
		3	44	MC	3,735	0.7	0.87	0.32
		3	45	MC	3,735	0.9	0.67	0.41
		3	46	CR	3,735	2.6	0.58	0.70
		3	47	MC	3,735	1.3	0.50	0.38
		3	48	MC	3,735	1.4	0.87	0.42
		3	49	MC	3,735	1.3	0.56	0.44
		3	50	MC	3,735	1.3	0.63	0.39
		3	51	CR	3,735	1.7	0.41	0.68

Table F-14. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 8 Writing

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
		1	1	MC	33,563	0.2	0.92	0.23
		1	2	MC	33,563	0.2	0.94	0.32
		1	3	MC	33,563	0.2	0.86	0.31
		1	4	MC	33,563	0.3	0.85	0.28
		1	5	MC	33,563	0.3	0.81	0.39
		1	6	MC	33,563	0.3	0.68	0.29
		1	7	MC	33,563	0.3	0.80	0.32
		1	8	MC	33,563	0.4	0.66	0.34
8	Writing	1	9	MC	33,563	0.3	0.62	0.31
		1	10	MC	33,563	0.6	0.49	0.32
		1	11	CR	33,563	0.5	0.61	0.64
		1	12	CR	33,563	1.5	0.64	0.67
		1	13	CR	33,563	1.4	0.54	0.58
		1	14	SA	33,563	0.5	0.96	0.36
		1	15	SA	33,563	0.5	0.96	0.37
		1	16	SA	33,563	0.5	0.94	0.34
		1	17	WR	33,563	0.5	0.53	0.70

Table F-15. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 11 Mathematics

Grade	*Content*	*Form*	*Position*	*Item Type*	*N*	*Omit Rate*	*Difficulty*	*Discrimination*
11	Mathematics	0	1	MC	33,771	0.2	0.75	0.36
		0	2	MC	33,771	0.5	0.53	0.39
		0	3	MC	33,771	0.8	0.54	0.56
		0	4	MC	33,771	1.3	0.58	0.39
		0	5	MC	33,771	1.4	0.29	0.30
		0	6	MC	33,771	0.4	0.41	0.48
		0	11	MC	33,771	1.4	0.55	0.51
		0	12	MC	33,771	1.0	0.57	0.36
		0	13	MC	33,771	1.1	0.30	0.36
		0	14	MC	33,771	1.3	0.31	0.30
		0	15	MC	33,771	1.1	0.25	0.53
		0	16	MC	33,771	1.7	0.34	0.28
		0	17	SA	33,771	10.2	0.47	0.62
		0	18	SA	33,771	10.5	0.14	0.45
		0	19	SA	33,771	10.1	0.24	0.54
		0	22	SA	33,771	8.2	0.15	0.46
		0	23	SA	33,771	8.1	0.35	0.59
		0	24	SA	33,771	16.3	0.08	0.42
		0	25	SA	33,771	12.9	0.33	0.50
		0	26	CR	33,771	6.7	0.45	0.70
		0	29	SA	33,771	15.4	0.30	0.56
		0	30	CR	33,771	14.2	0.18	0.68
		0	31	SA	33,771	26.3	0.09	0.45
		0	32	MC	33,771	0.7	0.83	0.32
		0	33	MC	33,771	1.4	0.61	0.38
		0	34	MC	33,771	1.7	0.38	0.43
		0	35	MC	33,771	1.2	0.63	0.40
		0	36	MC	33,771	1.2	0.81	0.32
		0	37	MC	33,771	1.0	0.56	0.30
		0	42	MC	33,771	1.2	0.31	0.35
		0	43	MC	33,771	1.3	0.59	0.42
		0	44	MC	33,771	1.2	0.53	0.49
		0	45	MC	33,771	1.4	0.49	0.18
		0	46	MC	33,771	1.5	0.42	0.57
		0	47	MC	33,771	1.8	0.34	0.42
		0	48	SA	33,771	4.9	0.29	0.40
		0	49	SA	33,771	4.9	0.59	0.57
		0	50	SA	33,771	2.7	0.80	0.32
		0	53	SA	33,771	9.3	0.32	0.50
		0	54	SA	33,771	5.2	0.54	0.57
		0	55	SA	33,771	7.3	0.37	0.55
		0	56	SA	33,771	6.7	0.60	0.56
		0	57	CR	33,771	7.1	0.39	0.66
		0	60	SA	33,771	14.4	0.24	0.59
		0	61	CR	33,771	10.8	0.26	0.63
		0	62	SA	33,771	17.1	0.16	0.50
		1	7	MC	4,270	1.0	0.66	0.40
		1	9	MC	4,270	1.4	0.20	0.32

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		1	20	SA	4,270	5.3	0.29	0.59
		1	27	SA	4,270	9.8	0.30	0.62
		1	38	MC	4,270	1.1	0.54	0.50
		1	40	MC	4,270	1.3	0.49	0.39
		1	51	SA	4,270	5.8	0.40	0.60
		1	59	CR	4,270	14.4	0.24	0.70
		2	7	MC	4,236	1.3	0.40	0.42
		2	9	MC	4,236	1.0	0.37	0.16
		2	20	SA	4,236	7.1	0.23	0.49
		2	28	CR	4,236	11.0	0.25	0.70
		2	38	MC	4,236	1.1	0.69	0.42
		2	40	MC	4,236	1.0	0.49	0.48
		2	51	SA	4,236	4.5	0.48	0.54
		2	58	SA	4,236	6.8	0.08	0.40
		3	7	MC	4,185	1.2	0.56	0.39
		3	9	MC	4,185	1.5	0.37	0.21
		3	20	SA	4,185	8.5	0.37	0.53
		3	28	CR	4,185	6.4	0.37	0.62
		3	38	MC	4,185	1.1	0.54	0.45
		3	40	MC	4,185	1.4	0.33	0.49
		3	51	SA	4,185	13.7	0.09	0.38
		3	58	SA	4,185	8.8	0.12	0.47
		4	7	MC	4,241	1.3	0.43	0.31
		4	9	MC	4,241	1.1	0.46	0.46
11	Mathematics	4	20	SA	4,241	6.2	0.59	0.49
		4	28	CR	4,241	20.2	0.07	0.56
		4	38	MC	4,241	1.7	0.35	0.34
		4	40	MC	4,241	1.0	0.31	0.32
		4	51	SA	4,241	8.3	0.15	0.54
		4	58	SA	4,241	12.3	0.24	0.55
		5	7	MC	4,181	1.1	0.49	0.39
		5	9	MC	4,181	1.0	0.40	0.05
		5	20	SA	4,181	8.4	0.24	0.59
		5	27	SA	4,181	9.7	0.18	0.49
		5	38	MC	4,181	1.3	0.38	0.27
		5	40	MC	4,181	1.1	0.29	0.34
		5	51	SA	4,181	5.4	0.24	0.56
		5	59	CR	4,181	6.4	0.48	0.66
		6	7	MC	4,224	1.9	0.41	0.34
		6	9	MC	4,224	1.1	0.54	0.47
		6	20	SA	4,224	6.9	0.34	0.49
		6	27	SA	4,224	16.8	0.13	0.36
		6	38	MC	4,224	1.0	0.64	0.47
		6	40	MC	4,224	1.1	0.31	0.17
		6	51	SA	4,224	7.6	0.41	0.61
		6	59	CR	4,224	12.7	0.17	0.68
		7	7	MC	4,212	1.0	0.67	0.41
		7	9	MC	4,212	1.4	0.21	0.37
		7	20	SA	4,212	5.3	0.29	0.57
		7	27	SA	4,212	9.9	0.29	0.60

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
11	Mathematics	7	38	MC	4,212	1.2	0.53	0.49
		7	40	MC	4,212	1.3	0.50	0.40
		7	51	SA	4,212	5.7	0.40	0.61
		7	59	CR	4,212	14.2	0.23	0.70
		8	7	MC	4,214	1.3	0.40	0.40
		8	9	MC	4,214	1.1	0.36	0.13
		8	20	SA	4,214	7.1	0.22	0.50
		8	28	CR	4,214	11.4	0.24	0.70
		8	38	MC	4,214	1.0	0.68	0.41
		8	40	MC	4,214	1.1	0.48	0.49
		8	51	SA	4,214	3.8	0.48	0.51
		8	58	SA	4,214	7.1	0.09	0.41

Table F-16. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 11 Reading

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
11	Reading	0	1	MC	33,702	0.5	0.86	0.38
		0	2	MC	33,702	0.4	0.54	0.20
		0	3	MC	33,702	0.4	0.84	0.32
		0	4	MC	33,702	0.6	0.68	0.20
		0	7	MC	33,702	0.3	0.77	0.31
		0	8	MC	33,702	0.5	0.71	0.41
		0	9	MC	33,702	0.4	0.72	0.42
		0	10	MC	33,702	0.8	0.69	0.39
		0	11	CR	33,702	1.6	0.57	0.60
		0	12	MC	33,702	0.5	0.85	0.46
		0	13	MC	33,702	0.5	0.83	0.45
		0	14	MC	33,702	0.5	0.56	0.26
		0	15	MC	33,702	0.4	0.92	0.34
		0	16	CR	33,702	2.8	0.56	0.66
		0	17	MC	33,702	0.5	0.87	0.39
		0	18	MC	33,702	0.6	0.69	0.37
		0	19	MC	33,702	0.6	0.74	0.36
		0	20	MC	33,702	0.8	0.80	0.47
		0	21	CR	33,702	2.5	0.61	0.68
		0	22	MC	33,702	1.3	0.70	0.36
		0	23	MC	33,702	1.2	0.76	0.33
		0	24	MC	33,702	1.3	0.74	0.40
		0	25	MC	33,702	1.4	0.66	0.35
		0	26	CR	33,702	4.4	0.47	0.66
		0	27	MC	33,702	0.8	0.76	0.23
		0	28	MC	33,702	0.9	0.75	0.32
		0	29	MC	33,702	1.0	0.75	0.46
		0	30	MC	33,702	0.9	0.76	0.39
		0	31	CR	33,702	2.4	0.56	0.62
		0	32	MC	33,702	0.9	0.88	0.44
		0	33	MC	33,702	1.0	0.82	0.44

(cont'd)

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		0	34	MC	33,702	1.0	0.71	0.49
		0	35	MC	33,702	1.0	0.56	0.29
		0	36	CR	33,702	3.8	0.49	0.65
		1	5	MC	4,267	0.4	0.82	0.30
		1	6	MC	4,267	0.6	0.67	0.33
		1	37	MC	4,267	1.2	0.66	0.33
		1	38	MC	4,267	1.4	0.44	0.41
		1	39	MC	4,267	1.2	0.91	0.49
		1	40	MC	4,267	1.4	0.61	0.34
		1	41	CR	4,267	3.4	0.46	0.70
		1	42	MC	4,267	1.5	0.72	0.39
		1	43	MC	4,267	2.0	0.49	0.31
		1	44	MC	4,267	1.6	0.79	0.56
		1	45	MC	4,267	1.6	0.84	0.45
		1	46	CR	4,267	4.8	0.44	0.68
		1	47	MC	4,267	2.2	0.72	0.46
		1	48	MC	4,267	2.3	0.82	0.45
		1	49	MC	4,267	2.3	0.66	0.50
11	Reading	1	50	MC	4,267	2.4	0.82	0.44
		1	51	CR	4,267	4.4	0.55	0.72
		2	5	MC	4,217	0.4	0.84	0.38
		2	6	MC	4,217	0.5	0.70	0.40
		2	37	MC	4,217	1.0	0.74	0.42
		2	38	MC	4,217	1.2	0.76	0.49
		2	39	MC	4,217	1.0	0.79	0.54
		2	40	MC	4,217	1.2	0.71	0.39
		2	41	CR	4,217	3.3	0.45	0.67
		2	42	MC	4,217	1.5	0.65	0.45
		2	43	MC	4,217	1.5	0.87	0.55
		2	44	MC	4,217	1.5	0.73	0.40
		2	45	MC	4,217	1.5	0.77	0.46
		2	46	CR	4,217	6.5	0.42	0.71
		2	47	MC	4,217	1.6	0.53	0.20
		2	48	MC	4,217	1.7	0.71	0.39
		2	49	MC	4,217	1.6	0.61	0.29
		2	50	MC	4,217	1.7	0.87	0.53
		2	51	CR	4,217	4.5	0.47	0.72

Table F-17. 2008-09 NECAP: Item Level Classical Stats by Grade, Content and Form—Grade 11 Writing

Grade	Content	Form	Position	Item Type	N	Omit Rate	Difficulty	Discrimination
		0	1	WR	33,653	0.0	0.50	
		1	2	WR	4,258	0.8	0.44	0.63
11	Writing	2	2	WR	4,218	1.2	0.41	0.61
		3	2	WR	4,166	0.8	0.46	0.61
		6	2	WR	4,208	0.8	0.49	0.58
		7	2	WR	4,202	0.6	0.48	0.65

Appendix G—Classical Item Statistic Descriptives Summarized by Grade, Content and Form

Table G-1. 2008-09: NECAP: Item Difficulty and Discriminations Indices—by Grade, Content and Form

Grade	*Content*	*Form*	*Statistic*	*Difficulty*	*Discrimination*
3	Mathematics	0	Points	77	77
			Mean	0.67	0.44
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.94	0.61
			Range	0.62	0.4
		1	Points	77	77
			Mean	0.67	0.44
			StDev	0.16	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38
		2	Points	77	77
			Mean	0.67	0.44
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.91	0.60
			Range	0.59	0.39
		3	Points	77	77
			Mean	0.67	0.43
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38
		4	Points	77	77
			Mean	0.67	0.43
			StDev	0.16	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38
		5	Points	77	77
			Mean	0.67	0.43
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38
		6	Points	77	77
			Mean	0.67	0.44
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.94	0.61
			Range	0.62	0.40
		7	Points	77	77
			Mean	0.67	0.44
			StDev	0.16	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
3	Mathematics	8	Points	77	77
			Mean	0.67	0.44
			StDev	0.17	0.09
			Min	0.32	0.21
			Max	0.94	0.61
			Range	0.62	0.40
		9	Points	77	77
			Mean	0.67	0.44
			StDev	0.16	0.09
			Min	0.32	0.21
			Max	0.91	0.59
			Range	0.59	0.38
3	Reading	0	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		1	Points	78	78
			Mean	0.71	0.45
			StDev	0.14	0.09
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		2	Points	78	78
			Mean	0.72	0.46
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		3	Points	78	78
			Mean	0.72	0.47
			StDev	0.13	0.09
			Min	0.38	0.26
			Max	0.92	0.65
			Range	0.54	0.39
		4	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		5	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
3	Reading	6	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		7	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		8	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
		9	Points	52	52
			Mean	0.72	0.45
			StDev	0.14	0.08
			Min	0.38	0.26
			Max	0.91	0.65
			Range	0.53	0.39
4	Mathematics	0	Points	65	65
			Mean	0.68	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.90	0.63
			Range	0.68	0.38
		1	Points	77	77
			Mean	0.69	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.90	0.63
			Range	0.68	0.38
		2	Points	77	77
			Mean	0.69	0.44
			StDev	0.16	0.09
			Min	0.22	0.25
			Max	0.94	0.63
			Range	0.72	0.38
		3	Points	77	77
			Mean	0.68	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.92	0.63
			Range	0.70	0.38

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
4	Mathematics	4	Points	77	77
			Mean	0.68	0.44
			StDev	0.17	0.08
			Min	0.22	0.25
			Max	0.90	0.63
			Range	0.68	0.38
		5	Points	77	77
			Mean	0.68	0.43
			StDev	0.17	0.09
			Min	0.22	0.25
			Max	0.92	0.63
			Range	0.70	0.38
		6	Points	77	77
			Mean	0.68	0.44
			StDev	0.17	0.08
			Min	0.22	0.25
			Max	0.93	0.63
			Range	0.71	0.38
		7	Points	77	77
			Mean	0.68	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.90	0.63
			Range	0.68	0.38
		8	Points	77	77
			Mean	0.68	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.93	0.63
			Range	0.71	0.38
		9	Points	77	77
			Mean	0.68	0.44
			StDev	0.16	0.08
			Min	0.22	0.25
			Max	0.91	0.63
			Range	0.69	0.38
4	Reading	0	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		1	Points	78	78
			Mean	0.70	0.42
			StDev	0.14	0.08
			Min	0.37	0.26
			Max	0.90	0.58
			Range	0.53	0.32

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
4	Reading	2	Points	78	78
			Mean	0.7	0.42
			StDev	0.13	0.07
			Min	0.37	0.26
			Max	0.90	0.56
			Range	0.53	0.30
		3	Points	78	78
			Mean	0.70	0.42
			StDev	0.14	0.08
			Min	0.37	0.26
			Max	0.92	0.59
			Range	0.55	0.33
		4	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		5	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		6	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		7	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		8	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29
		9	Points	52	52
			Mean	0.71	0.41
			StDev	0.14	0.07
			Min	0.37	0.26
			Max	0.90	0.55
			Range	0.53	0.29

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
5	Mathematics	0	Points	66	66
			Mean	0.59	0.42
			StDev	0.14	0.11
			Min	0.36	0.16
			Max	0.89	0.66
			Range	0.53	0.50
		1	Points	82	82
			Mean	0.58	0.43
			StDev	0.16	0.11
			Min	0.23	0.16
			Max	0.91	0.67
			Range	0.68	0.51
		2	Points	82	82
			Mean	0.58	0.42
			StDev	0.14	0.12
			Min	0.36	0.16
			Max	0.89	0.67
			Range	0.53	0.51
		3	Points	79	79
			Mean	0.59	0.42
			StDev	0.15	0.11
			Min	0.31	0.16
			Max	0.89	0.66
			Range	0.58	0.50
		4	Points	81	81
			Mean	0.59	0.42
			StDev	0.15	0.11
			Min	0.29	0.16
			Max	0.89	0.66
			Range	0.60	0.50
		5	Points	82	82
			Mean	0.58	0.42
			StDev	0.16	0.11
			Min	0.18	0.16
			Max	0.89	0.66
			Range	0.71	0.50
		6	Points	82	82
			Mean	0.58	0.42
			StDev	0.14	0.11
			Min	0.24	0.16
			Max	0.89	0.66
			Range	0.65	0.50
		7	Points	82	82
			Mean	0.58	0.42
			StDev	0.16	0.11
			Min	0.22	0.16
			Max	0.92	0.66
			Range	0.70	0.50

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
5	Mathematics	8	Points	82	82
			Mean	0.58	0.42
			StDev	0.14	0.12
			Min	0.36	0.16
			Max	0.89	0.66
			Range	0.53	0.50
		9	Points	79	79
			Mean	0.59	0.42
			StDev	0.15	0.11
			Min	0.32	0.16
			Max	0.89	0.66
			Range	0.57	0.50
5	Reading	0	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
		1	Points	78	78
			Mean	0.65	0.42
			StDev	0.15	0.10
			Min	0.33	0.17
			Max	0.91	0.64
			Range	0.58	0.47
		2	Points	78	78
			Mean	0.66	0.42
			StDev	0.16	0.10
			Min	0.33	0.17
			Max	0.92	0.61
			Range	0.59	0.44
		3	Points	78	78
			Mean	0.67	0.41
			StDev	0.16	0.10
			Min	0.33	0.16
			Max	0.91	0.66
			Range	0.58	0.50
		4	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
		5	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
5	Reading	6	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
		7	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
		8	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
		9	Points	52	52
			Mean	0.66	0.41
			StDev	0.17	0.10
			Min	0.33	0.17
			Max	0.91	0.61
			Range	0.58	0.44
5	Writing	1	Points	37	37
			Mean	0.75	0.37
			StDev	0.19	0.14
			Min	0.45	0.19
			Max	0.99	0.62
			Range	0.54	0.43
6	Mathematics	0	Points	66	66
			Mean	0.54	0.44
			StDev	0.16	0.12
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53
		1	Points	80	80
			Mean	0.53	0.45
			StDev	0.16	0.13
			Min	0.21	0.20
			Max	0.86	0.74
			Range	0.65	0.54
		2	Points	83	83
			Mean	0.53	0.45
			StDev	0.15	0.12
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
		3	Points	80	80
			Mean	0.54	0.44
			StDev	0.15	0.11
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53
		4	Points	84	84
			Mean	0.54	0.44
			StDev	0.17	0.12
			Min	0.17	0.20
			Max	0.86	0.73
			Range	0.69	0.53
		5	Points	81	81
			Mean	0.55	0.44
			StDev	0.16	0.12
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53
6	Mathematics	6	Points	84	84
			Mean	0.53	0.46
			StDev	0.16	0.12
			Min	0.20	0.20
			Max	0.86	0.74
			Range	0.66	0.54
		7	Points	80	80
			Mean	0.53	0.45
			StDev	0.16	0.12
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53
		8	Points	83	83
			Mean	0.53	0.45
			StDev	0.15	0.12
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53
		9	Points	80	80
			Mean	0.54	0.44
			StDev	0.15	0.11
			Min	0.21	0.20
			Max	0.86	0.73
			Range	0.65	0.53

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
6	Reading	0	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
		1	Points	78	78
			Mean	0.73	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.95	0.67
			Range	0.54	0.46
		2	Points	78	78
			Mean	0.72	0.45
			StDev	0.15	0.11
			Min	0.41	0.21
			Max	0.93	0.67
			Range	0.52	0.46
		3	Points	78	78
			Mean	0.70	0.42
			StDev	0.16	0.13
			Min	0.41	0.10
			Max	0.92	0.67
			Range	0.51	0.57
		4	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
		5	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
		6	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
		7	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
6	Reading	8	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
		9	Points	52	52
			Mean	0.72	0.44
			StDev	0.15	0.10
			Min	0.41	0.21
			Max	0.92	0.67
			Range	0.51	0.46
7	Mathematics	0	Points	66	66
			Mean	0.53	0.44
			StDev	0.17	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		1	Points	82	82
			Mean	0.51	0.44
			StDev	0.16	0.12
			Min	0.21	0.20
			Max	0.89	0.69
			Range	0.68	0.49
		2	Points	81	81
			Mean	0.52	0.44
			StDev	0.16	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		3	Points	81	81
			Mean	0.52	0.44
			StDev	0.17	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		4	Points	82	82
			Mean	0.53	0.44
			StDev	0.17	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		5	Points	83	83
			Mean	0.50	0.44
			StDev	0.17	0.12
			Min	0.20	0.20
			Max	0.89	0.69
			Range	0.69	0.49

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
7	Mathematics	6	Points	83	83
			Mean	0.53	0.43
			StDev	0.17	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		7	Points	82	82
			Mean	0.51	0.44
			StDev	0.17	0.12
			Min	0.21	0.19
			Max	0.89	0.69
			Range	0.68	0.50
		8	Points	81	81
			Mean	0.53	0.44
			StDev	0.16	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
		9	Points	81	81
			Mean	0.52	0.44
			StDev	0.17	0.12
			Min	0.21	0.23
			Max	0.89	0.69
			Range	0.68	0.46
7	Reading	0	Points	52	52
			Mean	0.72	0.4
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		1	Points	78	78
			Mean	0.70	0.41
			StDev	0.14	0.11
			Min	0.45	0.20
			Max	0.92	0.67
			Range	0.47	0.47
		2	Points	78	78
			Mean	0.70	0.41
			StDev	0.15	0.12
			Min	0.48	0.15
			Max	0.92	0.68
			Range	0.44	0.53
		3	Points	78	78
			Mean	0.73	0.41
			StDev	0.14	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
7	Reading	4	Points	52	52
			Mean	0.72	0.4
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		5	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		6	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		7	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		8	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
		9	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.12
			Min	0.48	0.20
			Max	0.92	0.66
			Range	0.44	0.46
8	Mathematics	0	Points	66	66
			Mean	0.54	0.45
			StDev	0.17	0.11
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50
		1	Points	83	83
			Mean	0.54	0.46
			StDev	0.17	0.11
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
8	Mathematics	2	Points	82	82
			Mean	0.53	0.46
			StDev	0.17	0.11
			Min	0.13	0.19
			Max	0.85	0.73
			Range	0.72	0.54
		3	Points	83	83
			Mean	0.53	0.45
			StDev	0.18	0.11
			Min	0.11	0.23
			Max	0.85	0.73
			Range	0.74	0.50
		4	Points	81	81
			Mean	0.55	0.46
			StDev	0.17	0.11
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50
		5	Points	79	79
			Mean	0.53	0.45
			StDev	0.17	0.11
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50
		6	Points	84	84
			Mean	0.53	0.46
			StDev	0.17	0.12
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50
		7	Points	83	83
			Mean	0.55	0.46
			StDev	0.17	0.11
			Min	0.13	0.23
			Max	0.85	0.73
			Range	0.72	0.50
		8	Points	82	82
			Mean	0.53	0.46
			StDev	0.17	0.11
			Min	0.13	0.20
			Max	0.85	0.73
			Range	0.72	0.53
		9	Points	83	83
			Mean	0.53	0.45
			StDev	0.18	0.11
			Min	0.11	0.22
			Max	0.85	0.73
			Range	0.74	0.51

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
8	Reading	0	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50
		1	Points	78	78
			Mean	0.72	0.42
			StDev	0.15	0.13
			Min	0.34	0.17
			Max	0.92	0.69
			Range	0.58	0.52
		21	Points	78	78
			Mean	0.71	0.41
			StDev	0.14	0.13
			Min	0.41	0.17
			Max	0.92	0.70
			Range	0.51	0.53
		3	Points	78	78
			Mean	0.71	0.41
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.70
			Range	0.51	0.53
		4	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50
		5	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50
		6	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50
		7	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
8	Reading	8	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.5
		9	Points	52	52
			Mean	0.72	0.40
			StDev	0.15	0.13
			Min	0.41	0.17
			Max	0.92	0.67
			Range	0.51	0.50
8	Writing	1	Points	37	37
			Mean	0.75	0.40
			StDev	0.17	0.15
			Min	0.49	0.23
			Max	0.96	0.70
			Range	0.47	0.47
11	Mathematics	0	Points	64	64
			Mean	0.42	0.46
			StDev	0.19	0.12
			Min	0.08	0.18
			Max	0.83	0.70
			Range	0.75	0.52
		1	Points	76	76
			Mean	0.41	0.47
			StDev	0.19	0.12
			Min	0.08	0.18
			Max	0.83	0.70
			Range	0.75	0.52
		2	Points	76	76
			Mean	0.41	0.46
			StDev	0.19	0.12
			Min	0.08	0.16
			Max	0.83	0.70
			Range	0.75	0.54
		3	Points	76	76
			Mean	0.41	0.46
			StDev	0.19	0.12
			Min	0.08	0.18
			Max	0.83	0.70
			Range	0.75	0.52
		4	Points	76	76
			Mean	0.40	0.46
			StDev	0.19	0.12
			Min	0.07	0.18
			Max	0.83	0.70
			Range	0.76	0.52

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
11	Mathematics	5	Points	76	76
			Mean	0.41	0.46
			StDev	0.18	0.13
			Min	0.08	0.05
			Max	0.83	0.70
			Range	0.75	0.65
		6	Points	76	76
			Mean	0.41	0.46
			StDev	0.19	0.12
			Min	0.08	0.17
			Max	0.83	0.70
			Range	0.75	0.53
		7	Points	76	76
			Mean	0.41	0.47
			StDev	0.19	0.12
			Min	0.08	0.18
			Max	0.83	0.70
			Range	0.75	0.52
		8	Points	76	76
			Mean	0.41	0.46
			StDev	0.19	0.12
			Min	0.08	0.13
			Max	0.83	0.70
			Range	0.75	0.57
11	Reading	0	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
		1	Points	78	78
			Mean	0.70	0.43
			StDev	0.13	0.13
			Min	0.44	0.20
			Max	0.92	0.72
			Range	0.48	0.52
		2	Points	78	78
			Mean	0.70	0.43
			StDev	0.13	0.13
			Min	0.42	0.20
			Max	0.92	0.72
			Range	0.50	0.52
		3	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
11	Reading	4	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
		5	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
		6	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
		7	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
		8	Points	52	52
			Mean	0.71	0.41
			StDev	0.12	0.13
			Min	0.47	0.20
			Max	0.92	0.68
			Range	0.45	0.48
11	Writing	0	Points	12	12
			Mean	0.50	
			StDev		
			Min	0.50	
			Max	0.50	
			Range	0.00	
		1	Points	18	18
			Mean	0.47	0.63
			StDev	0.04	
			Min	0.44	0.63
			Max	0.50	0.63
			Range	0.06	0.00
		2	Points	18	18
			Mean	0.46	0.61
			StDev	0.06	
			Min	0.41	0.61
			Max	0.50	0.61
			Range	0.09	0.00

(cont'd)

Grade	Content	Form	Statistic	Difficulty	Discrimination
11	Writing	3	Points	18	18
			Mean	0.48	0.61
			StDev	0.03	
			Min	0.46	0.61
			Max	0.50	0.61
			Range	0.04	0.00
		4	Points	12	12
			Mean	0.50	
			StDev		
			Min	0.50	
			Max	0.50	
			Range	0.00	
		5	Points	12	12
			Mean	0.50	
			StDev		
			Min	0.50	
			Max	0.50	
			Range	0.00	
		6	Points	18	18
			Mean	0.50	0.58
			StDev	0.01	
			Min	0.49	0.58
			Max	0.50	0.58
			Range	0.01	0.00
		7	Points	18	18
			Mean	0.49	0.65
			StDev	0.01	
			Min	0.48	0.65
			Max	0.50	0.65
			Range	0.02	0.00
		8	Points	12	12
			Mean	0.50	
			StDev		
			Min	0.50	
			Max	0.50	
			Range	0.00	

Appendix H—Number of Items Classified into DIF Categories by Subgroup Test Form and Item Type

Table H-1. 2008–09 NECAP: Number of Items Classified into DIF Categories within Subgroups by Grade, Content, Test Form, and Item Type—Male versus Female

Grade	*Content*	*Form*	*All A*	*All B*	*All C*	*All D*	*MC A*	*MC B*	*MC C*	*MC D*	*OR A*	*OR B*	*OR C*	*OR D*
3	Mathematics	00	52	3	0	0	33	2	0	0	19	1	0	0
		01	9	1	0	0	6	0	0	0	3	1	0	0
		02	10	0	0	0	6	0	0	0	4	0	0	0
		03	10	0	0	0	6	0	0	0	4	0	0	0
		04	10	0	0	0	6	0	0	0	4	0	0	0
		05	10	0	0	0	6	0	0	0	4	0	0	0
		06	10	0	0	0	6	0	0	0	4	0	0	0
		07	9	1	0	0	5	1	0	0	4	0	0	0
		08	10	0	0	0	6	0	0	0	4	0	0	0
		09	10	0	0	0	6	0	0	0	4	0	0	0
	Reading	00	34	0	0	0	28	0	0	0	6	0	0	0
		01	17	0	0	0	14	0	0	0	3	0	0	0
		02	17	0	0	0	14	0	0	0	3	0	0	0
		03	17	0	0	0	14	0	0	0	3	0	0	0
4	Mathematics	00	49	6	0	0	31	4	0	0	18	2	0	0
		01	10	0	0	0	6	0	0	0	4	0	0	0
		02	10	0	0	0	6	0	0	0	4	0	0	0
		03	7	3	0	0	4	2	0	0	3	1	0	0
		04	8	1	1	0	4	1	1	0	4	0	0	0
		05	10	0	0	0	6	0	0	0	4	0	0	0
		06	9	1	0	0	5	1	0	0	4	0	0	0
		07	10	0	0	0	6	0	0	0	4	0	0	0
		08	10	0	0	0	6	0	0	0	4	0	0	0
		09	9	1	0	0	6	0	0	0	3	1	0	0
	Reading	00	34	0	0	0	28	0	0	0	6	0	0	0
		01	16	1	0	0	13	1	0	0	3	0	0	0
		02	17	0	0	0	14	0	0	0	3	0	0	0
		03	16	1	0	0	13	1	0	0	3	0	0	0
5	Mathematics	00	45	3	0	0	29	3	0	0	16	0	0	0
		01	9	2	0	0	4	2	0	0	5	0	0	0
		02	10	1	0	0	5	1	0	0	5	0	0	0
		03	10	0	0	0	6	0	0	0	4	0	0	0
		04	9	2	0	0	5	1	0	0	4	1	0	0
		05	11	0	0	0	6	0	0	0	5	0	0	0
		06	10	1	0	0	6	0	0	0	4	1	0	0
		07	9	2	0	0	4	2	0	0	5	0	0	0
		08	10	1	0	0	5	1	0	0	5	0	0	0
		09	10	0	0	0	6	0	0	0	4	0	0	0
	Reading	00	33	1	0	0	27	1	0	0	6	0	0	0
		01	13	4	0	0	10	4	0	0	3	0	0	0
		02	14	3	0	0	11	3	0	0	3	0	0	0
		03	16	1	0	0	13	1	0	0	3	0	0	0
	Writing	01	17	0	0	0	10	0	0	0	7	0	0	0
6	Mathematics	*00*	37	11	0	0	22	10	0	0	15	1	0	0
		01	8	2	0	0	4	2	0	0	4	0	0	0
		02	10	1	0	0	5	1	0	0	5	0	0	0
		03	5	4	1	0	3	2	1	0	2	2	0	0
		04	12	0	0	0	6	0	0	0	6	0	0	0
		05	11	0	0	0	6	0	0	0	5	0	0	0

(cont'd)

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
6	Mathematics	06	11	1	0	0	5	1	0	0	6	0	0	0
		07	8	2	0	0	4	2	0	0	4	0	0	0
		08	10	1	0	0	5	1	0	0	5	0	0	0
		09	5	5	0	0	2	4	0	0	3	1	0	0
	Reading	00	32	2	0	0	26	2	0	0	6	0	0	0
		01	11	6	0	0	8	6	0	0	3	0	0	0
		02	14	1	2	0	11	1	2	0	3	0	0	0
		03	13	2	2	0	11	1	2	0	2	1	0	0
7	Mathematics	00	42	5	1	0	29	2	1	0	13	3	0	0
		01	7	4	0	0	2	4	0	0	5	0	0	0
		02	9	1	0	0	5	1	0	0	4	0	0	0
		03	10	0	0	0	6	0	0	0	4	0	0	0
		04	6	5	0	0	3	3	0	0	3	2	0	0
		05	10	2	0	0	5	1	0	0	5	1	0	0
		06	8	2	2	0	4	0	2	0	4	2	0	0
		07	6	5	0	0	2	4	0	0	4	1	0	0
		08	9	1	0	0	5	1	0	0	4	0	0	0
		09	9	1	0	0	5	1	0	0	4	0	0	0
	Reading	00	27	6	1	0	23	4	1	0	4	2	0	0
		01	14	3	0	0	12	2	0	0	2	1	0	0
		02	13	3	1	0	12	1	1	0	1	2	0	0
		03	15	1	1	0	12	1	1	0	3	0	0	0
8	Mathematics	00	42	5	1	0	29	2	1	0	13	3	0	0
		01	9	2	0	0	4	2	0	0	5	0	0	0
		02	10	1	0	0	5	1	0	0	5	0	0	0
		03	10	2	0	0	6	0	0	0	4	2	0	0
		04	8	3	0	0	5	1	0	0	3	2	0	0
		05	9	0	0	0	6	0	0	0	3	0	0	0
		06	12	0	0	0	6	0	0	0	6	0	0	0
		07	6	5	0	0	4	2	0	0	2	3	0	0
		08	11	0	0	0	6	0	0	0	5	0	0	0
		09	9	3	0	0	5	1	0	0	4	2	0	0
	Reading	00	27	6	1	0	24	3	1	0	3	3	0	0
		01	15	2	0	0	13	1	0	0	2	1	0	0
		02	14	2	1	0	12	1	1	0	2	1	0	0
		03	15	2	0	0	12	2	0	0	3	0	0	0
	Writing	01	15	2	0	0	9	1	0	0	6	1	0	0
11	Mathematics	00	41	5	0	0	21	3	0	0	20	2	0	0
		01	7	1	0	0	4	0	0	0	3	1	0	0
		02	6	2	0	0	2	2	0	0	4	0	0	0
		03	7	1	0	0	4	0	0	0	3	1	0	0
		04	7	1	0	0	3	1	0	0	4	0	0	0
		05	8	0	0	0	4	0	0	0	4	0	0	0
		06	8	0	0	0	4	0	0	0	4	0	0	0
		07	8	0	0	0	4	0	0	0	4	0	0	0
		08	6	2	0	0	2	2	0	0	4	0	0	0
		09	41	5	0	0	21	3	0	0	20	2	0	0
	Reading	00	22	9	3	0	18	7	3	0	4	2	0	0
		01	15	2	0	0	12	2	0	0	3	0	0	0
		02	11	5	1	0	8	5	1	0	3	0	0	0

All = MC and OR items; MC = multiple-choice items; OR = open-response items;
A = “negligible” DIF; B = “low” DIF; C = “high” DIF; D = not enough students to perform reliable DIF analysis

Table H-2. 2008–09 NECAP: Number of Items Classified into DIF Categories within Subgroups by Grade, Content, Test Form, and Item Type—White versus Black

Grade	*Content*	*Form*	*All A*	*All B*	*All C*	*All D*	*MC A*	*MC B*	*MC C*	*MC D*	*OR A*	*OR B*	*OR C*	*OR D*
3	Mathematics	00	53	2	0	0	34	1	0	0	19	1	0	0
		01	0	0	0	10	0	0	0	6	0	0	0	4
		02	0	0	0	10	0	0	0	6	0	0	0	4
		03	0	0	0	10	0	0	0	6	0	0	0	4
		04	0	0	0	10	0	0	0	6	0	0	0	4
		05	0	0	0	10	0	0	0	6	0	0	0	4
		06	0	0	0	10	0	0	0	6	0	0	0	4
		07	0	0	0	10	0	0	0	6	0	0	0	4
		08	0	0	0	10	0	0	0	6	0	0	0	4
		09	0	0	0	10	0	0	0	6	0	0	0	4
	Reading	00	31	2	1	0	26	1	1	0	5	1	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
4	Mathematics	00	45	10	0	0	29	6	0	0	16	4	0	0
		01	0	0	0	10	0	0	0	6	0	0	0	4
		02	0	0	0	10	0	0	0	6	0	0	0	4
		03	0	0	0	10	0	0	0	6	0	0	0	4
		04	0	0	0	10	0	0	0	6	0	0	0	4
		05	0	0	0	10	0	0	0	6	0	0	0	4
		06	0	0	0	10	0	0	0	6	0	0	0	4
		07	0	0	0	10	0	0	0	6	0	0	0	4
		08	0	0	0	10	0	0	0	6	0	0	0	4
		09	0	0	0	10	0	0	0	6	0	0	0	4
	Reading	00	30	3	1	0	24	3	1	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
5	Mathematics	00	44	4	0	0	31	1	0	0	13	3	0	0
		01	0	0	0	11	0	0	0	6	0	0	0	5
		02	0	0	0	11	0	0	0	6	0	0	0	5
		03	0	0	0	10	0	0	0	6	0	0	0	4
		04	0	0	0	11	0	0	0	6	0	0	0	5
		05	0	0	0	11	0	0	0	6	0	0	0	5
		06	0	0	0	11	0	0	0	6	0	0	0	5
		07	0	0	0	11	0	0	0	6	0	0	0	5
		08	0	0	0	11	0	0	0	6	0	0	0	5
		09	0	0	0	10	0	0	0	6	0	0	0	4
	Reading	00	29	4	1	0	23	4	1	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
	Writing	01	14	3	0	0	7	3	0	0	7	0	0	0
6	Mathematics	00	42	5	1	0	29	3	0	0	13	2	1	0
		01	0	0	0	10	0	0	0	6	0	0	0	4
		02	0	0	0	11	0	0	0	6	0	0	0	5
		03	0	0	0	10	0	0	0	6	0	0	0	4
		04	0	0	0	12	0	0	0	6	0	0	0	6

(cont'd)

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
6	Mathematics	05	0	0	0	11	0	0	0	6	0	0	0	5
		06	0	0	0	12	0	0	0	6	0	0	0	6
		07	0	0	0	10	0	0	0	6	0	0	0	4
		08	0	0	0	11	0	0	0	6	0	0	0	5
		09	0	0	0	10	0	0	0	6	0	0	0	4
	Reading	00	28	5	1	0	22	5	1	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
7	Mathematics	00	39	9	0	0	25	7	0	0	14	2	0	0
		01	0	0	0	11	0	0	0	6	0	0	0	5
		02	0	0	0	10	0	0	0	6	0	0	0	4
		03	0	0	0	10	0	0	0	6	0	0	0	4
		04	0	0	0	11	0	0	0	6	0	0	0	5
		05	0	0	0	12	0	0	0	6	0	0	0	6
		06	0	0	0	12	0	0	0	6	0	0	0	6
		07	0	0	0	11	0	0	0	6	0	0	0	5
		08	0	0	0	10	0	0	0	6	0	0	0	4
		09	0	0	0	10	0	0	0	6	0	0	0	4
	Reading	00	27	7	0	0	21	7	0	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
8	Mathematics	00	45	3	0	0	29	3	0	0	16	0	0	0
		01	0	0	0	11	0	0	0	6	0	0	0	5
		02	0	0	0	11	0	0	0	6	0	0	0	5
		03	0	0	0	12	0	0	0	6	0	0	0	6
		04	0	0	0	11	0	0	0	6	0	0	0	5
		05	0	0	0	9	0	0	0	6	0	0	0	3
		06	0	0	0	12	0	0	0	6	0	0	0	6
		07	0	0	0	11	0	0	0	6	0	0	0	5
		08	0	0	0	11	0	0	0	6	0	0	0	5
		09	0	0	0	11	0	0	0	6	0	0	0	5
	Reading	00	27	5	2	0	21	5	2	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
		03	0	0	0	17	0	0	0	14	0	0	0	3
	Writing	01	13	4	0	0	6	4	0	0	7	0	0	0
11	Mathematics	00	41	5	0	0	19	5	0	0	22	0	0	0
		01	0	0	0	8	0	0	0	4	0	0	0	4
		02	0	0	0	8	0	0	0	4	0	0	0	4
		03	0	0	0	8	0	0	0	4	0	0	0	4
		04	0	0	0	8	0	0	0	4	0	0	0	4
		05	0	0	0	8	0	0	0	4	0	0	0	4
		06	0	0	0	8	0	0	0	4	0	0	0	4
		07	0	0	0	8	0	0	0	4	0	0	0	4
		08	0	0	0	8	0	0	0	4	0	0	0	4
		09	41	5	0	0	19	5	0	0	22	0	0	0

(cont'd)

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
11	Reading	00	24	9	1	0	18	9	1	0	6	0	0	0
		01	0	0	0	17	0	0	0	14	0	0	0	3
		02	0	0	0	17	0	0	0	14	0	0	0	3
	Writing	00	0	1	0	0					0	1	0	0
		01	0	0	0	1					0	0	0	1
		02	0	0	0	1					0	0	0	1
		03	0	0	0	1					0	0	0	1
		06	0	0	0	1					0	0	0	1
		07	0	0	0	1					0	0	0	1

All = MC andOR items; MC = multiple-choice items; OR = open-response items;
A = “negligible” DIF; B = “low” DIF; C = “high” DIF; D = not enough students to perform reliable DIF analysis

Table H-3. 2008–09 NECAP: Number of Items Classified into DIF Categories within Subgroups by Grade, Content, Test Form, and Item Type —White versus Hispanic

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
3	Mathematics	00	47	8	0	0	30	5	0	0	17	3	0	0
		01	9	1	0	0	5	1	0	0	4	0	0	0
		02	6	4	0	0	5	1	0	0	1	3	0	0
		03	8	2	0	0	5	1	0	0	3	1	0	0
		04	9	1	0	0	6	0	0	0	3	1	0	0
		05	8	2	0	0	6	0	0	0	2	2	0	0
		06	7	3	0	0	3	3	0	0	4	0	0	0
		07	8	1	1	0	5	1	0	0	3	0	1	0
		08	7	3	0	0	5	1	0	0	2	2	0	0
		09	8	2	0	0	4	2	0	0	4	0	0	0
	Reading	00	27	7	0	0	22	6	0	0	5	1	0	0
		01	16	0	1	0	13	0	1	0	3	0	0	0
		02	14	3	0	0	11	3	0	0	3	0	0	0
		03	16	1	0	0	13	1	0	0	3	0	0	0
4	Mathematics	00	49	6	0	0	31	4	0	0	18	2	0	0
		01	7	3	0	0	4	2	0	0	3	1	0	0
		02	7	3	0	0	4	2	0	0	3	1	0	0
		03	9	0	1	0	5	0	1	0	4	0	0	0
		04	9	1	0	0	5	1	0	0	4	0	0	0
		05	7	2	1	0	4	1	1	0	3	1	0	0
		06	8	2	0	0	5	1	0	0	3	1	0	0
		07	10	0	0	0	6	0	0	0	4	0	0	0
		08	8	1	1	0	5	0	1	0	3	1	0	0
		09	8	1	1	0	4	1	1	0	4	0	0	0
	Reading	00	24	9	1	0	19	8	1	0	5	1	0	0
		01	15	2	0	0	12	2	0	0	3	0	0	0
		02	16	0	1	0	13	0	1	0	3	0	0	0
		03	14	2	1	0	11	2	1	0	3	0	0	0
5	Mathematics	00	43	5	0	0	30	2	0	0	13	3	0	0
		01	9	2	0	0	4	2	0	0	5	0	0	0
		02	9	2	0	0	5	1	0	0	4	1	0	0
		03	8	2	0	0	5	1	0	0	3	1	0	0
		04	10	1	0	0	6	0	0	0	4	1	0	0
		05	10	1	0	0	6	0	0	0	4	1	0	0
		06	10	1	0	0	5	1	0	0	5	0	0	0
		07	10	1	0	0	6	0	0	0	4	1	0	0

(cont’d)

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
5	Reading	08	9	2	0	0	5	1	0	0	4	1	0	0
		09	9	1	0	0	5	1	0	0	4	0	0	0
		00	26	7	1	0	20	7	1	0	6	0	0	0
		01	11	5	1	0	8	5	1	0	3	0	0	0
		02	12	4	1	0	10	3	1	0	2	1	0	0
		03	12	5	0	0	9	5	0	0	3	0	0	0
	Writing	01	13	3	1	0	6	3	1	0	7	0	0	0
6	Mathematics	00	41	6	1	0	27	5	0	0	14	1	1	0
		01	9	1	0	0	5	1	0	0	4	0	0	0
		02	5	5	1	0	3	3	0	0	2	2	1	0
		03	5	4	1	0	3	2	1	0	2	2	0	0
		04	10	2	0	0	5	1	0	0	5	1	0	0
		05	6	4	1	0	3	3	0	0	3	1	1	0
		06	9	3	0	0	3	3	0	0	6	0	0	0
		07	8	2	0	0	4	2	0	0	4	0	0	0
		08	8	2	1	0	4	1	1	0	4	1	0	0
		09	5	5	0	0	3	3	0	0	2	2	0	0
	Reading	00	23	8	3	0	17	8	3	0	6	0	0	0
		01	12	2	3	0	10	1	3	0	2	1	0	0
		02	13	2	2	0	10	2	2	0	3	0	0	0
		03	7	6	4	0	5	5	4	0	2	1	0	0
7	Mathematics	00	41	4	3	0	27	3	2	0	14	1	1	0
		01	10	1	0	0	6	0	0	0	4	1	0	0
		02	9	1	0	0	6	0	0	0	3	1	0	0
		03	9	1	0	0	6	0	0	0	3	1	0	0
		04	8	2	1	0	4	2	0	0	4	0	1	0
		05	10	1	1	0	5	1	0	0	5	0	1	0
		06	10	1	1	0	5	0	1	0	5	1	0	0
		07	11	0	0	0	6	0	0	0	5	0	0	0
		08	8	2	0	0	6	0	0	0	2	2	0	0
		09	9	1	0	0	5	1	0	0	4	0	0	0
	Reading	00	24	10	0	0	18	10	0	0	6	0	0	0
		01	11	5	1	0	8	5	1	0	3	0	0	0
		02	12	4	1	0	9	4	1	0	3	0	0	0
		03	14	2	1	0	12	1	1	0	2	1	0	0
8	Mathematics	00	45	3	0	0	29	3	0	0	16	0	0	0
		01	11	0	0	0	6	0	0	0	5	0	0	0
		02	11	0	0	0	6	0	0	0	5	0	0	0
		03	11	0	1	0	5	0	1	0	6	0	0	0
		04	7	4	0	0	4	2	0	0	3	2	0	0
		05	7	2	0	0	4	2	0	0	3	0	0	0
		06	11	1	0	0	5	1	0	0	6	0	0	0
		07	7	3	1	0	3	2	1	0	4	1	0	0
		08	9	2	0	0	5	1	0	0	4	1	0	0
		09	9	2	1	0	5	1	0	0	4	1	1	0
	Reading	00	25	7	2	0	19	7	2	0	6	0	0	0
		01	12	3	2	0	9	3	2	0	3	0	0	0
		02	14	2	1	0	11	2	1	0	3	0	0	0
		03	9	3	5	0	7	2	5	0	2	1	0	0
	Writing	01	12	5	0	0	5	5	0	0	7	0	0	0

(cont'd)

Grade	Content	Form	All A	All B	All C	All D	MC A	MC B	MC C	MC D	OR A	OR B	OR C	OR D
11	Mathematics	00	43	3	0	0	21	3	0	0	22	0	0	0
		01	6	2	0	0	3	1	0	0	3	1	0	0
		02	6	2	0	0	3	1	0	0	3	1	0	0
		03	6	2	0	0	2	2	0	0	4	0	0	0
		04	5	2	1	0	1	2	1	0	4	0	0	0
		05	7	1	0	0	3	1	0	0	4	0	0	0
		06	7	1	0	0	3	1	0	0	4	0	0	0
		07	5	3	0	0	2	2	0	0	3	1	0	0
		08	4	3	1	0	2	2	0	0	2	1	1	0
	Reading	00	17	10	7	0	13	8	7	0	4	2	0	0
		01	10	6	1	0	8	5	1	0	2	1	0	0
		02	11	5	1	0	9	4	1	0	2	1	0	0
	Writing	00	0	1	0	0					0	1	0	0
		01	1	0	0	0					1	0	0	0
		02	1	0	0	0					1	0	0	0
		03	1	0	0	0					1	0	0	0
		06	1	0	0	0					1	0	0	0
		07	1	0	0	0					1	0	0	0

All = MC and OR items; MC = multiple-choice items; OR = open-response items;
A = “negligible” DIF; B = “low” DIF; C = “high” DIF; D = not enough students to perform reliable DIF analysis

Appendix I—Common Item DIF Category Counts in the Male vs. Female Comparison

Table I-1. 2008–09 NECAP: Common Item DIF Category Counts—
Male vs. Female Comparison by Direction of Advantage, Item Type, Grade, and Content

Grade	*Content*	*Item Type*	*F_A*	*M_A*	*N_A*	*P_A*	*F_B*	*M_B*	*N_B*	*P_B*	*F_C*	*M_C*	*N_C*	*P_C*
3	Mathematics	MC	53	33	86	0.97	0	3	3	0.03	0	0	0	0.00
		SA	23	31	54	0.96	1	1	2	0.04	0	0	0	0.00
	Reading	MC	47	23	70	1.00	0	0	0	0.00	0	0	0	0.00
		CR	12	3	15	1.00	0	0	0	0.00	0	0	0	0.00
4	Mathematics	MC	46	34	80	0.90	2	6	8	0.09	0	1	1	0.01
		SA	39	13	52	0.93	3	1	4	0.07	0	0	0	0.00
	Reading	MC	37	31	68	0.97	0	2	2	0.03	0	0	0	0.00
		CR	11	4	15	1.00	0	0	0	0.00	0	0	0	0.00
5	Mathematics	MC	34	42	76	0.88	2	8	10	0.12	0	0	0	0.00
		OR	34	23	57	0.97	2	0	2	0.03	0	0	0	0.00
	Reading	MC	27	34	61	0.87	0	9	9	0.13	0	0	0	0.00
		CR	14	1	15	1.00	0	0	0	0.00	0	0	0	0.00
	Writing	MC	4	6	10	1.00	0	0	0	0.00	0	0	0	0.00
		OR	7	0	7	1.00	0	0	0	0.00	0	0	0	0.00
6	Mathematics	MC	29	33	62	0.72	8	15	23	0.27	1	0	1	0.01
		OR	35	20	55	0.93	3	1	4	0.07	0	0	0	0.00
	Reading	MC	16	40	56	0.80	0	10	10	0.14	0	4	4	0.06
		CR	14	0	14	0.93	1	0	1	0.07	0	0	0	0.00
7	Mathematics	MC	37	29	66	0.77	3	14	17	0.20	0	3	3	0.03
		OR	31	19	50	0.85	6	3	9	0.15	0	0	0	0.00
	Reading	MC	26	33	59	0.84	0	8	8	0.11	0	3	3	0.04
		CR	10	0	10	0.67	5	0	5	0.33	0	0	0	0.00
8	Mathematics	MC	37	39	76	0.88	0	9	9	0.10	0	1	1	0.01
		OR	36	14	50	0.81	6	6	12	0.19	0	0	0	0.00
	Reading	MC	23	38	61	0.87	0	7	7	0.10	0	2	2	0.03
		CR	10	0	10	0.67	5	0	5	0.33	0	0	0	0.00
	Writing	MC	3	6	9	0.90	0	1	1	0.10	0	0	0	0.00
		OR	6	0	6	0.86	1	0	1	0.14	0	0	0	0.00
11	Mathematics	MC	26	25	51	0.91	1	4	5	0.09	0	0	0	0.00
		OR	29	15	44	0.81	6	4	10	0.19	0	0	0	0.00
	Reading	MC	15	24	39	0.70	2	12	14	0.25	0	3	3	0.05
		CR	7	0	7	0.58	5	0	5	0.42	0	0	0	0.00
	Writing	OR	5	0	5	0.83	1	0	1	0.17	0	0	0	0.00

F_ = items on which females performed better than males (controlling for total test score); M_ = items on which males performed better than females, (controlling for total test score); N_ = number of items; P_ = proportion of items_ A = “negligible” DIF; _B = “low” DIF; _C = “high” DIF

Appendix J—Item Response Theory Parameters, TCCs, and TIFs

Table J-1. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Mathematics Grade 3

Item Number	*Parameters*		
	a	b	c
119847	0.570	-1.673	0.072
119678	1.266	-0.450	0.228
119839	0.866	-0.307	0.000
119752	0.724	-1.619	0.081
119725	1.066	0.325	0.146
119718	1.148	-0.775	0.115
119846	0.642	0.325	0.091
119772	0.570	-1.681	0.114
119723	0.934	0.287	0.079
119844	0.577	-2.545	0.000
119877	0.483	-2.524	0.000
119871	1.004	-1.072	0.031
119735	0.618	0.325	0.213
119866	1.059	1.104	0.199
119848	0.674	-2.112	0.000
119770	1.018	-0.622	0.097
119768	0.897	-0.468	0.269
119869	0.927	0.140	0.190
119840	1.007	-0.132	0.235
119745	0.833	0.028	0.183
119935	0.777	1.009	0.153
119681	0.546	-2.167	0.000
119858	0.401	-2.480	0.239
119743	0.818	-1.207	0.359
119685	0.906	-1.184	0.040
119702	1.252	0.615	0.201
119842	1.198	0.157	0.402
119694	0.857	-0.586	0.310
119764	1.002	-0.280	0.131
119769	1.125	0.024	0.182
119776	0.728	0.877	0.266
119741	0.622	-0.907	0.299
119778	1.182	0.559	0.103
119758	1.108	-0.867	0.049
119696	0.953	-1.567	0.061

a = discrimination; b = difficulty; c = guessing

Table J-2. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 3

Item Number	*Parameters*			
	a	b	D1	D2
119760	0.647	-0.640	N/A	N/A
119751	0.786	0.350	N/A	N/A
119775	0.653	-0.572	N/A	N/A
119852	0.875	0.841	N/A	N/A
119761	0.803	0.744	0.670	-0.670
119708	0.751	-1.086	1.027	-1.027
119782	0.629	0.967	1.536	-1.536
119864	0.701	-0.605	0.776	-0.776
119853	1.011	-1.265	0.593	-0.593
119731	0.902	-1.115	N/A	N/A
119699	0.700	-2.301	N/A	N/A
119850	1.010	0.480	N/A	N/A
119875	0.767	-0.662	0.956	-0.956
119862	0.882	0.018	0.790	-0.790
119806	0.878	-1.001	0.813	-0.813
119733	0.660	-0.879	N/A	N/A
119860	0.896	-0.393	N/A	N/A
119873	0.836	0.283	N/A	N/A
119836	0.640	-2.284	1.027	-1.027
119780	0.846	0.264	0.920	-0.920

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter

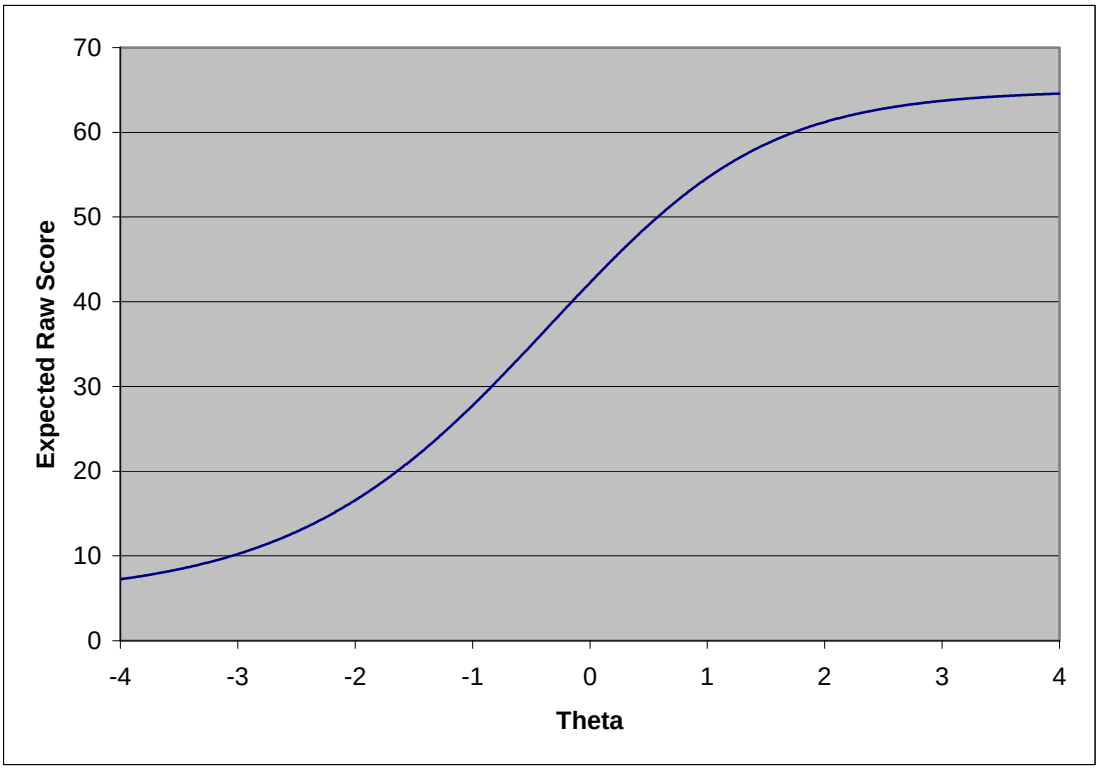

Figure J-1 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 3

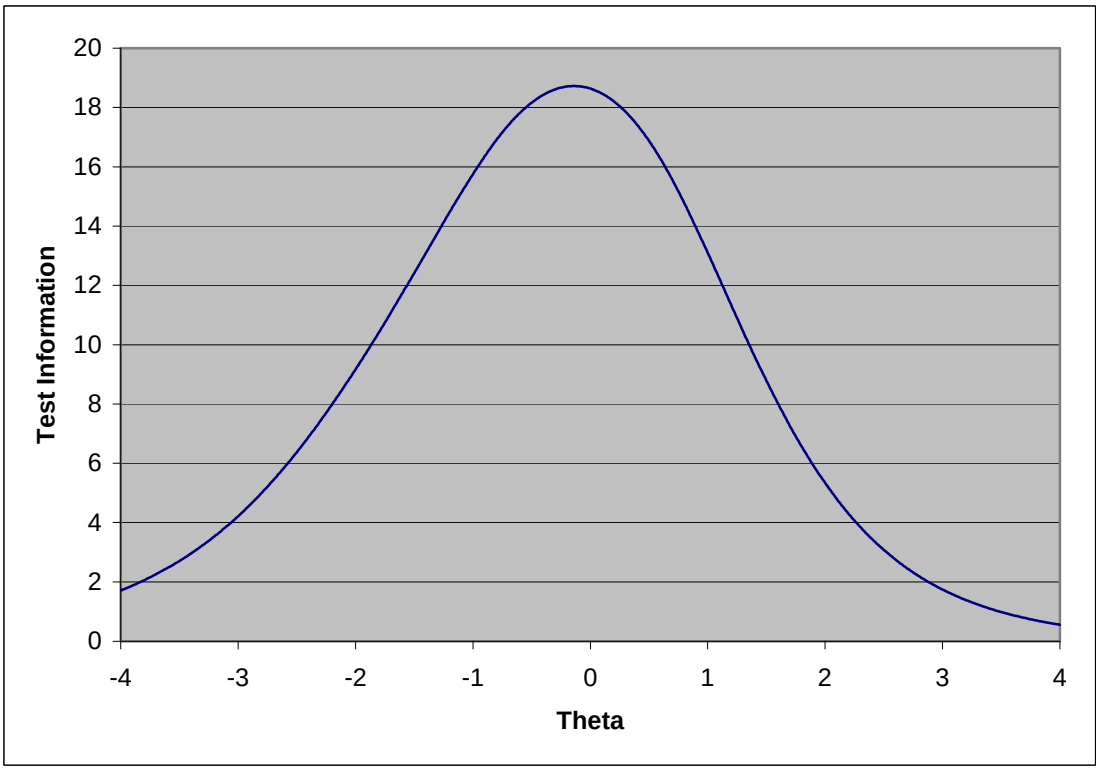

Figure J-2 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 3

Table J-3. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Mathematics Grade 4

Item Number	*Parameters*		
	a	b	c
120220	1.069	-1.277	0.065
120047	0.760	-0.855	0.107
120217	0.712	-0.573	0.102
120218	0.593	-0.062	0.153
120112	0.742	1.786	0.070
120219	0.908	0.423	0.144
120062	0.831	-0.283	0.081
120211	1.023	0.274	0.140
120070	1.252	0.413	0.194
120053	0.988	-1.566	0.179
120224	0.450	-1.890	0.230
120232	0.993	-1.332	0.090
120214	0.791	-1.556	0.047
120234	1.081	0.370	0.097
120158	0.982	-0.677	0.278
120183	0.746	-1.987	0.064
120243	0.809	-2.153	0.000
120087	0.480	-1.823	0.093
120215	1.450	1.384	0.108
120245	0.724	-1.135	0.042
120209	0.723	-1.036	0.186
120135	0.657	-0.884	0.146
120212	0.900	-1.090	0.395
120100	0.793	-1.376	0.104
120085	0.749	-1.056	0.178
120139	1.014	-0.530	0.081
120126	1.051	-1.390	0.168
120116	1.108	-0.461	0.187
120236	1.074	-0.101	0.401
120110	0.681	0.148	0.115
120067	0.639	-1.539	0.045
120146	0.378	0.906	0.068
120144	0.949	-1.250	0.078
120090	0.753	-0.696	0.072
120124	0.732	-1.556	0.000

a = discrimination; b = difficulty; c = guessing

Table J-4. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 4

Item Number	*Parameters*			
	a	b	D1	D2
120051	0.518	-1.940	N/A	N/A
120074	0.683	-0.712	N/A	N/A
120197	0.661	-0.497	N/A	N/A
120050	0.934	0.407	N/A	N/A
120061	0.735	-1.509	0.521	-0.521
120250	0.780	-0.415	1.367	-1.367
120083	0.695	-0.843	0.798	-0.798
120107	0.716	0.811	0.372	-0.372
120127	0.470	-2.458	1.541	-1.541
120237	0.671	-0.416	N/A	N/A
120159	0.792	-0.739	N/A	N/A
120072	0.538	-2.090	N/A	N/A
120222	0.868	0.017	1.515	-1.515
120205	1.076	-0.251	0.284	-0.284
120121	0.989	-0.191	0.613	-0.613
120228	0.541	-1.189	N/A	N/A
120129	0.463	-0.601	N/A	N/A
120071	0.855	-0.127	N/A	N/A
120099	0.801	-0.492	0.564	-0.564
120079	0.720	-0.689	0.729	-0.729

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter

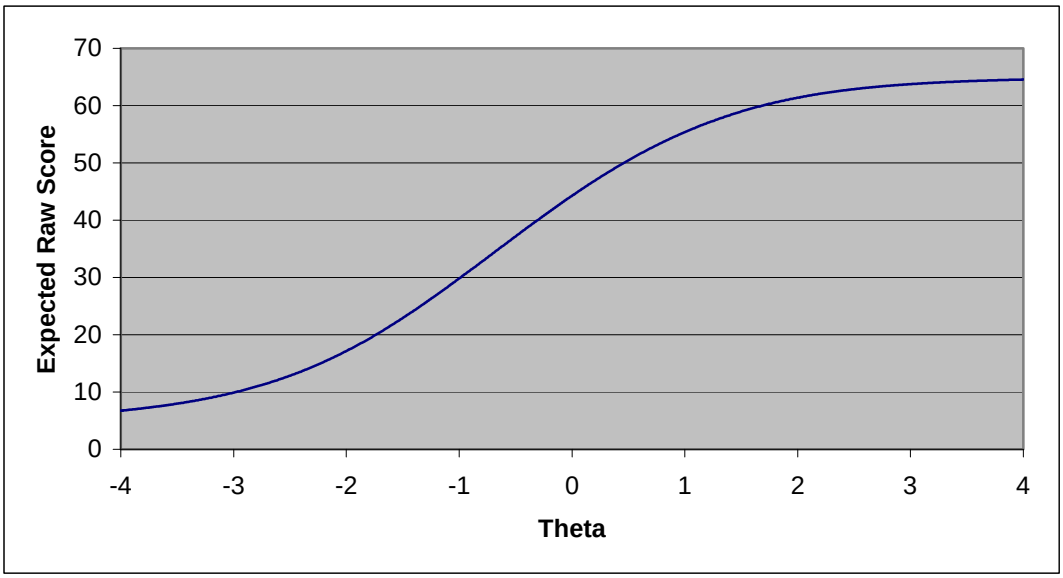

Figure J-3 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 4

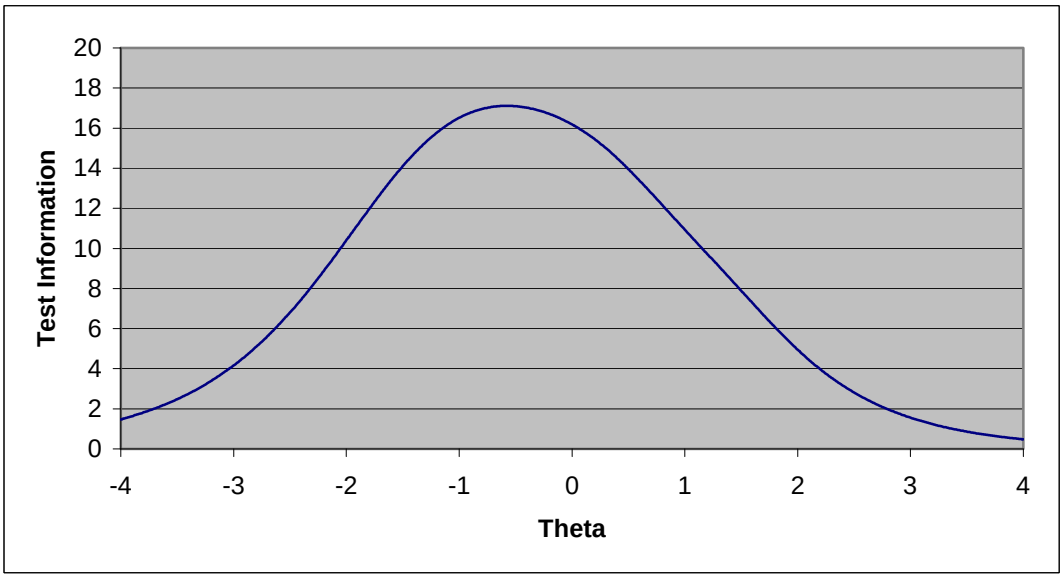

Figure J-4 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 4

Table J-5. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Mathematics Grade 5

Item Number	*Parameters*		
	a	b	c
120642	0.702	-1.756	0.092
120645	0.510	-2.219	0.000
120760	1.023	0.992	0.440
120650	1.007	1.210	0.187
120764	0.866	-0.451	0.217
120762	1.179	0.708	0.409
120713	1.138	1.027	0.144
120728	0.676	-0.055	0.223
120653	0.835	-0.282	0.180
120651	1.040	0.547	0.245
120699	0.650	0.456	0.146
120855	0.640	1.407	0.474
120757	0.669	-1.549	0.000
120769	0.702	0.353	0.103
120661	0.653	-0.140	0.185
120770	0.699	-0.861	0.112
120720	0.957	0.096	0.226
120648	0.888	-1.669	0.110
120683	0.974	0.813	0.238
120678	0.698	-0.012	0.053
120787	0.460	-1.411	0.000
120644	0.971	-0.603	0.111
120779	0.991	-0.390	0.532
120708	0.462	-0.902	0.144
120646	1.371	0.272	0.114
120775	1.141	1.073	0.267
120700	1.319	0.400	0.285
120672	0.969	0.680	0.166
120771	0.501	1.469	0.268
120777	0.519	0.043	0.263
120637	0.864	-0.371	0.164
120690	1.271	0.717	0.092

a = discrimination; b = difficulty; c = guessing

Table J-6. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 5

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
120686	0.779	-0.113	N/A	N/A	N/A	N/A
120717	0.754	0.195	N/A	N/A	N/A	N/A
120730	0.763	-0.268	0.135	-0.135	N/A	N/A
120785	0.786	-0.460	0.479	-0.479	N/A	N/A
120667	0.813	0.518	N/A	N/A	N/A	N/A
120704	0.834	-0.199	0.444	-0.444	N/A	N/A
120772	0.862	-0.847	0.845	-0.845	N/A	N/A
120755	0.726	0.397	N/A	N/A	N/A	N/A
120789	0.818	-0.120	N/A	N/A	N/A	N/A
120698	0.932	0.315	N/A	N/A	N/A	N/A
120639	0.888	-0.326	0.224	-0.224	N/A	N/A
120694	0.468	0.562	0.681	-0.681	N/A	N/A
120731	0.890	0.128	1.172	0.581	-0.531	-1.222
120857	0.973	-0.179	1.697	0.696	-1.030	-1.363
120695	1.099	0.532	0.965	0.457	-0.380	-1.043
120751	0.964	0.108	0.789	0.261	-0.363	-0.687

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

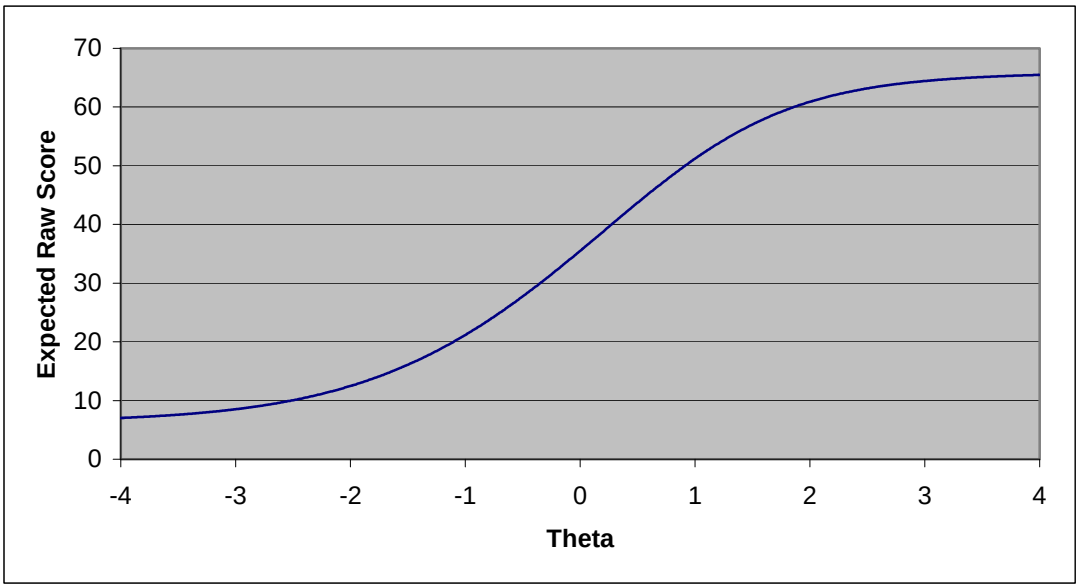

Figure J-5 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 5

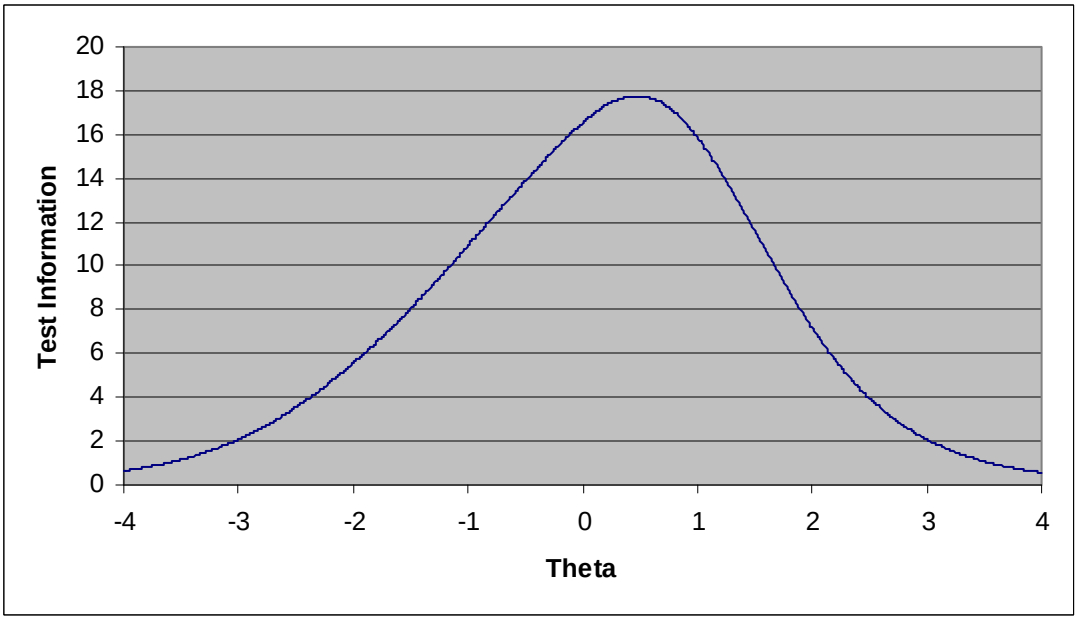

Figure J-6 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 5

Table J-7. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Mathematics Grade 6

Item Number	*Parameters*		
	a	b	c
119317	0.873	1.120	0.386
119236	0.544	-0.483	0.544
119313	0.970	0.773	0.229
119316	0.820	-0.022	0.099
119184	0.930	-0.915	0.191
119344	0.997	0.382	0.109
119214	0.816	0.275	0.115
119311	0.840	0.704	0.181
119185	1.169	0.839	0.158
119186	0.568	-1.771	0.092
119175	0.607	-0.269	0.111
119332	0.259	-1.549	0.000
119288	0.269	-0.478	0.000
119283	1.102	0.238	0.214
119258	1.256	0.215	0.139
119294	1.258	1.071	0.081
119179	1.228	0.272	0.078
119326	0.876	1.009	0.101
119246	1.076	-0.109	0.186
119241	0.734	0.461	0.288
119334	0.800	-1.577	0.121
119181	1.261	-0.723	0.082
119231	0.460	-0.540	0.106
119204	0.968	-0.084	0.208
119251	0.733	1.721	0.164
119211	0.883	0.111	0.175
119199	0.935	0.092	0.228
119333	0.357	-0.876	0.000
119238	0.896	0.363	0.079
119315	1.167	0.238	0.218
119262	0.586	0.793	0.206
119261	0.483	1.392	0.167

a = discrimination; b = difficulty; c = guessing

Table J-8. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 6

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
119272	0.695	-0.268	N/A	N/A	N/A	N/A
119337	0.517	-1.158	N/A	N/A	N/A	N/A
119230	0.798	-0.923	0.346	-0.346	N/A	N/A
119281	0.806	-0.223	0.816	-0.816	N/A	N/A
119319	0.778	-0.906	N/A	N/A	N/A	N/A
119240	0.948	0.943	0.168	-0.168	N/A	N/A
119320	0.707	0.070	0.781	-0.781	N/A	N/A
119263	0.472	0.950	N/A	N/A	N/A	N/A
119277	1.283	-0.599	N/A	N/A	N/A	N/A
119253	0.947	0.207	N/A	N/A	N/A	N/A
119205	0.808	0.594	0.275	-0.275	N/A	N/A
119203	0.657	0.060	1.097	-1.097	N/A	N/A
119322	1.276	0.800	1.218	0.306	-0.549	-0.975
119346	0.904	0.841	1.490	0.267	-0.564	-1.193
119269	0.824	1.787	1.628	0.651	-0.488	-1.790
119343	1.381	0.691	1.120	0.150	-0.212	-1.058

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

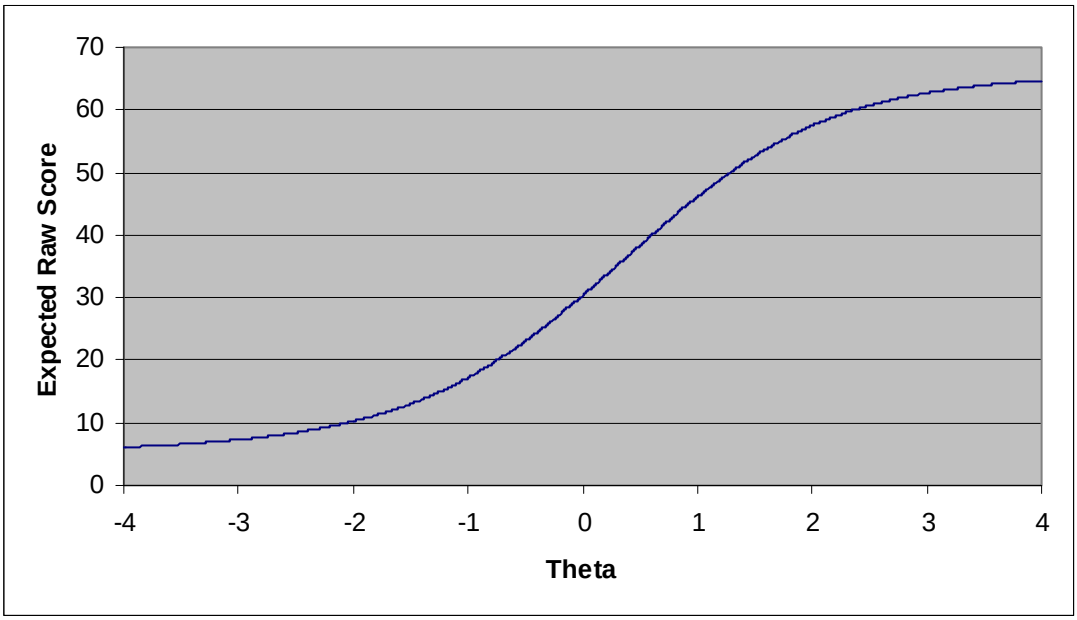

Figure J-7 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 6

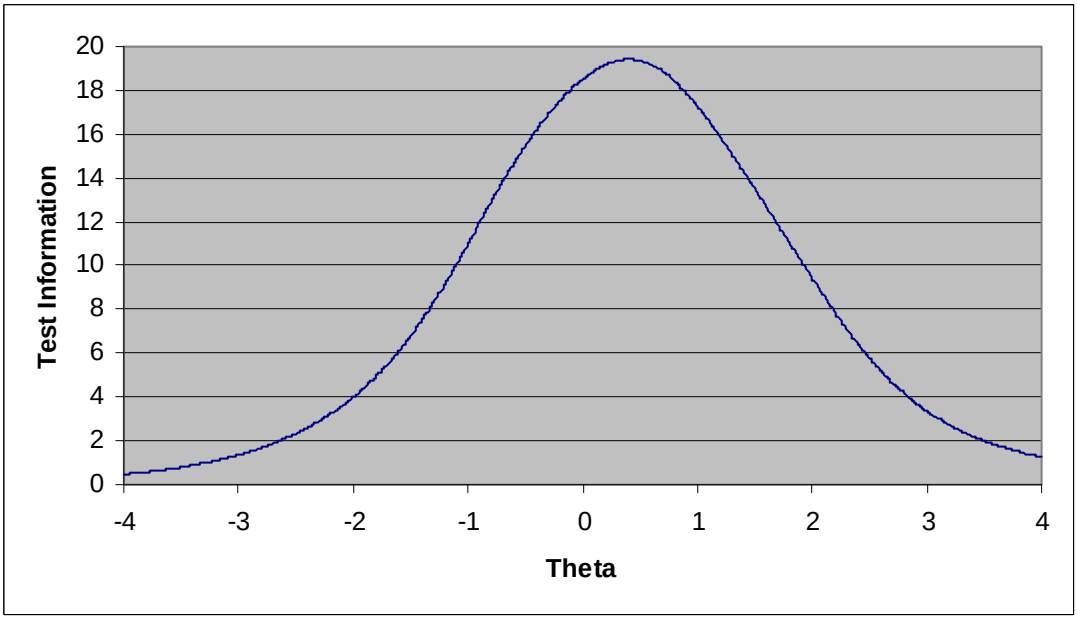

Figure J-8 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 6

Table J-9. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Mathematics Grade 7

Item Number	*Parameters*		
	a	b	c
120464	0.893	0.008	0.298
120441	0.983	0.594	0.071
120444	0.549	-0.188	0.108
120349	0.865	0.183	0.175
120351	0.951	-0.066	0.127
120385	0.808	0.966	0.181
120342	1.444	1.051	0.207
120346	0.379	-1.826	0.089
120446	1.241	0.554	0.144
120447	1.345	1.379	0.164
120460	0.857	-0.630	0.169
120472	0.599	-1.262	0.000
120335	1.087	0.431	0.250
120356	0.654	1.193	0.197
120450	0.509	-0.996	0.054
120411	1.037	-1.432	0.067
120377	0.690	0.085	0.191
120331	0.516	-2.516	0.000
120334	0.576	-0.201	0.053
120467	0.914	-0.859	0.125
120364	0.834	-0.573	0.139
120462	0.852	0.390	0.059
120407	1.452	0.729	0.105
120337	0.694	-0.421	0.164
120404	0.712	0.272	0.257
120455	0.842	0.387	0.210
120362	0.367	-0.081	0.000
120531	0.356	-0.932	0.000
120465	0.677	1.525	0.186
120394	0.717	1.705	0.262
120453	1.254	1.716	0.187
120361	0.889	-0.549	0.119

a = discrimination; b = difficulty; c = guessing

Table J-10. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 7

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
120434	0.985	0.129	N/A	N/A	N/A	N/A
120387	0.786	0.059	N/A	N/A	N/A	N/A
120376	0.770	0.760	0.531	-0.531	N/A	N/A
120423	1.137	0.791	0.482	-0.482	N/A	N/A
120338	0.626	-0.729	N/A	N/A	N/A	N/A
120369	0.905	0.473	0.250	-0.250	N/A	N/A
120360	0.894	1.509	0.641	-0.641	N/A	N/A
120456	0.485	0.207	N/A	N/A	N/A	N/A
120533	1.134	0.261	N/A	N/A	N/A	N/A
120366	0.800	0.690	N/A	N/A	N/A	N/A
120477	0.890	0.952	0.237	-0.237	N/A	N/A
120417	1.092	0.685	0.484	-0.484	N/A	N/A
120355	1.106	0.671	1.894	0.259	-0.251	-1.902
120458	0.978	1.228	0.877	0.193	-0.188	-0.881
120530	1.144	0.295	0.667	0.325	-0.241	-0.751
120393	0.984	0.741	1.476	0.374	-0.518	-1.333

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-9 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 7

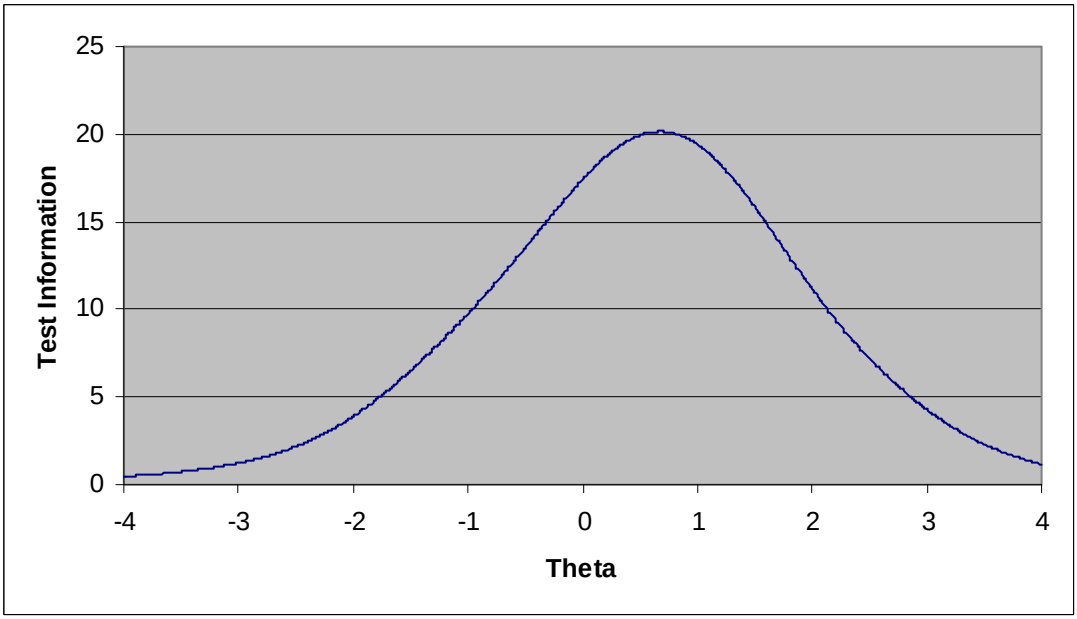

Figure J-10 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 7

Table J-11. 2008–09 NECAP: IRT Parameters Multiple-Choice Items—Mathematics Grade 8

Item Number	*Parameters*		
	a	b	c
120915	0.932	-1.056	0.219
120896	0.895	0.379	0.201
120899	0.917	-1.157	0.068
120897	1.123	1.426	0.317
120917	1.158	-0.658	0.058
120886	1.390	-0.493	0.107
120924	1.597	0.566	0.133
120994	0.965	1.266	0.065
120998	1.262	0.769	0.140
120902	0.974	-0.377	0.127
121024	0.657	-1.189	0.079
120894	1.019	0.477	0.256
120929	1.179	0.808	0.223
120961	1.040	0.526	0.283
121020	0.583	0.177	0.088
121022	0.913	0.114	0.141
120970	1.263	1.095	0.307
120974	1.204	0.374	0.175
121008	0.993	0.515	0.216
120976	1.218	0.396	0.237
121006	1.077	0.988	0.167
120911	1.012	-0.004	0.018
120941	1.322	1.352	0.102
121018	0.937	-1.279	0.034
120921	1.012	-0.857	0.296
120879	0.768	0.433	0.162
120952	0.466	0.730	0.166
121025	1.210	-0.041	0.259
120960	1.716	0.260	0.272
121004	0.427	0.381	0.078
121037	1.225	0.359	0.193
120946	0.893	0.327	0.197

a = discrimination; b = difficulty; c = guessing

Table J-12. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Mathematics Grade 8

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
120900	0.813	-0.524	N/A	N/A	N/A	N/A
120933	1.032	-0.073	N/A	N/A	N/A	N/A
120876	1.053	-0.063	0.782	-0.782	N/A	N/A
120890	0.716	-0.232	0.698	-0.698	N/A	N/A
120931	0.902	0.373	0.0000	0.0000	N/A	N/A
121031	1.207	-0.077	0.728	-0.728	N/A	N/A
120958	0.959	1.264	0.135	-0.135	N/A	N/A
120888	0.994	-1.069	N/A	N/A	N/A	N/A
121040	0.438	-1.016	N/A	N/A	N/A	N/A
120983	0.949	0.739	N/A	N/A	N/A	N/A
120881	0.606	1.522	0.883	-0.883	N/A	N/A
120937	1.112	0.374	0.249	-0.249	N/A	N/A
121002	1.253	0.668	1.433	0.529	-0.601	-1.361
121014	1.780	1.476	1.204	0.376	-0.737	-0.843
121042	1.204	0.547	1.749	0.568	-0.585	-1.731
121032	1.499	-0.150	1.204	0.376	-0.737	-0.843

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

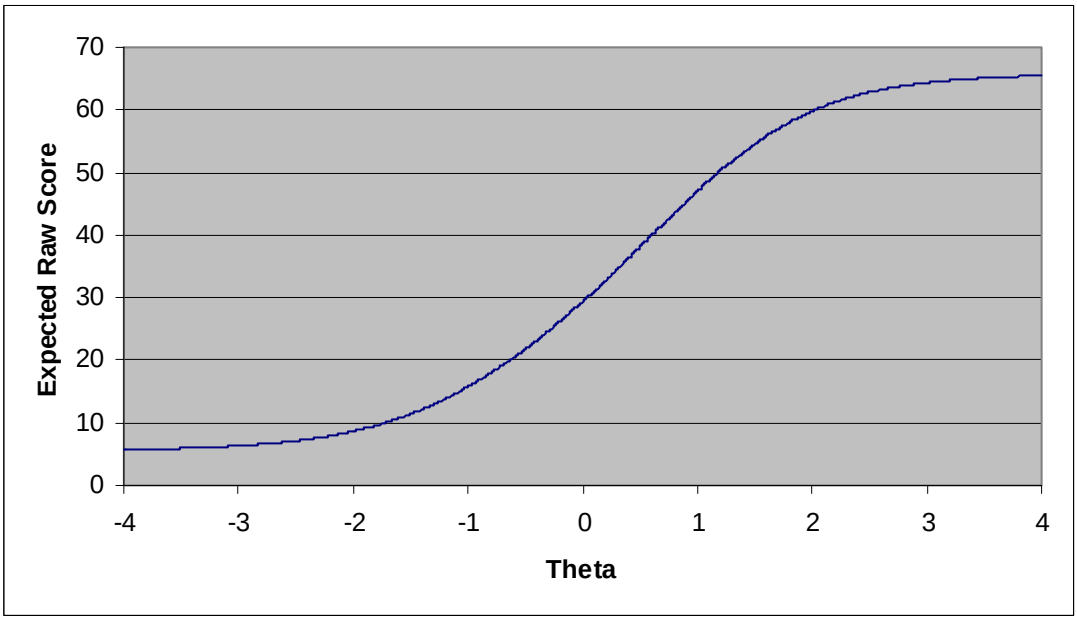

Figure J-11 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 8

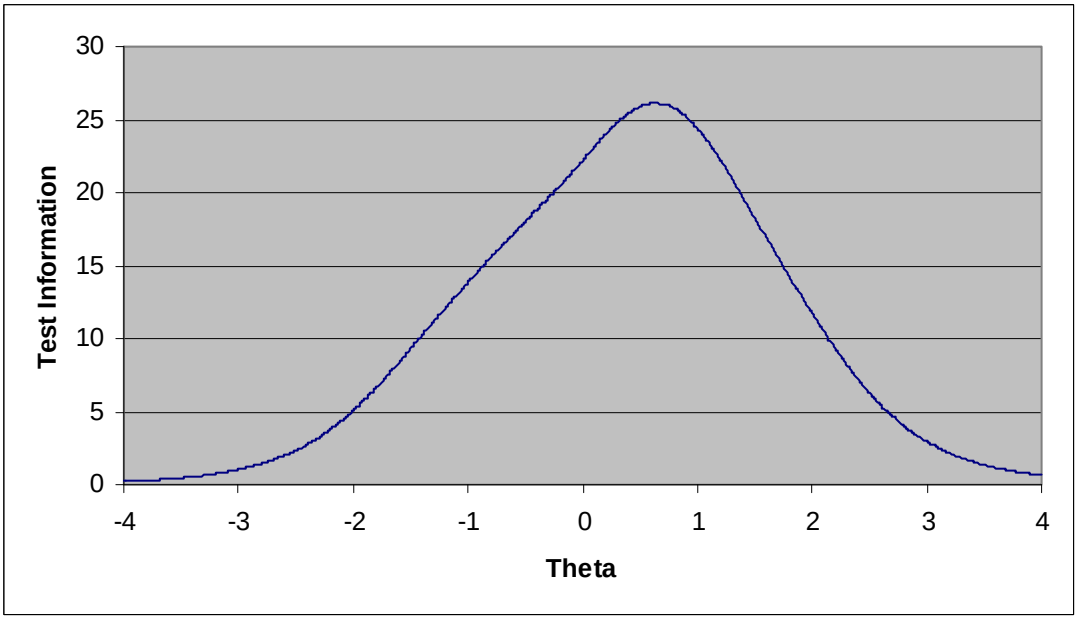

Figure J-12 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 8

Table J-13. 2008–09 NECAP: IRT Parameters Multiple-Choice Items—Mathematics Grade 11

Item Number	*Parameters*		
	a	b	c
119454	0.712	-0.857	0.147
119432	0.933	0.523	0.233
119484	1.581	0.236	0.152
119420	0.758	0.134	0.190
119508	1.241	1.425	0.173
119424	1.210	0.687	0.129
119536	1.388	0.282	0.198
119527	0.702	0.285	0.208
119451	1.592	1.241	0.162
119417	1.926	1.344	0.216
119531	1.700	1.037	0.057
119514	1.328	1.407	0.229
119534	0.670	-1.541	0.055
119437	1.054	0.409	0.333
119513	1.207	0.893	0.160
119529	0.840	-0.042	0.202
119419	0.694	-1.130	0.207
119549	0.639	0.587	0.274
119430	1.067	1.256	0.150
119452	0.783	-0.033	0.128
119532	1.069	0.243	0.128
119516	1.461	1.499	0.421
119450	1.835	0.599	0.128
119444	1.660	1.050	0.168

a = discrimination; b = difficulty; c = guessing

Table J-14. 2008–09 NECAP: IRT Parameters
Constructed Response Items—Mathematics Grade 11

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
119495	1.299	0.203	N/A	N/A	N/A	N/A
119441	1.119	1.521	N/A	N/A	N/A	N/A
119494	1.160	1.005	N/A	N/A	N/A	N/A
119490	1.082	1.465	N/A	N/A	N/A	N/A
119500	1.232	0.597	N/A	N/A	N/A	N/A
119509	1.516	1.750	N/A	N/A	N/A	N/A
119472	0.785	0.887	0.686	-0.686	N/A	N/A
119543	1.059	0.797	0.085	-0.085	N/A	N/A
119511	1.280	1.830	0.238	-0.238	N/A	N/A
119552	0.674	1.035	N/A	N/A	N/A	N/A
119518	1.197	-0.191	N/A	N/A	N/A	N/A
119492	0.696	-1.401	N/A	N/A	N/A	N/A
119470	0.912	0.772	N/A	N/A	N/A	N/A
119520	1.106	-0.032	N/A	N/A	N/A	N/A
119553	1.029	0.551	N/A	N/A	N/A	N/A
119544	0.979	-0.273	0.355	-0.355	N/A	N/A
119521	1.493	0.990	0.083	-0.083	N/A	N/A
119554	0.980	1.564	0.483	-0.483	N/A	N/A
119547	1.225	0.296	1.332	0.531	-0.576	-1.287
119460	1.357	1.572	1.384	0.399	-0.513	-1.270
119546	1.167	0.496	0.610	0.410	-0.319	-0.701
119523	1.089	1.051	0.835	0.425	-0.540	-0.720

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

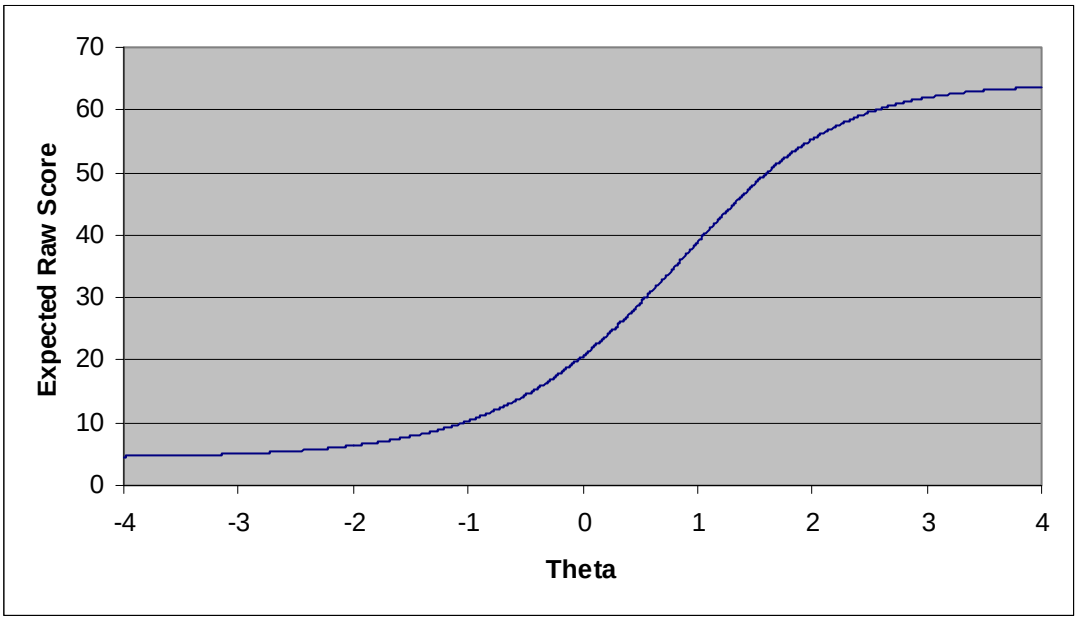

Figure J-13 2008–09 NECAP: Test Characteristic Curve (TCC)—Mathematics Grade 11

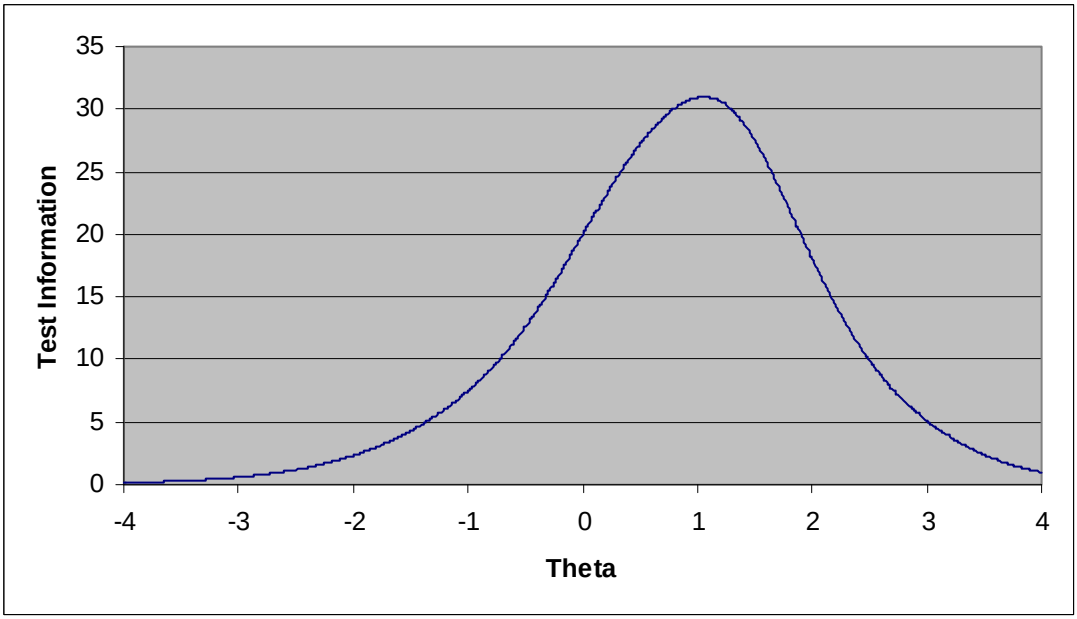

Figure J-14 2008–09 NECAP: Test Information Function (TIF)—Mathematics Grade 11

Table J-15. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 3

Item Number	*Parameters*		
	a	b	c
117649	0.976	-1.721	0.091
117663	1.346	-1.342	0.091
117671	1.285	-1.510	0.126
117672	1.297	-1.302	0.119
117673	0.852	-1.477	0.084
117674	0.612	0.622	0.197
117641	1.279	-0.127	0.187
117665	0.744	-0.010	0.233
117642	1.420	-0.250	0.273
117643	1.130	-1.557	0.239
117666	0.737	0.271	0.250
117644	0.763	-0.228	0.245
117667	0.559	0.446	0.261
117668	0.777	-1.417	0.059
117697	1.130	-1.150	0.090
117700	1.189	-0.274	0.170
117699	0.976	-0.275	0.191
117702	0.701	-1.282	0.000
117701	0.823	0.058	0.186
117703	1.339	-0.605	0.163
117698	1.061	-0.010	0.201
117704	0.691	-0.557	0.117
117733	0.831	-2.010	0.000
117734	0.607	-1.366	0.087
117687	0.722	-1.060	0.052
117688	0.698	-1.343	0.065
117690	0.738	-0.964	0.154
117689	0.767	-0.669	0.040

a = discrimination; b = difficulty; c = guessing

Table J-16. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 3

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
117675	0.702	0.883	2.770	0.677	-0.998	-2.450
117650	0.786	-0.580	1.982	0.366	-0.648	-1.700
117669	0.912	0.803	2.108	0.716	-0.802	-2.022
117705	0.477	-2.570	2.646	0.781	-0.993	-2.434
117706	1.105	-0.263	0.652	0.188	-0.235	-0.605
117691	0.712	-1.001	1.200	0.555	-0.230	-1.526

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-15 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 3

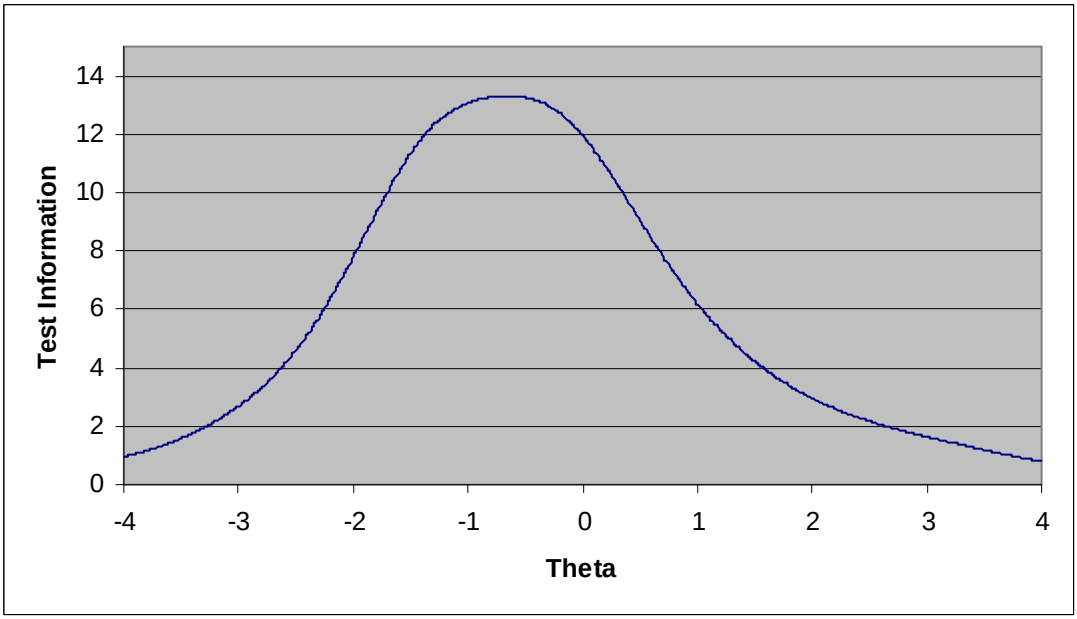

Figure J-16 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 3

Table J-17. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 4

Item Number	*Parameters*		
	a	b	c
117971	0.729	-1.578	0.000
117915	0.672	-0.326	0.181
117901	1.167	-1.135	0.118
117902	0.474	-1.446	0.000
117903	0.561	-0.223	0.152
117900	0.844	-0.790	0.250
117863	1.560	-0.831	0.233
117864	1.088	-1.237	0.145
117865	1.345	-0.796	0.189
117866	1.341	-0.591	0.137
117867	0.730	-1.045	0.000
117868	1.147	-0.867	0.144
117870	0.783	-0.789	0.078
117869	0.810	-1.865	0.000
117872	0.898	-0.148	0.158
117873	1.004	0.280	0.180
117894	0.897	0.292	0.194
117897	0.714	0.815	0.231
117895	0.832	1.184	0.269
117874	1.097	0.229	0.254
117899	1.010	-1.250	0.087
117898	0.867	-0.510	0.191
117974	0.518	-1.351	0.095
117973	0.779	0.074	0.122
118008	0.616	-1.161	0.205
118009	0.905	0.314	0.383
118010	0.697	-0.584	0.308
118011	0.766	-2.075	0.000

a = discrimination; b = difficulty; c = guessing

Table J-18. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 4

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
117904	0.705	-0.921	1.372	0.598	-0.366	-1.604
117879	0.605	-0.090	1.273	0.855	-0.755	-1.373
117871	0.779	0.488	2.281	0.788	-0.849	-2.220
117896	0.518	-0.132	2.194	1.410	-0.966	-2.638
117875	0.587	1.119	2.380	0.725	-0.789	-2.316
118012	0.403	-0.611	1.571	0.253	-0.405	-1.420

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

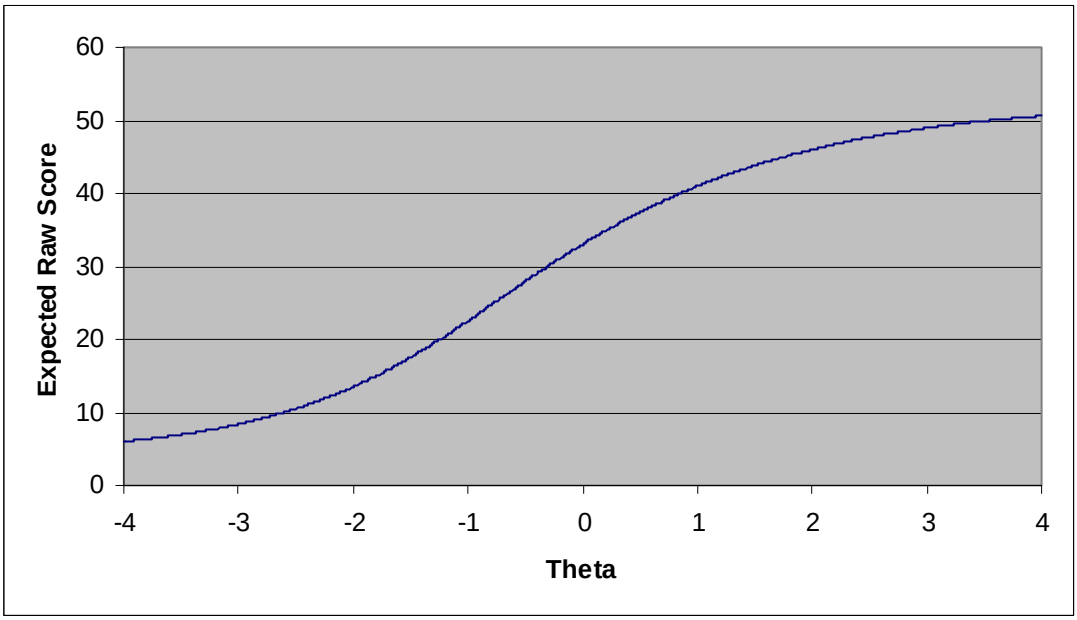

Figure J-17 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 4

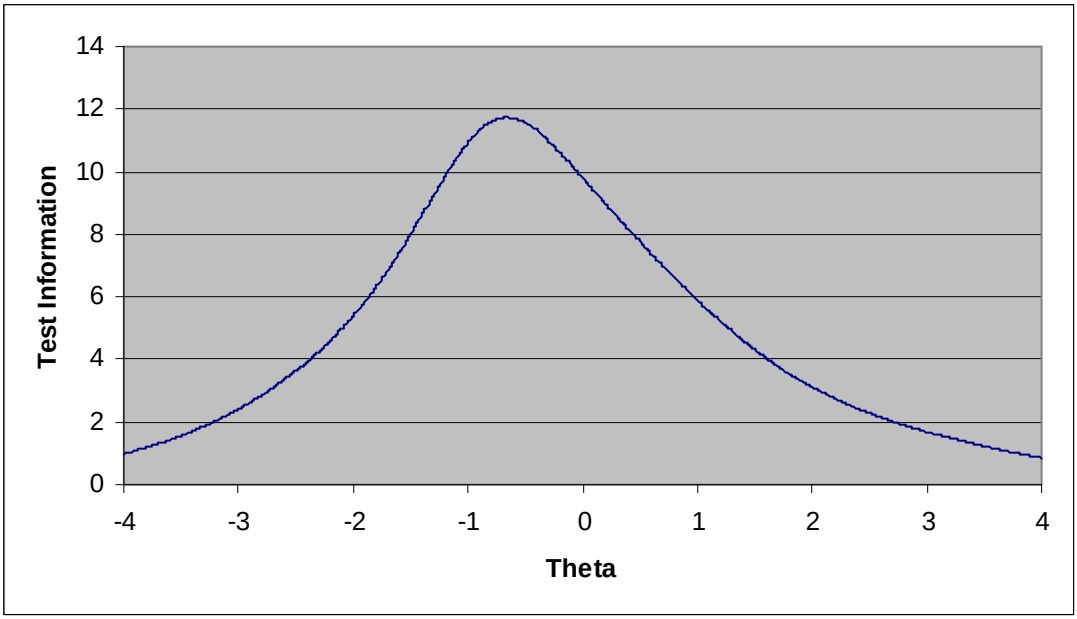

Figure J-18 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 4

Table J-19. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 5

Item Number	*Parameters*		
	a	b	c
118061	0.671	-0.660	0.133
118121	0.697	-0.666	0.066
118068	1.333	-1.330	0.238
118069	1.217	-0.372	0.240
118070	0.235	0.046	0.068
118071	0.850	-0.200	0.247
118109	0.486	0.519	0.100
118111	0.876	-2.036	0.000
118110	0.813	-0.739	0.188
118112	0.606	-0.843	0.179
118113	0.776	-0.595	0.103
118115	1.002	-1.069	0.112
118116	0.682	0.241	0.101
118114	0.350	0.238	0.075
118093	0.736	-1.168	0.000
118094	0.872	0.013	0.154
118096	0.878	-1.054	0.083
118095	0.834	-0.215	0.164
118097	1.300	-1.163	0.115
118098	1.045	-0.758	0.219
118099	0.795	0.006	0.231
118100	1.219	-1.073	0.149
118119	0.902	-0.593	0.177
118063	0.823	-0.832	0.222
118034	0.534	-0.405	0.158
118036	1.185	-0.555	0.180
118035	0.353	-2.034	0.063
118066	0.777	0.658	0.195

a = discrimination; b = difficulty; c = guessing

Table J-20. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 5

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
118072	0.739	0.954	2.803	0.713	-0.933	-2.583
118118	0.659	1.085	2.706	0.425	-0.873	-2.258
118117	0.822	1.318	2.827	0.788	-1.091	-2.524
118102	0.869	0.688	3.059	0.623	-1.088	-2.594
118101	0.895	0.729	2.863	0.853	-0.927	-2.789
118037	1.048	0.232	2.447	0.786	-0.958	-2.274

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-19 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 5

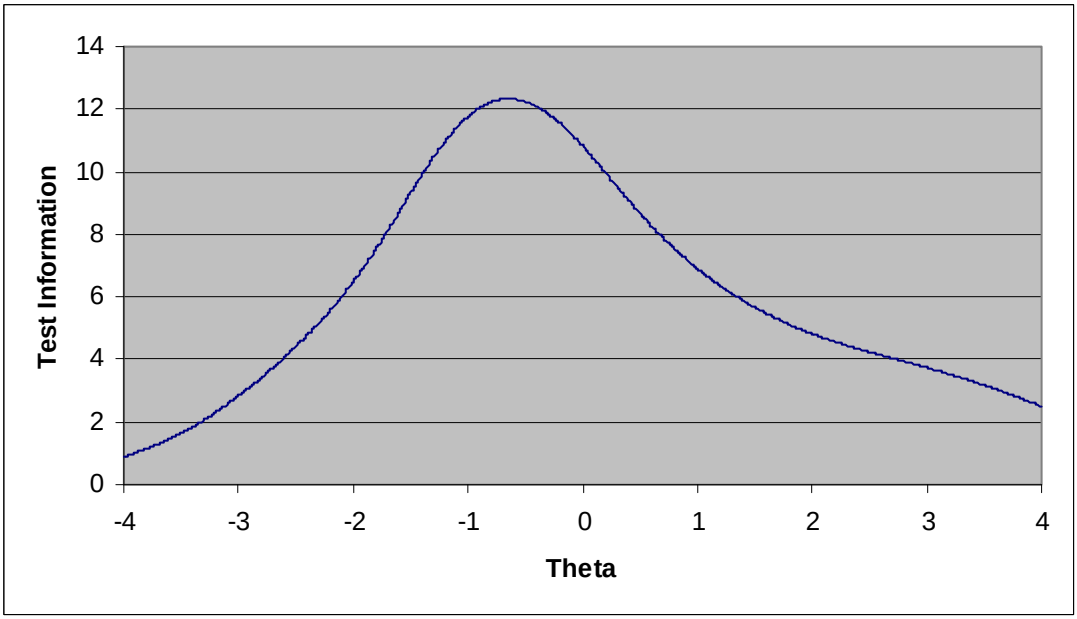

Figure J-20 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 5

Table J-21. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 6

Item Number	*Parameters*		
	a	b	c
118303	0.442	-1.755	0.111
118232	0.284	-1.868	0.062
118256	0.740	-0.707	0.090
118257	0.718	-0.834	0.133
118258	0.479	-1.985	0.000
118259	0.628	-1.211	0.053
118241	0.710	-1.083	0.058
118242	0.684	-2.334	0.000
118243	0.603	-1.653	0.085
118244	0.894	-1.701	0.062
118245	0.521	-1.150	0.000
118246	1.244	-1.813	0.067
118248	0.665	-1.635	0.000
118247	0.721	-1.437	0.000
118266	1.091	-1.011	0.117
118267	0.840	-0.744	0.155
118268	0.481	-0.993	0.000
118269	0.988	-1.331	0.093
118270	0.863	-0.908	0.081
118272	0.618	0.406	0.130
118271	1.396	-1.183	0.166
118273	0.876	-0.922	0.115
118302	0.718	-1.598	0.000
118305	0.783	-1.952	0.106
118276	0.593	-0.028	0.102
118277	0.506	-1.316	0.044
118278	0.819	-1.731	0.000
118279	0.907	-1.582	0.062

a = discrimination; b = difficulty; c = guessing

Table J-22. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 6

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
118260	0.941	0.422	2.395	0.736	-0.887	-2.243
118249	0.932	0.117	2.131	0.897	-0.680	-2.347
118250	1.030	0.508	2.300	0.613	-0.841	-2.072
118275	1.178	0.086	1.942	0.711	-0.745	-1.908
118274	1.102	0.338	2.405	0.659	-0.910	-2.155
118280	0.794	0.041	3.138	0.908	-1.169	-2.877

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-21 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 6

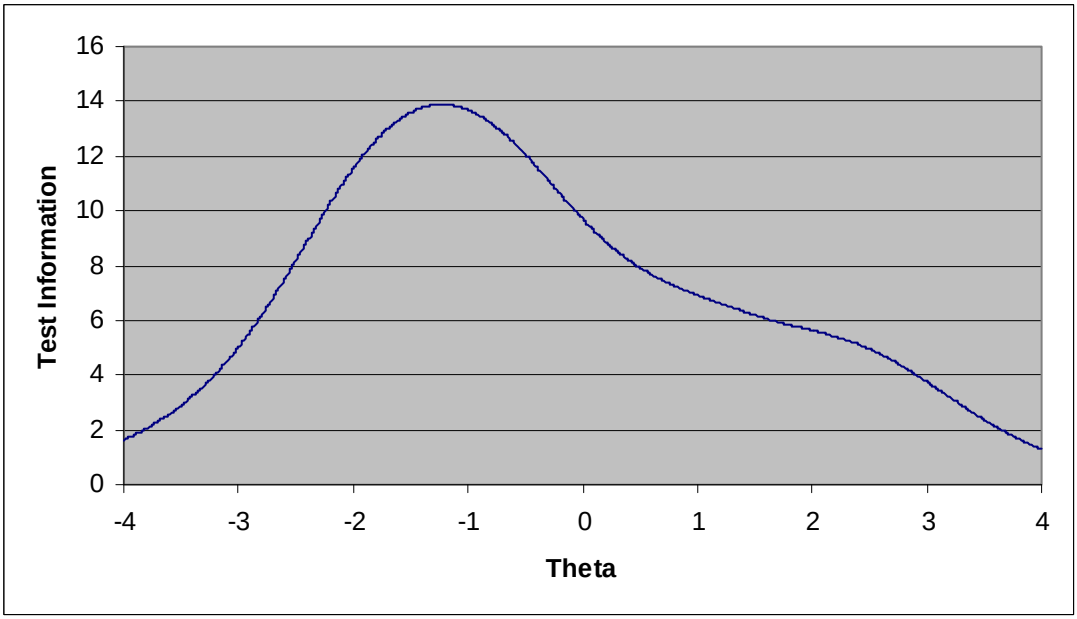

Figure J-22 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 6

Table J-23. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 7

Item Number	*Parameters*		
	a	b	c
118479	0.868	-1.612	0.093
118474	0.542	-1.374	0.090
118406	0.414	-1.369	0.000
118405	0.798	-0.933	0.212
118407	0.818	-2.150	0.000
118408	0.366	-0.115	0.178
118435	0.632	-2.178	0.000
118436	0.515	-0.279	0.061
118437	0.657	-2.022	0.000
118438	0.750	-1.519	0.057
118440	0.574	-0.641	0.000
118439	0.905	-1.374	0.050
118441	0.681	-0.531	0.088
118434	0.867	-0.959	0.095
118410	0.449	1.036	0.190
118411	0.850	-1.594	0.161
118412	0.819	-1.726	0.090
118446	0.455	-0.512	0.033
118413	0.483	-0.256	0.026
118414	0.977	-1.188	0.123
118415	0.923	-0.993	0.103
118416	0.696	-2.139	0.000
118444	0.511	-1.041	0.107
118476	0.286	-0.003	0.000
118453	0.282	-2.342	0.100
118454	0.589	-2.382	0.000
118455	0.351	-1.973	0.057
118456	0.370	-1.447	0.000

a = discrimination; b = difficulty; c = guessing

Table J-24. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 7

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
118409	0.963	-0.477	2.860	1.024	-1.129	-2.756
118443	1.210	0.180	1.877	0.867	-0.777	-1.967
118442	1.161	-0.003	2.177	0.620	-0.788	-2.009
118418	1.024	0.256	2.311	0.974	-1.050	-2.234
118417	0.944	0.018	2.374	0.647	-0.851	-2.170
118457	0.928	-0.383	2.700	0.971	-1.088	-2.583

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

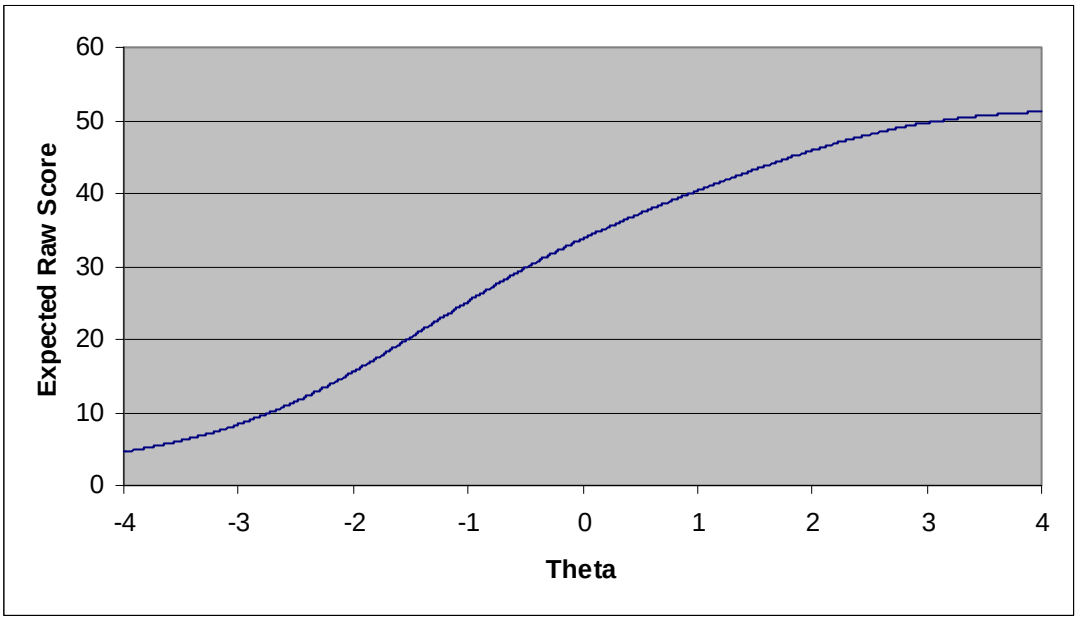

Figure J-23 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 7

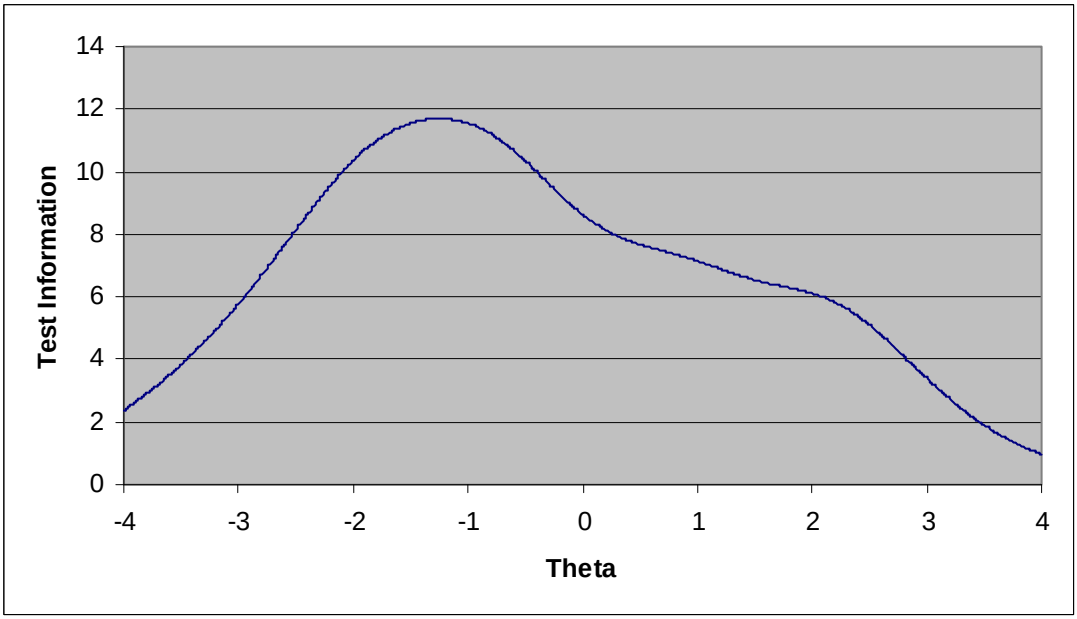

Figure J-24 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 7

Table J-25. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 8

Item Number	*Parameters*		
	a	b	c
118615	0.474	-2.787	0.000
118616	0.663	-2.282	0.201
118636	0.857	-2.207	0.100
118637	0.783	-0.896	0.273
118638	0.317	-2.484	0.078
118639	0.843	-2.078	0.065
118662	0.650	-1.179	0.153
118664	0.448	-1.707	0.054
118663	0.860	-0.944	0.143
118665	0.853	-0.687	0.141
118666	0.479	-1.172	0.000
118667	0.964	-1.484	0.123
118668	0.704	-1.054	0.105
118669	0.648	-2.289	0.000
118599	0.243	-0.471	0.220
118600	0.493	-1.747	0.079
118601	0.505	-1.987	0.000
118602	0.640	-0.913	0.122
118603	0.486	-2.677	0.000
118604	0.433	-2.189	0.000
118605	0.499	0.659	0.079
118606	0.955	-1.691	0.136
118661	0.841	-1.358	0.151
118658	0.263	-1.652	0.000
118631	0.825	-1.677	0.106
118632	0.218	-1.672	0.000
118633	0.656	-0.595	0.136
118634	0.325	-1.710	0.000

a = discrimination; b = difficulty; c = guessing

Table J-26. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 8

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
118640	1.002	-0.052	1.984	0.986	-0.779	-2.191
118670	1.217	0.014	2.424	0.779	-0.856	-2.347
118671	1.166	0.513	2.683	0.655	-0.981	-2.357
118608	0.856	0.344	2.719	1.106	-1.193	-2.632
118607	0.861	-0.111	2.739	0.782	-1.109	-2.413
118635	1.109	-0.263	2.211	0.655	-0.719	-2.148

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-25 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 8

Figure J-26 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 8

Table J-27. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Reading Grade 11

Item Number	*Parameters*		
	a	b	c
118827	0.784	-1.103	0.354
118786	0.374	1.037	0.267
118784	0.555	-1.561	0.196
118828	0.241	-1.795	0.000
118801	0.486	-1.333	0.087
118807	0.594	-0.894	0.000
118802	0.660	-0.725	0.089
118803	0.604	-0.642	0.084
118804	0.891	-1.261	0.130
118806	0.798	-1.252	0.083
118805	0.357	0.094	0.099
118808	0.664	-2.469	0.000
118778	0.643	-2.095	0.000
118779	0.546	-0.724	0.064
118780	0.510	-1.266	0.000
118781	0.894	-0.909	0.156
118755	0.593	-0.516	0.168
118754	0.471	-1.505	0.000
118753	0.639	-0.865	0.090
118756	0.561	-0.429	0.120
118816	0.307	-2.162	0.000
118817	0.491	-1.046	0.134
118818	0.824	-0.662	0.165
118819	0.613	-1.056	0.084
118820	0.871	-1.561	0.076
118822	0.731	-1.393	0.000
118821	0.876	-0.495	0.101
118823	0.370	-0.282	0.000

a = discrimination; b = difficulty; c = guessing

Table J-28. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Reading Grade 11

Item Number	*Parameters*					
	a	b	D1	D2	D3	D4
118809	1.000	-0.226	2.202	0.994	-0.834	-2.362
118810	1.205	-0.159	1.837	0.762	-0.706	-1.893
118782	1.295	-0.369	1.690	0.664	-0.648	-1.705
118757	1.307	0.343	2.000	0.741	-0.760	-1.981
118824	1.094	-0.207	2.223	0.989	-0.894	-2.319
118825	1.249	0.266	1.895	0.797	-0.715	-1.976

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter

Figure J-27 2008–09 NECAP: Test Characteristic Curve (TCC)—Reading Grade 11

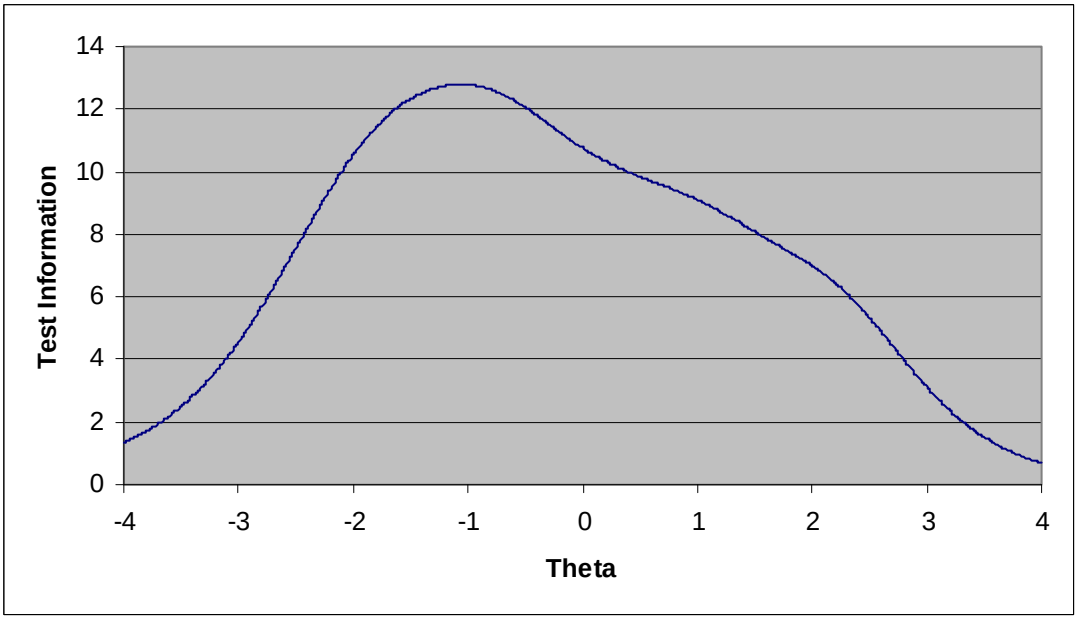

Figure J-28 2008–09 NECAP: Test Information Function (TIF)—Reading Grade 11

Table J-29. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Writing Grade 5

Item Number	*Parameters*		
	a	b	c
213392	0.686	-0.831	0.035
213395	0.688	-1.607	0.042
202780	0.674	-0.802	0.117
202839	0.692	-1.858	0.058
213142	0.410	-2.115	0.110
202776	0.395	1.181	0.092
213157	0.539	-1.479	0.070
202834	0.735	-1.668	0.070
213382	0.800	-1.536	0.071
202790	0.574	-0.944	0.078

a = discrimination; b = difficulty; c = guessing

Table J-30. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Writing Grade 5

Item Number	*Parameters*											
	a	b	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
201852	0.656	0.000	3.319	1.299	-2.212	-4.706	N/A	N/A	N/A	N/A	N/A	N/A
201810	0.767	0.000	3.389	0.802	-1.605	-3.823	N/A	N/A	N/A	N/A	N/A	N/A
201953	0.642	0.000	2.944	1.435	-1.891	-4.462	N/A	N/A	N/A	N/A	N/A	N/A
213658	0.551	0.000	2.999	2.106	0.999	0.389	-0.788	-1.394	-2.663	-3.45228	-4.99514	-6.03287

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter; …; D10 = 10th category step parameter
Note: Short-answer items are not included in this table because they were not part of the final calibration.

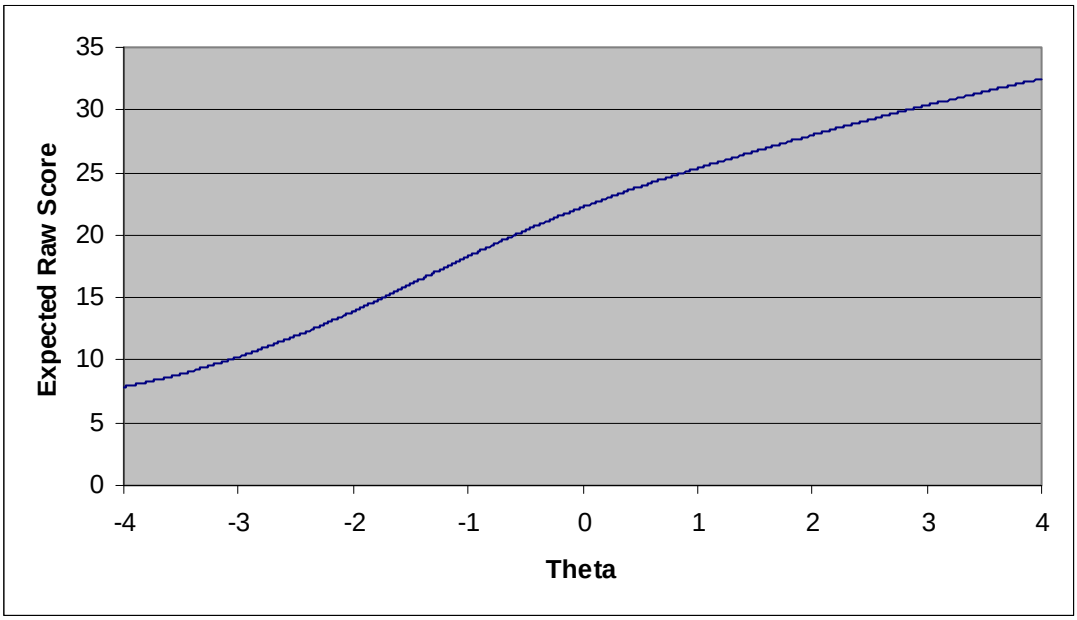

Figure J-29 2008–09 NECAP: Test Characteristic Curve (TCC)—Writing Grade 5

Figure J-30 2008–09 NECAP: Test Information Function (TIF)—Writing Grade 5

Table J-31. 2008–09 NECAP: IRT Parameters
Multiple-Choice Items—Writing Grade 8

Item Number	*Parameters*		
	a	b	c
202653	0.805	-2.012	0.200
202602	0.808	-1.818	0.031
202641	0.592	-1.921	0.125
212964	0.550	0.507	0.076
212975	0.585	-1.582	0.084
202647	0.629	-0.731	0.063
212960	0.907	-0.962	0.027
202670	0.770	-0.483	0.079
212970	1.325	-1.569	0.000
202642	0.581	-0.160	0.094

a = discrimination; b = difficulty; c = guessing

Table J-32. 2008–09 NECAP: IRT Parameters
Constructed-Response Items—Writing Grade 8

Item Number	*Parameters*											
	a	b	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
202423	1.006	0.000	2.878	1.252	-0.813	-2.531	N/A	N/A	N/A	N/A	N/A	N/A
202450	1.021	0.000	2.495	1.236	-0.321	-2.143	N/A	N/A	N/A	N/A	N/A	N/A
202453	0.875	0.000	2.752	1.327	-0.886	-2.649	N/A	N/A	N/A	N/A	N/A	N/A
213712	0.654	0.000	2.853	2.168	1.209	0.610	-0.243	-0.667	-1.495	-2.00649	-3.01014	-4.09554

a = discrimination; b = difficulty; D1 = 1st category step parameter; D2 = 2nd category step parameter; D3 = 3rd category step parameter; D4 = 4th category step parameter; …; D10 = 10th category step parameter
Note: Short-answer items are not included in this table because they were not part of the final calibration.

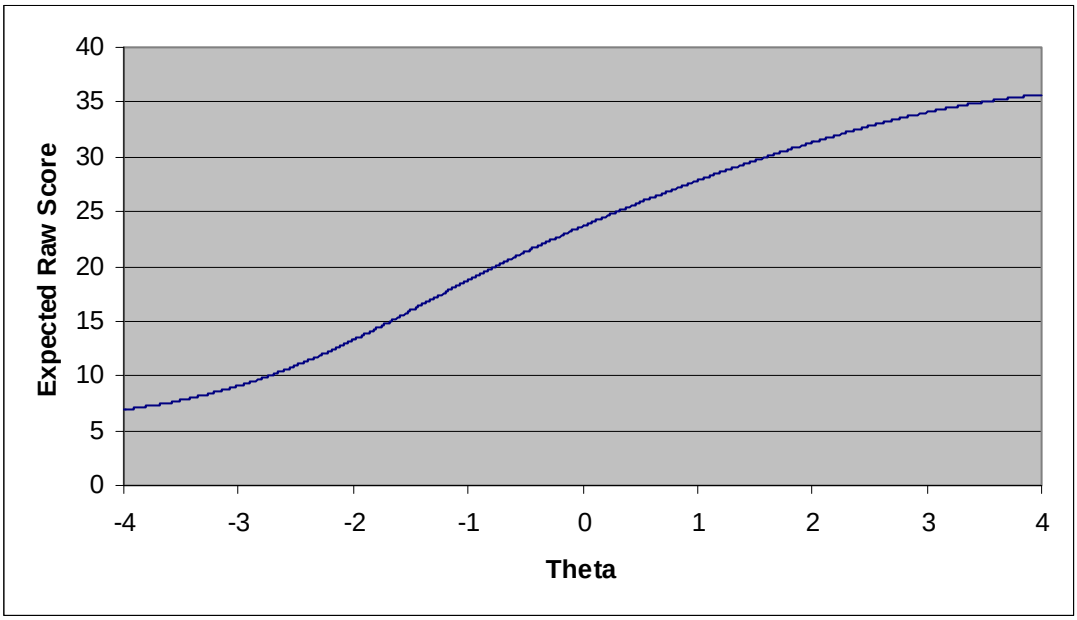

Figure J-31 2008–09 NECAP: Test Characteristic Curve (TCC)—Writing Grade 8

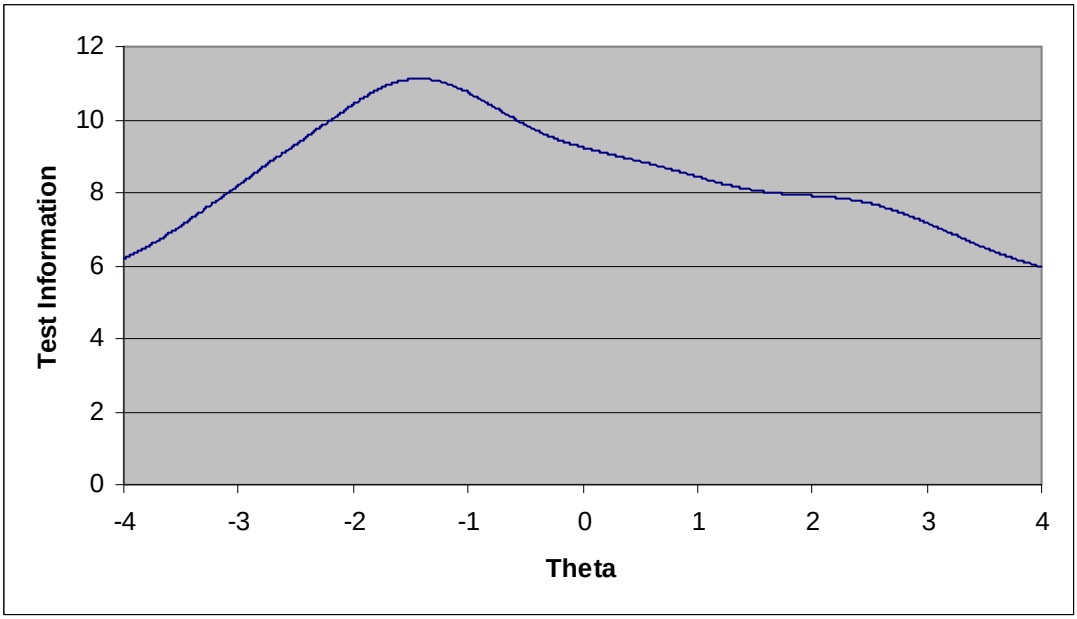

Figure J-32 2008–09 NECAP: Test Information Function (TIF)—Writing Grade 8

Appendix K—Delta Analyses and Rescore Analysis Results

Table K-1. 2008-09 NECAP: Delta Analysis—Mathematics Grade 3

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
202089	2	1.0634	1.0537	0.9632	0.9329	-0.0101	0.0098	NO
231019	2	1.4439	1.4195	0.792	0.832	-0.0308	0.0244	NO
234216	2	0.9415	0.9463	0.898	0.8955	0.0054	0.0049	NO
242311	2	1.3415	1.2976	0.6481	0.6509	-0.0677	0.0439	NO
198507	2	1.5659	1.561	0.6565	0.657	-0.0074	0.0049	NO
242779	2	0.9415	0.9561	0.9352	0.9229	0.0156	0.0146	NO
256001	2	0.9463	0.9122	0.8335	0.8335	-0.041	0.0341	NO
198631	2	0.7707	0.7902	0.8211	0.8084	0.0238	0.0195	NO
198631	2	0.7707	0.7902	0.8211	0.8084	0.0238	0.0195	NO
198505	2	1.5122	1.4878	0.7366	0.7232	-0.0331	0.0244	NO
198521	2	0.8829	0.8732	0.8473	0.8574	-0.0115	0.0098	NO
257424	2	0.6585	0.6683	0.7843	0.7885	0.0124	0.0098	NO
223936	2	1.5024	1.478	0.6218	0.6369	-0.0392	0.0244	NO

Table K-2. 2008-09 NECAP: Delta Analysis—Mathematics Grade 4

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
202368	2	1.0196	1.0539	0.7857	0.7554	0.0437	0.0343	NO
198445	2	1.3971	1.4853	0.7696	0.7308	0.1146	0.0882	NO
232429	2	1.2683	1.2634	0.8896	0.8828	-0.0055	0.0049	NO
224099	2	1.678	1.6634	0.6276	0.6835	-0.0233	0.0146	NO
198427	2	1.5073	1.5366	0.7498	0.7358	0.039	0.0293	NO
227082	2	1.3614	1.3663	0.5568	0.558	0.0089	0.005	NO
202370	2	0.9317	0.9366	0.8863	0.8839	0.0055	0.0049	NO
270069	2	1.0539	1.0245	0.8977	0.9044	-0.0328	0.0294	NO
198431	2	0.7756	0.9317	0.8195	0.7491	0.1905	0.1561	NO
227116	2	1.0439	1.0488	0.7734	0.7948	0.0063	0.0049	NO
255743	2	1.2524	1.3204	0.7723	0.7202	0.088	0.068	NO
202436	2	0.9024	0.8927	0.7192	0.738	-0.0136	0.0098	NO
202436	2	0.9024	0.8927	0.7192	0.738	-0.0136	0.0098	NO

Table K-3. 2008-09 NECAP: Delta Analysis—Mathematics Grade 5

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
269407	4	1.3676	1.3922	1.4167	1.4393	0.0173	0.0245	NO
272113	2	0.797	0.7921	0.7729	0.7812	-0.0064	0.005	NO
230964	4	1.7171	1.722	1.151	1.1585	0.0042	0.0049	NO
255255	2	0.9366	0.8537	0.809	0.7638	-0.1025	0.0829	NO
255249	2	1.0294	0.9804	0.874	0.8964	-0.0561	0.049	NO
255150	2	0.7756	0.7756	0.8075	0.8135	0	0	NO
203621	2	0.5707	0.5951	0.797	0.8006	0.0306	0.0244	NO
269405	2	1.0847	1.0282	0.9136	0.9171	-0.0618	0.0565	NO
198603	2	1.3463	1.4	0.8335	0.7815	0.0644	0.0537	NO
198603	2	1.3463	1.4	0.8335	0.7815	0.0644	0.0537	NO
198569	4	1.5539	1.6078	1.2493	1.3406	0.0432	0.0539	NO
198566	2	0.8966	0.8621	0.862	0.8657	-0.04	0.0345	NO
203432	2	0.5441	0.4902	0.7878	0.7701	-0.0684	0.0539	NO
203432	2	0.5441	0.4902	0.7878	0.7701	-0.0684	0.0539	NO
269314	4	1.8676	1.8039	1.5137	1.5213	-0.0421	0.0637	NO
234368	2	0.7192	0.6552	0.8509	0.836	-0.0753	0.064	NO
234368	2	0.7192	0.6552	0.8509	0.836	-0.0753	0.064	NO
255178	2	1.0441	1.1127	0.8062	0.8177	0.0851	0.0686	NO
203855	4	1.8824	1.8578	1.2271	1.2023	-0.02	0.0245	NO

Table K-4. 2008-09 NECAP: Delta Analysis—Mathematics Grade 6

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
225334	4	1.7192	1.6749	1.4971	1.4631	-0.0296	0.0443	NO
198727	2	0.6195	0.6146	0.7787	0.7669	-0.0063	0.0049	NO
198637	2	0.8878	0.9268	0.901	0.8942	0.0433	0.039	NO
270545	4	1.3676	1.348	1.2975	1.3141	-0.0151	0.0196	NO
271718	2	0.6488	0.6195	0.7009	0.6638	-0.0418	0.0293	NO
198665	2	0.5951	0.7122	0.7102	0.6402	0.1648	0.1171	NO
270596	4	1.678	1.5951	1.3006	1.3314	-0.0638	0.0829	NO
270694	2	1.3756	1.161	0.778	0.783	-0.2759	0.2146	NO
203259	2	1.1667	1.0735	0.8754	0.8629	-0.1064	0.0931	NO
225370	2	0.2718	0.2524	0.6638	0.6266	-0.0293	0.0194	NO
225370	2	0.2718	0.2524	0.6638	0.6266	-0.0293	0.0194	NO
198706	4	1.2745	1.2598	1.4863	1.484	-0.0099	0.0147	NO
198706	4	1.2745	1.2598	1.4863	1.484	-0.0099	0.0147	NO
203255	2	0.7366	0.7951	0.8316	0.8364	0.0704	0.0585	NO
255530	4	1.3951	1.4488	1.5633	1.5968	0.0343	0.0537	NO
203502	2	0.3756	0.3073	0.7193	0.6613	-0.0949	0.0683	NO
225371	2	0.4755	0.4853	0.8251	0.8134	0.0119	0.0098	NO
234419	2	1.2365	1.2857	0.8951	0.8917	0.055	0.0493	NO
234419	2	1.2365	1.2857	0.8951	0.8917	0.055	0.0493	NO
234414	4	1.7902	1.7756	1.347	1.3246	-0.0109	0.0146	NO

Table K-5. 2008-09 NECAP: Delta Analysis—Mathematics Grade 7

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
225135	2	0.6585	0.6098	0.8557	0.8519	-0.057	0.0488	NO
224856	2	0.5343	0.549	0.8653	0.8646	0.017	0.0147	NO
255899	2	0.5415	0.5366	0.7016	0.7018	-0.007	0.0049	NO
256004	2	0.6244	0.6683	0.8325	0.8365	0.0527	0.0439	NO
206195	4	2.2634	2.2439	1.5707	1.5489	-0.0124	0.0195	NO
269312	2	0.5512	0.5659	0.6352	0.6712	0.023	0.0146	NO
225137	2	0.4853	0.5147	0.77	0.8253	0.0382	0.0294	NO
256015	4	1.0296	1.0493	1.2824	1.1942	0.0154	0.0197	NO
269083	4	0.8976	0.9122	1.1829	1.1943	0.0124	0.0146	NO
269083	4	0.8976	0.9122	1.1829	1.1943	0.0124	0.0146	NO
206189	2	0.7941	0.7843	0.8948	0.8532	-0.011	0.0098	NO
234455	2	0.5561	0.5366	0.5526	0.5272	-0.0353	0.0195	NO
256122	4	1.8293	2.1805	1.0523	1.2343	0.3337	0.3512	NO
199932	2	0.5756	0.6049	0.7394	0.7619	0.0396	0.0293	NO
206215	2	0.8235	0.8775	0.9332	0.9392	0.0578	0.0539	NO
256095	2	0.5931	0.5637	0.6232	0.619	-0.0472	0.0294	NO
206213	2	1.1225	1.2647	0.9019	0.8903	0.1576	0.1422	NO
206213	2	1.1225	1.2647	0.9019	0.8903	0.1576	0.1422	NO
224876	4	1.1379	1.069	1.4691	1.409	-0.0469	0.069	NO
224876	4	1.1379	1.069	1.4691	1.409	-0.0469	0.069	NO

Table K-6. 2008-09 NECAP: Delta Analysis—Mathematics Grade 8

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
206240	2	1.5074	1.5222	0.8025	0.7773	0.0184	0.0148	NO
234148	2	0.7366	0.7024	0.5664	0.5879	-0.0603	0.0341	NO
206352	4	1.2573	1.3544	1.1728	1.0864	0.0828	0.0971	NO
256536	2	1.0878	1.122	0.9536	0.9212	0.0358	0.0341	NO
199783	2	0.6545	0.6303	0.5123	0.5306	-0.0473	0.0242	NO
269197	4	1.4537	1.4341	1.4697	1.4354	-0.0133	0.0195	NO
256062	2	0.6422	0.6422	0.7569	0.7439	0	0	NO
256107	4	1.7463	1.8195	1.5505	1.5017	0.0472	0.0732	NO
269172	2	0.0634	0.3122	0.2809	0.576	0.8857	0.2488	YES
269325	4	1.1912	1.1961	1.3568	1.3397	0.0036	0.0049	NO
206324	2	1.0833	1.1765	0.8448	0.8679	0.1103	0.0931	NO
206331	4	2.1854	2.1854	1.0931	1.1196	0	0	NO
224947	2	0.6878	0.6976	0.8668	0.8475	0.0113	0.0098	NO
269098	2	0.3951	0.4488	0.7424	0.8046	0.0723	0.0537	NO
269090	2	0.6829	0.7073	0.8216	0.7665	0.0297	0.0244	NO
206335	4	1.8382	1.8382	1.4479	1.5108	0	0	NO

Table K-7. 2008-09 NECAP: Delta Analysis—Mathematics Grade 11

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
260008	2	0.561	0.5463	0.7207	0.6943	-0.0203	0.0146	NO
259998	4	1.0637	1.0294	1.4249	1.3752	-0.0241	0.0343	NO
270716	4	1.0683	1.0634	1.2078	1.202	-0.004	0.0049	NO
259936	2	0.161	0.1512	0.5118	0.5147	-0.0191	0.0098	NO
259986	4	1.5025	1.532	0.9436	0.9378	0.0313	0.0296	NO
259928	2	0.2488	0.2439	0.5424	0.5401	-0.009	0.0049	NO
259921	2	0.3284	0.3676	0.6223	0.6621	0.063	0.0392	NO
260009	4	2.0098	2.0293	1.2257	1.2607	0.0159	0.0195	NO
260675	2	0.2244	0.2098	0.4923	0.4838	-0.0297	0.0146	NO
260675	2	0.2244	0.2098	0.4923	0.4838	-0.0297	0.0146	NO
259958	4	0.7549	0.7598	1.2322	1.2549	0.004	0.0049	NO

Table K-8. 2008-09 NECAP: Delta Analysis—Reading Grade 3

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
270066	4	1.5147	1.5539	0.7889	0.8237	0.0497	0.0392	NO
255338	4	3.2341	3.239	1.1744	1.1714	0.0042	0.0049	NO
255336	4	2.2341	2.7122	1.1191	1.0454	0.4272	0.478	NO
230973	4	1.8284	1.9412	1.0914	1.0876	0.1033	0.1127	NO
225242	4	3.6488	3.6488	0.6725	0.6725	0	0	NO
225253	4	1.8537	1.8	1.0491	1.0188	-0.0511	0.0537	NO
201764	4	2.4118	2.451	1.3781	1.3145	0.0285	0.0392	NO
270392	4	3.6195	3.6146	0.7269	0.7344	-0.0067	0.0049	NO
270404	4	2.0488	2.1366	1.1338	1.1815	0.0774	0.0878	NO

Table K-9. 2008-09 NECAP: Delta Analysis—Reading Grade 4

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
225776	4	2.3234	2.3731	1.0074	1.0246	0.0494	0.0498	NO
225778	4	1.5245	1.4755	0.9622	0.9622	-0.0509	0.049	NO
270626	4	1.3756	1.4244	1.0171	1.0171	0.048	0.0488	NO
270511	4	2.6244	2.6146	1.1264	1.1317	-0.0087	0.0098	NO
270516	4	1.6927	1.6195	0.8429	0.8095	-0.0868	0.0732	NO
200843	4	2.478	2.2341	1.3273	1.3877	-0.1838	0.2439	NO
205951	4	2.7463	2.8049	1.2703	1.2419	0.0461	0.0585	NO
203684	4	1.6829	1.5805	1.1902	1.1389	-0.0861	0.1024	NO

Table K-10. 2008-09 NECAP: Delta Analysis—Reading Grade 5

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
201937	4	1.8529	1.8382	1.0882	1.0186	-0.0135	0.0147	NO
202072	4	1.7402	1.8088	1.0319	1.0135	0.0665	0.0686	NO
202075	4	1.9461	2.0588	0.8117	0.8892	0.1389	0.1127	NO
201769	4	1.922	1.7024	1.0187	0.8913	-0.2155	0.2195	NO
256415	4	1.6098	1.4829	0.9339	0.8867	-0.1358	0.1268	NO
256370	4	1.5588	1.6275	0.8174	0.8393	0.084	0.0686	NO
269205	4	1.6127	1.6176	0.8756	0.9346	0.0056	0.0049	NO
226515	4	1.3463	1.3561	1.0131	0.9292	0.0096	0.0098	NO
226517	4	1.6716	1.6569	0.866	0.9074	-0.017	0.0147	NO

Table K-11. 2008-09 NECAP: Delta Analysis—Reading Grade 6

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
226669	4	1.902	1.8431	0.8746	0.7508	-0.0673	0.0588	NO
269782	4	1.7718	1.8641	0.925	0.9456	0.0997	0.0922	NO
269778	4	1.9073	1.8244	1.0435	1.1168	-0.0795	0.0829	NO
200348	4	2.0439	1.9902	1.0136	0.9318	-0.0529	0.0537	NO
204026	4	1.9512	1.9561	0.813	0.8796	0.006	0.0049	NO
204022	4	1.8732	1.9415	0.9232	0.9088	0.074	0.0683	NO
269525	4	1.8927	1.922	0.8769	0.8965	0.0334	0.0293	NO
200324	4	2.0833	2.1373	1.0087	0.9955	0.0535	0.0539	NO
200325	4	1.7024	1.8	1.1323	1.1236	0.0862	0.0976	NO

Table K-12. 2008-09 NECAP: Delta Analysis—Reading Grade 7

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
201535	4	1.9951	1.9608	1.0314	0.9894	-0.0333	0.0343	NO
270702	4	2.0293	2.0976	0.8718	0.9161	0.0783	0.0683	NO
270706	4	1.7073	1.6878	0.7537	0.7259	-0.0259	0.0195	NO
256108	4	1.878	1.7902	0.9727	0.8439	-0.0903	0.0878	NO
199535	4	1.878	1.8244	0.8023	0.8545	-0.0669	0.0537	NO
199536	4	2.0343	1.8775	0.8424	0.8223	-0.1862	0.1569	NO
201492	4	1.9951	1.9561	1.0887	0.9941	-0.0358	0.039	NO
201490	4	2.0634	1.9756	0.927	0.9548	-0.0947	0.0878	NO

Table K-13. 2008-09 NECAP: Delta Analysis—Reading Grade 8

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
204155	4	1.9902	1.8971	1.0146	1.1264	-0.0918	0.0931	NO
269963	4	2.134	2.11	1.1159	1.1249	-0.0214	0.0239	NO
269964	4	1.9707	1.9707	1.0449	1.0771	0	0	NO
256003	4	2.2	2.1902	0.8746	0.9041	-0.0112	0.0098	NO
204128	4	1.8725	1.9314	1.1041	0.9626	0.0533	0.0588	NO
204133	4	2.152	2.1569	0.864	1.0121	0.0057	0.0049	NO

Table K-14. 2008-09 NECAP: Delta Analysis—Reading Grade 11

IREF	Max Number of Points	Last Usage P Old Mean	Current Usage P New Mean	Last Usage Delta Old Stdev	Current Usage Delta New Stdev	Effect Size	Abs Diff	Discard
258568	4	1.6108	1.7685	0.8072	0.8826	0.1953	0.1576	NO
258570	4	1.6488	1.6439	0.846	0.8639	-0.0058	0.0049	NO
269663	4	1.9366	2.0537	0.8839	0.9117	0.1324	0.1171	NO
258662	4	1.5659	1.5902	0.9008	0.9202	0.0271	0.0244	NO
258664	4	1.6293	1.6244	0.8549	0.8726	-0.0057	0.0049	NO
269465	4	1.8942	1.8942	0.8707	0.8978	0	0	NO

Appendix L—Raw to Scaled Score Look-up Tables

Table L-1. 2008–09 NECAP: Look-up Table—Mathematics Grade 3

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	300	300	310	1
1	-4.00	300	300	310	1
2	-4.00	300	300	310	1
3	-4.00	300	300	310	1
4	-4.00	300	300	310	1
5	-4.00	300	300	310	1
6	-4.00	300	300	310	1
7	-4.00	300	300	309	1
8	-3.66	304	300	311	1
9	-3.32	307	301	313	1
10	-3.05	310	305	315	1
11	-2.83	313	308	318	1
12	-2.64	315	311	320	1
13	-2.48	316	312	320	1
14	-2.33	318	314	322	1
15	-2.19	319	315	323	1
16	-2.07	321	317	325	1
17	-1.95	322	319	326	1
18	-1.84	323	320	326	1
19	-1.74	324	321	327	1
20	-1.64	325	322	328	1
21	-1.55	326	323	329	1
22	-1.46	327	324	330	1
23	-1.37	328	325	331	1
24	-1.29	329	326	332	1
25	-1.21	330	327	333	1
26	-1.13	331	328	334	1
27	-1.06	331	328	334	1
28	-0.98	332	329	335	2
29	-0.91	333	330	336	2
30	-0.84	334	331	337	2
31	-0.77	335	332	338	2
32	-0.70	335	332	338	2
33	-0.63	336	334	339	2
34	-0.56	337	335	340	2
35	-0.49	338	336	341	2
36	-0.42	338	336	341	2
37	-0.36	339	337	342	2
38	-0.29	339	337	342	2
39	-0.22	341	339	344	3
40	-0.15	341	339	344	3
41	-0.09	342	340	345	3
42	-0.02	343	341	346	3
43	0.05	343	341	346	3
44	0.12	344	342	347	3
45	0.19	345	343	348	3
46	0.27	346	344	349	3

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
47	0.34	347	345	350	3
48	0.42	347	344	350	3
49	0.50	348	345	351	3
50	0.58	349	346	352	3
51	0.66	350	347	353	3
52	0.75	351	348	354	3
53	0.84	352	349	355	3
54	0.94	352	349	355	3
55	1.04	354	351	357	4
56	1.15	355	352	358	4
57	1.27	357	354	360	4
58	1.41	358	355	362	4
59	1.56	360	356	364	4
60	1.74	361	357	365	4
61	1.95	364	360	369	4
62	2.22	367	362	372	4
63	2.59	371	365	377	4
64	3.23	377	368	380	4
65	4.00	380	370	380	4

Table L-2. 2008–09 NECAP: Look-up Table—Mathematics Grade 4

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	400	400	410	1
1	-4.00	400	400	410	1
2	-4.00	400	400	410	1
3	-4.00	400	400	410	1
4	-4.00	400	400	410	1
5	-4.00	400	400	410	1
6	-4.00	400	400	410	1
7	-3.87	401	400	410	1
8	-3.49	406	400	413	1
9	-3.20	409	403	415	1
10	-2.98	411	406	416	1
11	-2.79	413	408	418	1
12	-2.62	415	411	420	1
13	-2.48	417	413	421	1
14	-2.35	418	414	422	1
15	-2.23	420	416	424	1
16	-2.12	421	417	425	1
17	-2.02	422	419	425	1
18	-1.92	423	420	426	1
19	-1.83	424	421	427	1
20	-1.74	425	422	428	1
21	-1.65	426	423	429	1
22	-1.57	427	424	430	1
23	-1.49	428	425	431	1

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
24	-1.42	429	426	432	1
25	-1.34	429	426	432	1
26	-1.27	430	427	433	1
27	-1.20	430	427	433	1
28	-1.13	432	429	435	2
29	-1.06	433	430	436	2
30	-0.99	433	430	436	2
31	-0.92	434	431	437	2
32	-0.85	435	432	438	2
33	-0.78	436	433	439	2
34	-0.72	436	433	439	2
35	-0.65	437	434	440	2
36	-0.58	438	435	441	2
37	-0.51	439	436	442	2
38	-0.44	439	436	442	2
39	-0.37	440	437	443	3
40	-0.31	441	438	444	3
41	-0.24	442	439	445	3
42	-0.16	442	439	445	3
43	-0.09	443	440	446	3
44	-0.02	444	441	447	3
45	0.06	445	442	448	3
46	0.13	446	443	449	3
47	0.21	446	443	449	3
48	0.29	447	444	450	3
49	0.37	448	445	451	3
50	0.46	449	446	452	3
51	0.55	450	447	453	3
52	0.64	451	448	454	3
53	0.74	452	449	455	3
54	0.85	454	451	457	3
55	0.96	455	452	458	4
56	1.08	456	453	459	4
57	1.20	457	454	461	4
58	1.35	459	455	463	4
59	1.51	461	457	465	4
60	1.69	463	459	467	4
61	1.90	465	460	470	4
62	2.18	468	463	474	4
63	2.56	472	465	479	4
64	3.21	480	470	480	4
65	4.00	480	470	480	4

Table L-3. 2008–09 NECAP: Look-up Table—Mathematics Grade 5

Raw Score	*Θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-8.75	500	500	510	1
1	-8.08	500	500	510	1
2	-7.40	500	500	510	1
3	-6.73	500	500	510	1
4	-6.05	500	500	510	1
5	-5.37	500	500	510	1
6	-4.70	500	500	510	1
7	-4.02	500	500	510	1
8	-3.24	508	500	517	1
9	-2.81	513	506	520	1
10	-2.51	516	510	522	1
11	-2.28	519	514	524	1
12	-2.08	521	516	526	1
13	-1.92	522	518	526	1
14	-1.77	524	520	528	1
15	-1.63	525	521	529	1
16	-1.51	527	523	531	1
17	-1.40	528	524	532	1
18	-1.29	529	525	533	1
19	-1.20	530	527	533	1
20	-1.10	531	528	534	1
21	-1.01	532	529	535	1
22	-0.93	532	529	535	1
23	-0.85	534	531	537	2
24	-0.77	535	532	538	2
25	-0.70	536	533	539	2
26	-0.62	536	533	539	2
27	-0.55	537	534	540	2
28	-0.48	538	535	541	2
29	-0.41	539	536	542	2
30	-0.35	539	536	542	2
31	-0.28	540	537	543	3
32	-0.22	541	538	544	3
33	-0.16	541	538	544	3
34	-0.09	542	539	545	3
35	-0.03	543	540	546	3
36	0.03	543	540	546	3
37	0.09	544	541	547	3
38	0.15	545	542	548	3
39	0.21	545	542	548	3
40	0.27	546	543	549	3
41	0.33	547	544	550	3
42	0.40	547	544	550	3
43	0.46	548	545	551	3
44	0.52	549	546	552	3
45	0.58	549	546	552	3
46	0.65	550	547	553	3
47	0.71	551	548	554	3
48	0.78	551	548	554	3
49	0.84	552	549	555	3

(cont'd)

Raw Score	Θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
50	0.91	553	550	556	3
51	0.99	553	550	556	3
52	1.06	554	551	557	4
53	1.14	555	552	558	4
54	1.22	556	553	559	4
55	1.31	557	554	560	4
56	1.40	558	555	561	4
57	1.50	559	556	562	4
58	1.60	560	557	563	4
59	1.72	562	559	566	4
60	1.86	563	559	567	4
61	2.02	565	561	569	4
62	2.21	567	563	572	4
63	2.45	569	564	574	4
64	2.79	573	567	580	4
65	3.40	580	570	580	4
66	4.00	580	570	580	4

Table L-4. 2008–09 NECAP: Look-up Table—Mathematics Grade 6

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	600	600	610	1
1	-4.00	600	600	610	1
2	-4.00	600	600	610	1
3	-4.00	600	600	610	1
4	-4.00	600	600	610	1
5	-4.00	600	600	610	1
6	-4.00	600	600	610	1
7	-3.18	609	600	619	1
8	-2.67	614	606	622	1
9	-2.32	618	612	624	1
10	-2.05	621	616	627	1
11	-1.84	623	618	628	1
12	-1.66	625	621	629	1
13	-1.50	626	622	630	1
14	-1.36	628	624	632	1
15	-1.24	629	625	633	1
16	-1.12	630	627	633	1
17	-1.02	632	629	635	1
18	-0.92	632	629	635	1
19	-0.83	634	631	637	2
20	-0.74	635	632	638	2
21	-0.66	635	632	638	2
22	-0.58	636	633	639	2
23	-0.50	637	634	640	2
24	-0.43	638	635	641	2
25	-0.36	639	636	642	2

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
26	-0.29	639	636	642	2
27	-0.22	640	638	643	3
28	-0.15	641	639	644	3
29	-0.09	641	639	644	3
30	-0.02	642	640	645	3
31	0.04	643	641	646	3
32	0.11	643	641	645	3
33	0.17	644	642	646	3
34	0.23	645	643	647	3
35	0.29	645	643	647	3
36	0.36	646	644	648	3
37	0.42	647	645	649	3
38	0.48	647	645	649	3
39	0.54	648	646	650	3
40	0.60	649	647	651	3
41	0.67	649	647	651	3
42	0.73	650	648	652	3
43	0.79	651	649	654	3
44	0.86	651	649	654	3
45	0.93	652	650	655	3
46	0.99	652	650	655	3
47	1.06	654	651	657	4
48	1.14	654	651	657	4
49	1.21	655	652	658	4
50	1.29	656	653	659	4
51	1.37	657	654	660	4
52	1.45	658	655	661	4
53	1.54	659	656	662	4
54	1.63	660	657	663	4
55	1.72	661	658	664	4
56	1.83	662	659	665	4
57	1.94	663	660	666	4
58	2.07	664	660	668	4
59	2.21	666	662	670	4
60	2.37	667	663	671	4
61	2.56	669	665	674	4
62	2.79	672	667	677	4
63	3.10	675	669	680	4
64	3.53	680	672	680	4
65	4.00	680	671	680	4
66	4.00	680	671	680	4

Table L-5. 2008–09 NECAP: Look-up Table—Mathematics Grade 7

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-8.81	700	700	710	1
1	-7.90	700	700	710	1
2	-7.00	700	700	710	1
3	-6.09	700	700	710	1
4	-5.18	700	700	710	1
5	-4.27	700	700	710	1
6	-3.27	707	700	717	1
7	-2.68	713	705	721	1
8	-2.30	717	711	723	1
9	-2.02	720	715	725	1
10	-1.79	723	718	728	1
11	-1.60	724	720	728	1
12	-1.44	726	722	730	1
13	-1.30	728	724	732	1
14	-1.17	729	726	733	1
15	-1.05	730	727	733	1
16	-0.94	731	728	734	1
17	-0.84	732	729	735	1
18	-0.74	733	730	736	1
19	-0.65	734	731	737	2
20	-0.56	735	732	738	2
21	-0.48	736	733	739	2
22	-0.40	737	734	740	2
23	-0.33	737	734	740	2
24	-0.26	738	735	741	2
25	-0.19	739	737	742	2
26	-0.12	739	737	742	2
27	-0.05	740	738	743	3
28	0.01	741	739	743	3
29	0.08	742	740	744	3
30	0.14	742	740	744	3
31	0.20	743	741	745	3
32	0.26	743	741	745	3
33	0.32	744	742	746	3
34	0.38	745	743	747	3
35	0.44	745	743	747	3
36	0.49	746	744	748	3
37	0.55	746	744	748	3
38	0.61	747	745	749	3
39	0.67	748	746	750	3
40	0.73	748	746	750	3
41	0.79	749	747	751	3
42	0.84	749	747	751	3
43	0.90	750	748	752	3
44	0.96	751	749	753	3
45	1.03	751	749	753	3
46	1.09	751	749	753	3
47	1.15	753	751	755	4
48	1.22	753	751	755	4
49	1.29	754	752	756	4

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
50	1.36	755	753	758	4
51	1.43	755	753	758	4
52	1.51	756	753	759	4
53	1.58	757	754	760	4
54	1.67	758	755	761	4
55	1.76	759	756	762	4
56	1.85	760	757	763	4
57	1.96	761	758	764	4
58	2.07	762	759	765	4
59	2.19	763	760	766	4
60	2.33	765	762	769	4
61	2.49	766	762	770	4
62	2.68	768	764	772	4
63	2.92	771	766	776	4
64	3.24	774	768	780	4
65	3.80	780	772	780	4
66	4.00	780	771	780	4

Table L-6. 2008–09 NECAP: Look-up Table—Mathematics Grade 8

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	800	800	810	1
1	-4.00	800	800	810	1
2	-4.00	800	800	810	1
3	-4.00	800	800	810	1
4	-4.00	800	800	810	1
5	-4.00	800	800	810	1
6	-3.38	806	800	816	1
7	-2.56	815	808	822	1
8	-2.17	818	813	823	1
9	-1.92	821	817	825	1
10	-1.72	823	819	827	1
11	-1.56	825	822	828	1
12	-1.42	826	823	829	1
13	-1.29	827	824	830	1
14	-1.18	828	825	831	1
15	-1.07	829	826	832	1
16	-0.98	830	827	833	1
17	-0.88	831	828	834	1
18	-0.80	832	829	835	1
19	-0.71	833	831	836	1
20	-0.63	834	832	836	2
21	-0.56	835	833	837	2
22	-0.48	835	833	837	2
23	-0.41	836	834	838	2
24	-0.34	837	835	839	2
25	-0.27	838	836	840	2

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
26	-0.21	838	836	840	2
27	-0.14	839	837	841	2
28	-0.08	839	837	841	2
29	-0.02	840	838	842	3
30	0.04	841	839	843	3
31	0.10	841	839	843	3
32	0.16	842	840	844	3
33	0.22	842	840	844	3
34	0.28	843	841	845	3
35	0.33	844	842	846	3
36	0.39	844	842	846	3
37	0.44	845	843	847	3
38	0.50	845	843	847	3
39	0.55	846	844	848	3
40	0.61	846	844	848	3
41	0.66	847	845	849	3
42	0.72	848	846	850	3
43	0.77	848	846	850	3
44	0.83	849	847	851	3
45	0.89	849	847	851	3
46	0.94	850	848	852	3
47	1.00	850	848	852	3
48	1.06	851	849	853	3
49	1.12	852	850	854	4
50	1.19	852	850	854	4
51	1.25	853	851	855	4
52	1.32	854	852	856	4
53	1.39	854	852	856	4
54	1.46	855	853	857	4
55	1.54	856	854	858	4
56	1.62	857	855	860	4
57	1.71	857	854	860	4
58	1.80	858	855	861	4
59	1.91	860	857	863	4
60	2.03	861	858	864	4
61	2.16	862	859	865	4
62	2.33	864	860	868	4
63	2.54	866	862	870	4
64	2.85	869	864	874	4
65	3.40	875	867	880	4
66	4.00	880	870	880	4

Table L-7. 2008–09 NECAP: Look-up Table—Mathematics Grade 11

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	1100	1100	1110	1
1	-4.00	1100	1100	1110	1
2	-4.00	1100	1100	1110	1
3	-4.00	1100	1100	1110	1
4	-4.00	1100	1100	1110	1
5	-3.07	1108	1100	1118	1
6	-2.18	1116	1110	1122	1
7	-1.74	1120	1115	1125	1
8	-1.45	1122	1118	1126	1
9	-1.23	1124	1120	1128	1
10	-1.05	1126	1123	1129	1
11	-0.89	1127	1124	1130	1
12	-0.76	1128	1125	1131	1
13	-0.64	1129	1126	1132	1
14	-0.53	1130	1128	1133	1
15	-0.44	1131	1129	1133	1
16	-0.35	1132	1130	1134	1
17	-0.26	1132	1130	1134	1
18	-0.18	1133	1131	1135	1
19	-0.11	1134	1132	1136	2
20	-0.04	1134	1132	1136	2
21	0.03	1135	1133	1137	2
22	0.09	1135	1133	1137	2
23	0.16	1136	1134	1138	2
24	0.22	1137	1135	1139	2
25	0.28	1137	1135	1139	2
26	0.33	1138	1136	1140	2
27	0.39	1138	1136	1140	2
28	0.45	1138	1136	1140	2
29	0.50	1139	1137	1141	2
30	0.55	1139	1137	1141	2
31	0.60	1139	1137	1141	2
32	0.66	1140	1138	1142	3
33	0.71	1141	1139	1143	3
34	0.76	1141	1139	1143	3
35	0.81	1142	1140	1144	3
36	0.86	1142	1140	1144	3
37	0.91	1143	1141	1145	3
38	0.96	1143	1141	1145	3
39	1.01	1143	1141	1145	3
40	1.06	1144	1142	1146	3
41	1.11	1144	1142	1146	3
42	1.16	1145	1143	1147	3
43	1.21	1145	1143	1147	3
44	1.27	1146	1144	1148	3
45	1.32	1146	1144	1148	3
46	1.37	1147	1145	1149	3
47	1.43	1147	1145	1149	3
48	1.49	1148	1146	1150	3
49	1.55	1148	1146	1150	3

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
50	1.61	1149	1147	1151	3
51	1.68	1149	1147	1151	3
52	1.74	1150	1148	1152	3
53	1.82	1150	1148	1152	3
54	1.89	1151	1149	1153	3
55	1.98	1151	1149	1153	3
56	2.07	1153	1151	1155	4
57	2.17	1153	1151	1155	4
58	2.28	1154	1152	1156	4
59	2.41	1156	1153	1159	4
60	2.56	1157	1154	1160	4
61	2.75	1158	1155	1161	4
62	3.01	1161	1157	1165	4
63	3.43	1164	1159	1169	4
64	4.00	1180	1171	1180	4

Table L-8. 2008–09 NECAP: Look-up Table—Reading Grade 3

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	300	300	310	1
1	-4.00	300	300	310	1
2	-4.00	300	300	310	1
3	-4.00	300	300	310	1
4	-4.00	300	300	310	1
5	-4.00	300	300	310	1
6	-4.00	300	300	310	1
7	-3.48	306	300	315	1
8	-3.11	310	303	317	1
9	-2.83	313	307	319	1
10	-2.61	316	310	322	1
11	-2.42	318	313	323	1
12	-2.26	320	315	325	1
13	-2.11	322	318	326	1
14	-1.98	323	319	327	1
15	-1.86	324	320	328	1
16	-1.74	326	322	330	1
17	-1.64	327	324	331	1
18	-1.54	328	325	331	1
19	-1.44	329	326	332	1
20	-1.35	330	327	333	1
21	-1.26	331	328	334	2
22	-1.17	332	329	335	2
23	-1.08	333	330	336	2
24	-0.99	334	331	337	2
25	-0.91	335	332	338	2
26	-0.82	336	333	339	2
27	-0.74	337	334	340	2

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
28	-0.65	338	335	341	2
29	-0.57	339	336	342	2
30	-0.48	340	337	343	3
31	-0.40	341	338	344	3
32	-0.31	342	339	345	3
33	-0.22	343	340	346	3
34	-0.13	344	341	347	3
35	-0.04	345	342	348	3
36	0.06	346	343	349	3
37	0.16	347	344	350	3
38	0.26	349	346	353	3
39	0.37	350	346	354	3
40	0.49	351	347	355	3
41	0.62	353	349	357	3
42	0.75	354	350	358	3
43	0.91	356	352	360	3
44	1.08	358	353	363	4
45	1.27	360	355	365	4
46	1.49	363	358	369	4
47	1.75	366	360	372	4
48	2.07	369	362	376	4
49	2.45	374	366	380	4
50	2.96	379	370	380	4
51	3.71	380	370	380	4
52	4.00	380	370	380	4

Table L-9. 2008–09 NECAP: Look-up Table—Reading Grade 4

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	400	400	410	1
1	-4.00	400	400	410	1
2	-4.00	400	400	410	1
3	-4.00	400	400	410	1
4	-4.00	400	400	410	1
5	-4.00	400	400	410	1
6	-4.00	400	400	410	1
7	-3.48	406	400	415	1
8	-3.11	410	403	417	1
9	-2.83	413	407	420	1
10	-2.60	415	409	421	1
11	-2.40	417	412	423	1
12	-2.23	419	414	424	1
13	-2.07	421	416	426	1
14	-1.93	422	418	427	1
15	-1.80	424	420	428	1
16	-1.67	425	421	429	1
17	-1.56	427	423	431	1

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
18	-1.45	428	424	432	1
19	-1.34	429	425	433	1
20	-1.24	430	427	434	1
21	-1.14	431	428	434	2
22	-1.05	432	429	435	2
23	-0.95	433	430	436	2
24	-0.86	434	431	437	2
25	-0.77	435	432	438	2
26	-0.68	436	433	439	2
27	-0.59	437	434	440	2
28	-0.49	438	435	441	2
29	-0.40	439	436	442	2
30	-0.30	440	437	443	3
31	-0.20	441	438	444	3
32	-0.10	442	439	445	3
33	0.00	443	440	447	3
34	0.11	445	442	449	3
35	0.22	446	442	450	3
36	0.33	447	443	451	3
37	0.45	448	444	452	3
38	0.58	450	446	454	3
39	0.71	451	447	455	3
40	0.85	453	449	457	3
41	1.00	454	450	459	3
42	1.17	456	451	461	4
43	1.34	458	453	463	4
44	1.54	460	455	465	4
45	1.75	462	456	468	4
46	2.00	465	459	471	4
47	2.28	468	461	475	4
48	2.61	472	465	479	4
49	3.01	476	468	480	4
50	3.55	480	470	480	4
51	4.00	480	470	480	4
52	4.00	480	470	480	4

Table L-10. 2008–09 NECAP: Look-up Table—Reading Grade 5

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	500	500	510	1
1	-4.00	500	500	510	1
2	-4.00	500	500	510	1
3	-4.00	500	500	510	1
4	-4.00	500	500	510	1
5	-4.00	500	500	510	1
6	-3.34	507	500	515	1
7	-2.94	512	506	519	1
8	-2.65	515	509	521	1
9	-2.42	518	513	523	1
10	-2.23	520	515	525	1
11	-2.05	522	518	527	1
12	-1.90	524	520	528	1
13	-1.76	525	521	529	1
14	-1.62	527	523	531	1
15	-1.50	528	524	532	1
16	-1.38	529	526	533	1
17	-1.27	531	528	534	2
18	-1.16	532	529	535	2
19	-1.06	533	530	536	2
20	-0.96	534	531	537	2
21	-0.86	535	532	538	2
22	-0.76	536	533	539	2
23	-0.66	537	534	540	2
24	-0.55	539	536	542	2
25	-0.45	539	536	542	2
26	-0.35	541	538	544	3
27	-0.25	542	539	545	3
28	-0.14	543	540	546	3
29	-0.03	544	541	547	3
30	0.08	546	543	550	3
31	0.20	547	544	551	3
32	0.32	548	544	552	3
33	0.45	550	546	554	3
34	0.59	551	547	555	3
35	0.73	553	549	557	3
36	0.88	555	551	559	3
37	1.03	555	551	559	3
38	1.20	558	554	562	4
39	1.37	560	555	565	4
40	1.55	562	557	567	4
41	1.74	564	559	569	4
42	1.95	567	562	572	4
43	2.16	569	564	574	4
44	2.38	571	566	576	4
45	2.62	574	569	580	4
46	2.87	577	571	580	4

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
47	3.15	580	574	580	4
48	3.46	580	574	580	4
49	3.81	580	573	580	4
50	4.00	580	573	580	4
51	4.00	580	573	580	4
52	4.00	580	573	580	4

Table L-11. 2008–09 NECAP: Look-up Table—Reading Grade 6

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	600	600	610	1
1	-4.00	600	600	610	1
2	-4.00	600	600	610	1
3	-4.00	600	600	610	1
4	-3.87	601	600	609	1
5	-3.47	606	600	613	1
6	-3.18	609	603	615	1
7	-2.95	612	607	617	1
8	-2.76	614	610	619	1
9	-2.60	616	612	620	1
10	-2.45	618	614	622	1
11	-2.32	619	615	623	1
12	-2.20	621	617	625	1
13	-2.09	622	619	626	1
14	-1.98	623	620	626	1
15	-1.88	624	621	627	1
16	-1.78	626	623	629	1
17	-1.68	627	624	630	1
18	-1.59	628	625	631	1
19	-1.49	628	625	631	1
20	-1.40	630	627	633	2
21	-1.31	631	628	634	2
22	-1.22	632	629	635	2
23	-1.13	633	630	636	2
24	-1.03	634	631	637	2
25	-0.94	635	632	638	2
26	-0.84	636	633	639	2
27	-0.75	637	634	640	2
28	-0.65	638	635	641	2
29	-0.55	639	636	642	2
30	-0.44	641	638	644	3
31	-0.33	642	639	645	3
32	-0.22	643	640	647	3
33	-0.10	645	641	649	3
34	0.02	646	642	650	3
35	0.15	648	644	652	3
36	0.29	649	645	653	3
37	0.44	651	647	655	3
38	0.59	653	649	657	3

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
39	0.75	655	651	659	3
40	0.91	656	652	660	3
41	1.08	658	654	662	3
42	1.26	660	656	665	4
43	1.45	663	658	668	4
44	1.64	665	660	670	4
45	1.85	667	662	672	4
46	2.06	670	665	675	4
47	2.30	672	667	677	4
48	2.55	675	670	680	4
49	2.84	679	673	680	4
50	3.21	680	674	680	4
51	3.78	680	671	680	4
52	4.00	680	670	680	4

Table L-12. 2008–09 NECAP: Look-up Table—Reading Grade 7

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	700	700	710	1
1	-4.00	700	700	710	1
2	-4.00	700	700	710	1
3	-4.00	700	700	710	1
4	-4.00	700	700	708	1
5	-3.85	702	700	709	1
6	-3.53	705	700	711	1
7	-3.28	708	703	713	1
8	-3.07	711	706	716	1
9	-2.88	713	708	718	1
10	-2.72	715	711	719	1
11	-2.57	716	712	720	1
12	-2.43	718	714	722	1
13	-2.30	720	716	724	1
14	-2.18	721	717	725	1
15	-2.06	722	718	726	1
16	-1.95	724	721	728	1
17	-1.84	725	722	729	1
18	-1.73	726	723	729	1
19	-1.63	727	724	730	1
20	-1.52	728	725	731	1
21	-1.42	730	727	733	2
22	-1.32	731	728	734	2
23	-1.22	732	729	735	2
24	-1.12	733	730	736	2
25	-1.01	734	731	737	2
26	-0.91	736	733	739	2
27	-0.80	737	734	740	2
28	-0.70	738	735	742	2
29	-0.59	739	736	743	2
30	-0.47	741	737	745	3

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
31	-0.35	742	738	746	3
32	-0.23	743	739	747	3
33	-0.10	745	741	749	3
34	0.03	746	742	750	3
35	0.17	748	744	752	3
36	0.32	750	746	754	3
37	0.47	751	747	755	3
38	0.62	753	749	757	3
39	0.78	755	751	759	3
40	0.94	757	753	761	3
41	1.11	759	755	763	3
42	1.28	761	757	765	4
43	1.45	763	759	768	4
44	1.64	765	761	770	4
45	1.83	767	762	772	4
46	2.03	769	764	774	4
47	2.24	772	767	777	4
48	2.48	775	770	780	4
49	2.77	778	772	780	4
50	3.14	780	773	780	4
51	3.78	780	770	780	4
52	4.00	780	770	780	4

Table L-13. 2008–09 NECAP: Look-up Table—Reading Grade 8

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
0	-4.00	800	800	810	1
1	-4.00	800	800	810	1
2	-4.00	800	800	810	1
3	-4.00	800	800	810	1
4	-4.00	800	800	810	1
5	-4.00	800	800	809	1
6	-3.97	800	800	808	1
7	-3.63	804	800	811	1
8	-3.37	807	801	813	1
9	-3.16	810	805	815	1
10	-2.97	812	807	817	1
11	-2.81	814	810	818	1
12	-2.66	815	811	819	1
13	-2.52	817	813	821	1
14	-2.39	818	814	822	1
15	-2.27	820	816	824	1
16	-2.15	821	817	825	1
17	-2.04	823	819	827	1
18	-1.92	824	820	828	1
19	-1.81	825	821	829	1
20	-1.70	826	822	830	1
21	-1.59	827	823	831	1
22	-1.48	829	825	833	2

(cont'd)

Raw Score	θ	Scaled Score	Error Band		Performance Level
			Lower Bound	Upper Bound	
23	-1.37	830	826	834	2
24	-1.26	831	828	835	2
25	-1.15	833	830	837	2
26	-1.04	834	830	838	2
27	-0.93	835	831	839	2
28	-0.82	837	833	841	2
29	-0.70	838	834	842	2
30	-0.58	839	835	843	2
31	-0.46	841	837	845	3
32	-0.33	842	838	846	3
33	-0.20	844	840	848	3
34	-0.06	845	841	849	3
35	0.09	847	843	851	3
36	0.24	849	845	853	3
37	0.40	851	847	855	3
38	0.56	852	848	856	3
39	0.73	854	850	858	3
40	0.91	856	852	860	3
41	1.09	858	854	863	3
42	1.28	861	856	866	4
43	1.48	863	858	868	4
44	1.68	865	860	870	4
45	1.89	868	863	873	4
46	2.12	870	865	875	4
47	2.35	873	868	878	4
48	2.62	876	871	880	4
49	2.92	880	874	880	4
50	3.32	880	873	880	4
51	3.99	880	870	880	4
52	4.00	880	870	880	4

Table L-14. 2008–09 NECAP: Look-up Table—Reading Grade 11

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	1100	1100	1110	1
1	-4.00	1100	1100	1110	1
2	-4.00	1100	1100	1110	1
3	-4.00	1100	1100	1110	1
4	-4.00	1100	1100	1110	1
5	-3.76	1103	1100	1111	1
6	-3.36	1107	1101	1113	1
7	-3.07	1110	1105	1115	1
8	-2.85	1112	1107	1117	1
9	-2.66	1115	1111	1119	1
10	-2.50	1116	1112	1120	1
11	-2.35	1118	1114	1122	1
12	-2.22	1119	1115	1123	1
13	-2.09	1121	1118	1124	1
14	-1.98	1122	1119	1125	1
15	-1.87	1123	1120	1126	1
16	-1.76	1124	1121	1127	1
17	-1.66	1125	1122	1128	1
18	-1.56	1126	1123	1129	1
19	-1.46	1128	1125	1131	1
20	-1.37	1129	1126	1132	1
21	-1.27	1129	1126	1132	1
22	-1.17	1131	1128	1134	2
23	-1.08	1132	1129	1135	2
24	-0.98	1133	1130	1136	2
25	-0.89	1134	1131	1137	2
26	-0.79	1135	1132	1138	2
27	-0.69	1136	1133	1139	2
28	-0.59	1137	1134	1140	2
29	-0.49	1138	1135	1141	2
30	-0.39	1139	1136	1142	2
31	-0.28	1140	1137	1143	3
32	-0.17	1142	1139	1145	3
33	-0.05	1143	1140	1146	3
34	0.07	1144	1141	1147	3
35	0.19	1145	1142	1148	3
36	0.31	1147	1144	1150	3
37	0.44	1148	1145	1151	3
38	0.57	1150	1147	1154	3
39	0.71	1151	1148	1155	3
40	0.85	1153	1150	1157	3
41	1.00	1153	1149	1157	3
42	1.15	1156	1152	1160	4
43	1.30	1158	1154	1162	4
44	1.47	1159	1155	1163	4
45	1.65	1161	1157	1165	4
46	1.84	1163	1159	1167	4
47	2.04	1165	1161	1169	4
48	2.27	1168	1164	1172	4
49	2.54	1171	1166	1176	4
50	2.89	1175	1169	1180	4
51	3.51	1180	1171	1180	4
52	4.00	1180	1170	1180	4

Table L-15. 2008–09 NECAP: Look-up Table—Writing Grade 5

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	500	500	510	1
1	-4.00	500	500	510	1
2	-4.00	500	500	510	1
3	-4.00	500	500	510	1
4	-4.00	500	500	510	1
5	-4.00	500	500	510	1
6	-4.00	500	500	510	1
7	-4.00	500	500	510	1
8	-3.93	501	500	510	1
9	-3.46	505	500	513	1
10	-3.09	509	502	516	1
11	-2.77	512	506	518	1
12	-2.48	515	509	521	1
13	-2.22	518	512	524	1
14	-1.97	520	515	526	1
15	-1.74	523	518	528	1
16	-1.51	525	520	530	1
17	-1.29	527	522	532	1
18	-1.06	530	525	535	2
19	-0.83	532	527	537	2
20	-0.59	534	529	540	2
21	-0.34	537	531	543	2
22	-0.06	539	533	545	2
23	0.23	543	537	549	3
24	0.55	546	539	553	3
25	0.89	549	542	556	3
26	1.24	553	546	560	3
27	1.62	557	550	565	4
28	2.01	560	552	568	4
29	2.41	565	557	573	4
30	2.84	569	561	577	4
31	3.29	573	565	580	4
32	3.76	578	570	580	4
33	4.00	580	572	580	4
34	4.00	580	572	580	4
35	4.00	580	572	580	4
36	4.00	580	572	580	4
37	4.00	580	572	580	4

Table L-16. 2008–09 NECAP: Look-up Table—Writing Grade 8

Raw Score	*θ*	*Scaled Score*	*Error Band*		*Performance Level*
			Lower Bound	*Upper Bound*	
0	-4.00	800	800	810	1
1	-4.00	800	800	810	1
2	-4.00	800	800	810	1
3	-4.00	800	800	810	1
4	-4.00	800	800	810	1
5	-4.00	800	800	810	1
6	-4.00	800	800	810	1
7	-3.94	801	800	810	1
8	-3.39	806	800	812	1
9	-3.03	810	805	816	1
10	-2.73	812	807	817	1
11	-2.48	815	810	820	1
12	-2.26	817	813	821	1
13	-2.05	819	815	823	1
14	-1.85	821	817	825	1
15	-1.67	823	819	827	1
16	-1.49	825	821	829	1
17	-1.31	826	822	830	1
18	-1.13	828	824	832	1
19	-0.95	830	826	834	2
20	-0.76	832	828	836	2
21	-0.56	834	830	838	2
22	-0.36	836	832	841	2
23	-0.14	838	833	843	2
24	0.08	839	834	844	2
25	0.31	842	837	847	3
26	0.54	844	839	849	3
27	0.79	847	842	852	3
28	1.04	849	844	854	3
29	1.32	852	847	858	3
30	1.60	855	849	861	3
31	1.91	858	852	864	4
32	2.23	861	855	867	4
33	2.57	864	858	870	4
34	2.97	868	861	875	4
35	3.48	873	865	880	4
36	4.00	878	868	880	4
37	4.00	880	870	880	4

Appendix M—Scaled Score Percentages and Cumulative Percentages

Table M-1. 2008–09 NECAP: Scaled Score Distribution—Mathematics Grade 3

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
300	0.3	0.3	341	4.8	38.1
301	0.0	0.3	342	2.5	40.7
302	0.0	0.3	343	5.3	46.0
303	0.0	0.3	344	3.1	49.0
304	0.2	0.5	345	2.7	51.7
305	0.0	0.5	346	3.0	54.7
306	0.0	0.5	347	6.3	61.0
307	0.2	0.7	348	3.3	64.3
308	0.0	0.7	349	3.2	67.5
309	0.0	0.7	350	3.4	70.9
310	0.3	0.9	351	3.3	74.2
311	0.0	0.9	352	6.4	80.6
312	0.0	0.9	353	0.0	80.6
313	0.3	1.3	354	3.2	83.8
314	0.0	1.3	355	3.1	86.8
315	0.4	1.6	356	0.0	86.8
316	0.4	2.0	357	2.7	89.5
317	0.0	2.0	358	2.7	92.2
318	0.5	2.5	359	0.0	92.2
319	0.6	3.1	360	2.3	94.6
320	0.0	3.1	361	2.0	96.6
321	0.6	3.6	362	0.0	96.6
322	0.7	4.4	363	0.0	96.6
323	0.8	5.1	364	1.5	98.1
324	0.8	5.9	365	0.0	98.1
325	0.8	6.6	366	0.0	98.1
326	0.8	7.4	367	1.1	99.2
327	1.0	8.4	368	0.0	99.2
328	1.0	9.5	369	0.0	99.2
329	1.1	10.5	370	0.0	99.2
330	1.1	11.7	371	0.6	99.7
331	2.5	14.1	372	0.0	99.7
332	1.3	15.4	373	0.0	99.7
333	1.4	16.9	374	0.0	99.7
334	1.6	18.4	375	0.0	99.7
335	3.2	21.7	376	0.0	99.7
336	1.8	23.5	377	0.2	100.0
337	1.7	25.2	378	0.0	100.0
338	3.9	29.1	379	0.0	100.0
339	4.2	33.3	380	0.0	100.0
340	0.0	33.3			

Table M-2. 2008–09 NECAP—Scaled Score Distribution—Mathematics Grade 4

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
400	0.3	0.3	441	2.4	35.3
401	0.2	0.5	442	5.0	40.4
402	0.0	0.5	443	2.7	43.0
403	0.0	0.5	444	2.8	45.9
404	0.0	0.5	445	2.9	48.8
405	0.0	0.5	446	5.9	54.6
406	0.2	0.7	447	3.0	57.6
407	0.0	0.7	448	3.2	60.9
408	0.0	0.7	449	3.2	64.1
409	0.3	1.0	450	3.5	67.6
410	0.0	1.0	451	3.4	71.0
411	0.3	1.3	452	3.4	74.4
412	0.0	1.3	453	0.0	74.4
413	0.4	1.7	454	3.4	77.8
414	0.0	1.7	455	3.4	81.1
415	0.4	2.2	456	3.1	84.3
416	0.0	2.2	457	3.1	87.3
417	0.5	2.6	458	0.0	87.3
418	0.5	3.1	459	2.8	90.1
419	0.0	3.1	460	0.0	90.1
420	0.4	3.5	461	2.8	92.9
421	0.6	4.1	462	0.0	92.9
422	0.5	4.6	463	2.3	95.2
423	0.7	5.3	464	0.0	95.2
424	0.7	6.0	465	1.8	96.9
425	0.7	6.7	466	0.0	96.9
426	0.8	7.5	467	0.0	96.9
427	0.9	8.4	468	1.4	98.4
428	0.9	9.2	469	0.0	98.4
429	1.8	11.0	470	0.0	98.4
430	2.1	13.2	471	0.0	98.4
431	0.0	13.2	472	0.9	99.3
432	1.2	14.3	473	0.0	99.3
433	2.6	16.9	474	0.0	99.3
434	1.5	18.4	475	0.0	99.3
435	1.4	19.8	476	0.0	99.3
436	3.4	23.2	477	0.0	99.3
437	1.9	25.1	478	0.0	99.3
438	1.9	26.9	479	0.0	99.3
439	4.0	30.9	480	0.7	100.0
440	2.1	33.0			

Table M-3. 2008–09 NECAP: Scaled Score Distribution—Mathematics Grade 5

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
500	0.6	0.6	541	4.4	38.6
501	0.0	0.6	542	2.2	40.7
502	0.0	0.6	543	4.3	45.0
503	0.0	0.6	544	2.4	47.5
504	0.0	0.6	545	4.4	51.9
505	0.0	0.6	546	2.4	54.3
506	0.0	0.6	547	4.6	58.9
507	0.0	0.6	548	2.3	61.2
508	0.4	1.0	549	4.8	66.0
509	0.0	1.0	550	2.3	68.3
510	0.0	1.0	551	4.9	73.3
511	0.0	1.0	552	2.4	75.7
512	0.0	1.0	553	4.7	80.4
513	0.5	1.5	554	2.2	82.6
514	0.0	1.5	555	2.2	84.8
515	0.0	1.5	556	2.2	87.0
516	0.6	2.1	557	2.1	89.1
517	0.0	2.1	558	1.9	91.0
518	0.0	2.1	559	1.8	92.8
519	0.8	3.0	560	1.7	94.5
520	0.0	3.0	561	0.0	94.5
521	0.9	3.8	562	1.5	96.0
522	1.0	4.8	563	1.1	97.0
523	0.0	4.8	564	0.0	97.0
524	1.1	5.9	565	1.0	98.1
525	1.2	7.1	566	0.0	98.1
526	0.0	7.1	567	0.8	98.9
527	1.3	8.4	568	0.0	98.9
528	1.2	9.6	569	0.6	99.5
529	1.3	11.0	570	0.0	99.5
530	1.5	12.5	571	0.0	99.5
531	1.5	14.0	572	0.0	99.5
532	3.2	17.3	573	0.3	99.8
533	0.0	17.3	574	0.0	99.8
534	1.6	18.9	575	0.0	99.8
535	1.8	20.7	576	0.0	99.8
536	3.7	24.4	577	0.0	99.8
537	1.9	26.3	578	0.0	99.8
538	1.9	28.2	579	0.0	99.8
539	4.0	32.2	580	0.2	100.0
540	2.0	34.2			

Table M-4. 2008–09 NECAP: Scaled Score Distribution—Mathematics Grade 6

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
600	1.0	1.0	641	4.6	43.3
601	0.0	1.0	642	2.5	45.7
602	0.0	1.0	643	4.5	50.2
603	0.0	1.0	644	2.2	52.4
604	0.0	1.0	645	4.5	56.9
605	0.0	1.0	646	2.2	59.1
606	0.0	1.0	647	4.5	63.6
607	0.0	1.0	648	2.1	65.7
608	0.0	1.0	649	4.2	70.0
609	0.7	1.7	650	2.1	72.0
610	0.0	1.7	651	4.0	76.0
611	0.0	1.7	652	3.9	79.9
612	0.0	1.7	653	0.0	79.9
613	0.0	1.7	654	3.6	83.5
614	0.9	2.6	655	1.7	85.2
615	0.0	2.6	656	1.7	86.9
616	0.0	2.6	657	1.7	88.6
617	0.0	2.6	658	1.7	90.3
618	1.1	3.6	659	1.5	91.8
619	0.0	3.6	660	1.4	93.2
620	0.0	3.6	661	1.2	94.4
621	1.3	4.9	662	1.2	95.5
622	0.0	4.9	663	1.0	96.5
623	1.4	6.3	664	0.9	97.4
624	0.0	6.3	665	0.0	97.4
625	1.5	7.9	666	0.7	98.1
626	1.6	9.5	667	0.6	98.8
627	0.0	9.5	668	0.0	98.8
628	1.6	11.1	669	0.6	99.3
629	1.8	12.9	670	0.0	99.3
630	1.8	14.7	671	0.0	99.3
631	0.0	14.7	672	0.3	99.6
632	4.0	18.7	673	0.0	99.6
633	0.0	18.7	674	0.0	99.6
634	2.1	20.8	675	0.2	99.8
635	4.4	25.2	676	0.0	99.8
636	2.2	27.4	677	0.0	99.8
637	2.2	29.6	678	0.0	99.8
638	2.2	31.8	679	0.0	99.8
639	4.7	36.5	680	0.2	100.0
640	2.2	38.7			

Table M-5. 2008–09 NECAP: Scaled Score Distribution—Mathematics Grade 7

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
700	0.7	0.7	741	2.4	44.4
701	0.0	0.7	742	4.7	49.1
702	0.0	0.7	743	4.4	53.5
703	0.0	0.7	744	2.2	55.7
704	0.0	0.7	745	4.3	60.0
705	0.0	0.7	746	4.1	64.1
706	0.0	0.7	747	2.0	66.1
707	0.5	1.2	748	3.9	70.0
708	0.0	1.2	749	3.8	73.9
709	0.0	1.2	750	1.9	75.8
710	0.0	1.2	751	5.3	81.1
711	0.0	1.2	752	0.0	81.1
712	0.0	1.2	753	3.4	84.5
713	0.7	2.0	754	1.5	86.0
714	0.0	2.0	755	3.2	89.2
715	0.0	2.0	756	1.4	90.7
716	0.0	2.0	757	1.5	92.1
717	1.0	3.0	758	1.2	93.4
718	0.0	3.0	759	1.2	94.6
719	0.0	3.0	760	1.0	95.6
720	1.3	4.3	761	0.9	96.5
721	0.0	4.3	762	0.8	97.4
722	0.0	4.3	763	0.7	98.1
723	1.5	5.8	764	0.0	98.1
724	1.6	7.4	765	0.6	98.6
725	0.0	7.4	766	0.5	99.1
726	1.8	9.2	767	0.0	99.1
727	0.0	9.2	768	0.3	99.5
728	1.8	11.0	769	0.0	99.5
729	1.9	12.9	770	0.0	99.5
730	2.0	14.9	771	0.3	99.7
731	2.0	16.9	772	0.0	99.7
732	2.2	19.1	773	0.0	99.7
733	2.2	21.3	774	0.2	99.9
734	2.1	23.4	775	0.0	99.9
735	2.2	25.6	776	0.0	99.9
736	2.2	27.8	777	0.0	99.9
737	4.8	32.6	778	0.0	99.9
738	2.3	34.9	779	0.0	99.9
739	4.8	39.7	780	0.1	100.0
740	2.3	42.0			

Table M-6. 2008–09 NECAP: Scaled Score Distribution—Mathematics Grade 8

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
800	0.6	0.6	841	4.6	46.9
801	0.0	0.6	842	4.6	51.6
802	0.0	0.6	843	2.2	53.7
803	0.0	0.6	844	4.8	58.5
804	0.0	0.6	845	4.3	62.8
805	0.0	0.6	846	4.1	66.9
806	0.5	1.1	847	2.2	69.1
807	0.0	1.1	848	4.2	73.4
808	0.0	1.1	849	4.0	77.3
809	0.0	1.1	850	3.8	81.1
810	0.0	1.1	851	1.9	83.1
811	0.0	1.1	852	3.6	86.7
812	0.0	1.1	853	1.6	88.2
813	0.0	1.1	854	2.8	91.0
814	0.0	1.1	855	1.4	92.4
815	0.8	1.9	856	1.3	93.7
816	0.0	1.9	857	2.4	96.1
817	0.0	1.9	858	0.9	97.0
818	0.9	2.7	859	0.0	97.0
819	0.0	2.7	860	0.9	97.9
820	0.0	2.7	861	0.6	98.5
821	1.1	3.8	862	0.6	99.1
822	0.0	3.8	863	0.0	99.1
823	1.2	5.1	864	0.4	99.5
824	0.0	5.1	865	0.0	99.5
825	1.1	6.2	866	0.3	99.7
826	1.4	7.6	867	0.0	99.7
827	1.4	9.0	868	0.0	99.7
828	1.6	10.5	869	0.2	99.9
829	1.5	12.1	870	0.0	99.9
830	1.7	13.7	871	0.0	99.9
831	1.8	15.6	872	0.0	99.9
832	1.9	17.5	873	0.0	99.9
833	2.0	19.5	874	0.0	99.9
834	2.0	21.5	875	0.1	100.0
835	4.3	25.8	876	0.0	100.0
836	2.4	28.2	877	0.0	100.0
837	2.4	30.6	878	0.0	100.0
838	4.4	35.0	879	0.0	100.0
839	4.7	39.7	880	0.0	100.0
840	2.7	42.4			

Table M-7. 2008–09 NECAP: Scaled ScoreDistribution—Mathematics Grade 11

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
1100	1.8	1.8	1141	3.6	74.6
1101	0.0	1.8	1142	3.5	78.2
1102	0.0	1.8	1143	4.8	82.9
1103	0.0	1.8	1144	2.8	85.7
1104	0.0	1.8	1145	2.7	88.4
1105	0.0	1.8	1146	2.3	90.7
1106	0.0	1.8	1147	2.1	92.8
1107	0.0	1.8	1148	1.8	94.6
1108	1.4	3.2	1149	1.5	96.0
1109	0.0	3.2	1150	1.4	97.4
1110	0.0	3.2	1151	1.0	98.5
1111	0.0	3.2	1152	0.0	98.5
1112	0.0	3.2	1153	0.8	99.3
1113	0.0	3.2	1154	0.2	99.5
1114	0.0	3.2	1155	0.0	99.5
1115	0.0	3.2	1156	0.2	99.6
1116	2.0	5.2	1157	0.2	99.8
1117	0.0	5.2	1158	0.1	99.9
1118	0.0	5.2	1159	0.0	99.9
1119	0.0	5.2	1160	0.0	99.9
1120	2.5	7.6	1161	0.1	100.0
1121	0.0	7.6	1162	0.0	100.0
1122	2.9	10.5	1163	0.0	100.0
1123	0.0	10.5	1164	0.0	100.0
1124	3.0	13.6	1165	0.0	100.0
1125	0.0	13.6	1166	0.0	100.0
1126	3.2	16.7	1167	0.0	100.0
1127	3.2	19.9	1168	0.0	100.0
1128	3.2	23.1	1169	0.0	100.0
1129	3.1	26.2	1170	0.0	100.0
1130	3.0	29.2	1171	0.0	100.0
1131	2.8	32.0	1172	0.0	100.0
1132	5.5	37.5	1173	0.0	100.0
1133	2.6	40.2	1174	0.0	100.0
1134	5.1	45.2	1175	0.0	100.0
1135	4.7	49.9	1176	0.0	100.0
1136	2.5	52.4	1177	0.0	100.0
1137	4.4	56.8	1178	0.0	100.0
1138	6.5	63.3	1179	0.0	100.0
1139	5.9	69.2	1180	0.0	100.0
1140	1.9	71.0			

Table M-8. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 3

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
300	0.4	0.4	341	2.7	31.1
301	0.0	0.4	342	2.7	33.8
302	0.0	0.4	343	3.0	36.8
303	0.0	0.4	344	3.1	39.9
304	0.0	0.4	345	3.6	43.5
305	0.0	0.4	346	3.7	47.3
306	0.2	0.6	347	3.8	51.1
307	0.0	0.6	348	0.0	51.1
308	0.0	0.6	349	4.2	55.3
309	0.0	0.6	350	4.3	59.6
310	0.3	0.9	351	4.6	64.2
311	0.0	0.9	352	0.0	64.2
312	0.0	0.9	353	4.7	68.9
313	0.5	1.4	354	5.2	74.1
314	0.0	1.4	355	0.0	74.1
315	0.0	1.4	356	5.0	79.1
316	0.6	2.0	357	0.0	79.1
317	0.0	2.0	358	5.0	84.1
318	0.7	2.7	359	0.0	84.1
319	0.0	2.7	360	4.6	88.7
320	0.7	3.4	361	0.0	88.7
321	0.0	3.4	362	0.0	88.7
322	0.8	4.1	363	4.1	92.8
323	0.8	4.9	364	0.0	92.8
324	0.9	5.8	365	0.0	92.8
325	0.0	5.8	366	3.0	95.8
326	0.9	6.7	367	0.0	95.8
327	0.9	7.6	368	0.0	95.8
328	1.0	8.6	369	2.0	97.9
329	1.1	9.7	370	0.0	97.9
330	1.1	10.8	371	0.0	97.9
331	1.2	12.1	372	0.0	97.9
332	1.3	13.4	373	0.0	97.9
333	1.5	14.9	374	1.3	99.1
334	1.5	16.4	375	0.0	99.1
335	1.7	18.0	376	0.0	99.1
336	1.9	19.9	377	0.0	99.1
337	1.9	21.9	378	0.0	99.1
338	2.0	23.9	379	0.6	99.8
339	2.2	26.1	380	0.2	100.0
340	2.3	28.4			

Table M-9. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 4

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
400	0.6	0.6	441	3.4	34.8
401	0.0	0.6	442	3.6	38.5
402	0.0	0.6	443	3.9	42.4
403	0.0	0.6	444	0.0	42.4
404	0.0	0.6	445	4.1	46.5
405	0.0	0.6	446	4.4	50.8
406	0.3	0.9	447	4.7	55.5
407	0.0	0.9	448	4.4	59.9
408	0.0	0.9	449	0.0	59.9
409	0.0	0.9	450	4.6	64.5
410	0.4	1.3	451	4.8	69.3
411	0.0	1.3	452	0.0	69.3
412	0.0	1.3	453	4.7	74.1
413	0.4	1.7	454	4.4	78.5
414	0.0	1.7	455	0.0	78.5
415	0.5	2.1	456	3.9	82.4
416	0.0	2.1	457	0.0	82.4
417	0.5	2.6	458	3.8	86.2
418	0.0	2.6	459	0.0	86.2
419	0.6	3.2	460	3.5	89.6
420	0.0	3.2	461	0.0	89.6
421	0.6	3.8	462	2.8	92.5
422	0.7	4.5	463	0.0	92.5
423	0.0	4.5	464	0.0	92.5
424	0.7	5.2	465	2.5	94.9
425	0.8	6.0	466	0.0	94.9
426	0.0	6.0	467	0.0	94.9
427	0.8	6.8	468	1.9	96.8
428	0.9	7.8	469	0.0	96.8
429	1.2	8.9	470	0.0	96.8
430	1.3	10.3	471	0.0	96.8
431	1.3	11.6	472	1.4	98.2
432	1.4	13.0	473	0.0	98.2
433	1.6	14.5	474	0.0	98.2
434	1.8	16.3	475	0.0	98.2
435	2.0	18.3	476	0.9	99.1
436	2.2	20.5	477	0.0	99.1
437	2.4	22.9	478	0.0	99.1
438	2.6	25.4	479	0.0	99.1
439	2.9	28.3	480	0.9	100.0
440	3.1	31.5			

Table M-10. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 5

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
500	0.4	0.4	541	3.6	32.0
501	0.0	0.4	542	3.8	35.8
502	0.0	0.4	543	4.3	40.1
503	0.0	0.4	544	4.7	44.8
504	0.0	0.4	545	0.0	44.8
505	0.0	0.4	546	4.9	49.7
506	0.0	0.4	547	5.1	54.8
507	0.2	0.6	548	5.2	60.1
508	0.0	0.6	549	0.0	60.1
509	0.0	0.6	550	5.4	65.5
510	0.0	0.6	551	5.2	70.7
511	0.0	0.6	552	0.0	70.7
512	0.4	1.0	553	4.7	75.5
513	0.0	1.0	554	0.0	75.5
514	0.0	1.0	555	8.7	84.1
515	0.4	1.5	556	0.0	84.1
516	0.0	1.5	557	0.0	84.1
517	0.0	1.5	558	3.5	87.6
518	0.5	2.0	559	0.0	87.6
519	0.0	2.0	560	3.1	90.7
520	0.7	2.6	561	0.0	90.7
521	0.0	2.6	562	2.5	93.2
522	0.7	3.3	563	0.0	93.2
523	0.0	3.3	564	2.0	95.2
524	0.8	4.1	565	0.0	95.2
525	0.9	5.0	566	0.0	95.2
526	0.0	5.0	567	1.5	96.8
527	1.0	6.1	568	0.0	96.8
528	1.1	7.2	569	1.1	97.9
529	1.3	8.5	570	0.0	97.9
530	0.0	8.5	571	0.7	98.6
531	1.4	9.9	572	0.0	98.6
532	1.5	11.4	573	0.0	98.6
533	1.9	13.3	574	0.6	99.2
534	1.9	15.2	575	0.0	99.2
535	2.1	17.3	576	0.0	99.2
536	2.3	19.6	577	0.4	99.5
537	2.6	22.3	578	0.0	99.5
538	0.0	22.3	579	0.0	99.5
539	6.1	28.3	580	0.5	100.0
540	0.0	28.3			

Table M-11. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 6

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
600	0.1	0.1	641	3.5	32.6
601	0.1	0.2	642	3.9	36.5
602	0.0	0.2	643	4.2	40.7
603	0.0	0.2	644	0.0	40.7
604	0.0	0.2	645	4.8	45.6
605	0.0	0.2	646	5.2	50.7
606	0.2	0.4	647	0.0	50.7
607	0.0	0.4	648	5.4	56.2
608	0.0	0.4	649	5.3	61.5
609	0.2	0.6	650	0.0	61.5
610	0.0	0.6	651	5.7	67.2
611	0.0	0.6	652	0.0	67.2
612	0.4	1.0	653	5.2	72.4
613	0.0	1.0	654	0.0	72.4
614	0.4	1.4	655	5.1	77.5
615	0.0	1.4	656	4.6	82.1
616	0.5	1.8	657	0.0	82.1
617	0.0	1.8	658	4.2	86.3
618	0.4	2.2	659	0.0	86.3
619	0.6	2.8	660	3.2	89.5
620	0.0	2.8	661	0.0	89.5
621	0.6	3.4	662	0.0	89.5
622	0.7	4.1	663	3.0	92.6
623	0.7	4.8	664	0.0	92.6
624	0.7	5.5	665	2.2	94.8
625	0.0	5.5	666	0.0	94.8
626	0.8	6.3	667	1.7	96.5
627	0.9	7.2	668	0.0	96.5
628	2.0	9.3	669	0.0	96.5
629	0.0	9.3	670	1.2	97.7
630	1.1	10.4	671	0.0	97.7
631	1.4	11.8	672	1.0	98.7
632	1.4	13.2	673	0.0	98.7
633	1.6	14.7	674	0.0	98.7
634	1.8	16.5	675	0.6	99.3
635	1.9	18.4	676	0.0	99.3
636	2.2	20.6	677	0.0	99.3
637	2.4	23.1	678	0.0	99.3
638	2.9	26.0	679	0.4	99.7
639	3.1	29.0	680	0.3	100.0
640	0.0	29.0			

Table M-12. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 7

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
700	0.1	0.1	741	3.6	28.8
701	0.0	0.1	742	3.9	32.7
702	0.1	0.2	743	4.5	37.2
703	0.0	0.2	744	0.0	37.2
704	0.0	0.2	745	4.7	41.9
705	0.1	0.3	746	5.3	47.1
706	0.0	0.3	747	0.0	47.1
707	0.0	0.3	748	5.1	52.3
708	0.1	0.4	749	0.0	52.3
709	0.0	0.4	750	5.6	57.8
710	0.0	0.4	751	5.8	63.7
711	0.2	0.6	752	0.0	63.7
712	0.0	0.6	753	5.5	69.2
713	0.2	0.8	754	0.0	69.2
714	0.0	0.8	755	5.1	74.3
715	0.3	1.1	756	0.0	74.3
716	0.3	1.4	757	4.6	78.9
717	0.0	1.4	758	0.0	78.9
718	0.4	1.8	759	4.2	83.1
719	0.0	1.8	760	0.0	83.1
720	0.4	2.2	761	3.7	86.8
721	0.5	2.8	762	0.0	86.8
722	0.6	3.3	763	3.2	90.0
723	0.0	3.3	764	0.0	90.0
724	0.7	4.0	765	2.6	92.6
725	0.8	4.8	766	0.0	92.6
726	0.8	5.6	767	2.3	94.8
727	0.9	6.5	768	0.0	94.8
728	1.0	7.5	769	1.7	96.5
729	0.0	7.5	770	0.0	96.5
730	1.1	8.6	771	0.0	96.5
731	1.2	9.8	772	1.4	97.9
732	1.4	11.2	773	0.0	97.9
733	1.6	12.8	774	0.0	97.9
734	1.8	14.6	775	0.9	98.8
735	0.0	14.6	776	0.0	98.8
736	2.1	16.7	777	0.0	98.8
737	2.5	19.2	778	0.6	99.5
738	2.9	22.0	779	0.0	99.5
739	3.2	25.2	780	0.5	100.0
740	0.0	25.2			

Table M-13. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 8

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
800	0.3	0.3	841	4.1	35.8
801	0.0	0.3	842	4.4	40.2
802	0.0	0.3	843	0.0	40.2
803	0.0	0.3	844	4.7	44.9
804	0.2	0.5	845	5.0	49.8
805	0.0	0.5	846	0.0	49.8
806	0.0	0.5	847	5.4	55.2
807	0.2	0.7	848	0.0	55.2
808	0.0	0.7	849	5.3	60.5
809	0.0	0.7	850	0.0	60.5
810	0.3	1.0	851	5.5	65.9
811	0.0	1.0	852	5.4	71.3
812	0.3	1.3	853	0.0	71.3
813	0.0	1.3	854	5.1	76.5
814	0.4	1.7	855	0.0	76.5
815	0.4	2.1	856	4.7	81.2
816	0.0	2.1	857	0.0	81.2
817	0.5	2.6	858	4.2	85.4
818	0.5	3.1	859	0.0	85.4
819	0.0	3.1	860	0.0	85.4
820	0.6	3.7	861	3.4	88.9
821	0.7	4.4	862	0.0	88.9
822	0.0	4.4	863	3.0	91.8
823	0.8	5.1	864	0.0	91.8
824	0.9	6.0	865	2.5	94.3
825	0.9	6.9	866	0.0	94.3
826	1.1	8.0	867	0.0	94.3
827	1.3	9.3	868	2.0	96.3
828	0.0	9.3	869	0.0	96.3
829	1.4	10.8	870	1.5	97.8
830	1.5	12.3	871	0.0	97.8
831	1.9	14.2	872	0.0	97.8
832	0.0	14.2	873	1.0	98.8
833	2.1	16.4	874	0.0	98.8
834	2.4	18.8	875	0.0	98.8
835	2.7	21.5	876	0.6	99.4
836	0.0	21.5	877	0.0	99.4
837	2.9	24.4	878	0.0	99.4
838	3.5	27.9	879	0.0	99.4
839	3.8	31.7	880	0.6	100.0
840	0.0	31.7			

Table M-14. 2008–09 NECAP: Scaled Score Distribution—Reading Grade 11

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
1100	0.2	0.2	1141	0.0	32.1
1101	0.0	0.2	1142	3.8	36.0
1102	0.0	0.2	1143	4.2	40.2
1103	0.1	0.3	1144	4.7	44.9
1104	0.0	0.3	1145	5.0	49.9
1105	0.0	0.3	1146	0.0	49.9
1106	0.0	0.3	1147	5.1	55.0
1107	0.2	0.5	1148	5.5	60.5
1108	0.0	0.5	1149	0.0	60.5
1109	0.0	0.5	1150	5.3	65.9
1110	0.3	0.7	1151	5.3	71.2
1111	0.0	0.7	1152	0.0	71.2
1112	0.3	1.1	1153	9.9	81.1
1113	0.0	1.1	1154	0.0	81.1
1114	0.0	1.1	1155	0.0	81.1
1115	0.4	1.5	1156	4.0	85.1
1116	0.5	1.9	1157	0.0	85.1
1117	0.0	1.9	1158	3.8	88.9
1118	0.5	2.4	1159	3.1	92.0
1119	0.4	2.8	1160	0.0	92.0
1120	0.0	2.8	1161	2.7	94.7
1121	0.6	3.4	1162	0.0	94.7
1122	0.6	4.0	1163	2.0	96.7
1123	0.7	4.7	1164	0.0	96.7
1124	0.7	5.4	1165	1.4	98.0
1125	0.9	6.3	1166	0.0	98.0
1126	0.9	7.2	1167	0.0	98.0
1127	0.0	7.2	1168	0.9	99.0
1128	0.9	8.1	1169	0.0	99.0
1129	2.1	10.2	1170	0.0	99.0
1130	0.0	10.2	1171	0.6	99.6
1131	1.3	11.5	1172	0.0	99.6
1132	1.4	12.9	1173	0.0	99.6
1133	1.6	14.5	1174	0.0	99.6
1134	1.7	16.2	1175	0.3	99.8
1135	1.9	18.1	1176	0.0	99.8
1136	2.4	20.5	1177	0.0	99.8
1137	2.4	22.9	1178	0.0	99.8
1138	2.8	25.7	1179	0.0	99.8
1139	3.1	28.8	1180	0.2	100.0
1140	3.3	32.1			

Table M-15. 2008–09 NECAP: Scaled Score Distribution—Writing Grade 5

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
500	4.6	4.6	541	0.0	100.0
501	2.9	7.5	542	0.0	100.0
502	0.0	7.5	543	0.0	100.0
503	0.0	7.5	544	0.0	100.0
504	0.0	7.5	545	0.0	100.0
505	4.4	11.9	546	0.0	100.0
506	0.0	11.9	547	0.0	100.0
507	0.0	11.9	548	0.0	100.0
508	0.0	11.9	549	0.0	100.0
509	6.4	18.3	550	0.0	100.0
510	0.0	18.3	551	0.0	100.0
511	0.0	18.3	552	0.0	100.0
512	8.5	26.8	553	0.0	100.0
513	0.0	26.8	554	0.0	100.0
514	0.0	26.8	555	0.0	100.0
515	11.6	38.5	556	0.0	100.0
516	0.0	38.5	557	0.0	100.0
517	0.0	38.5	558	0.0	100.0
518	16.0	54.5	559	0.0	100.0
519	0.0	54.5	560	0.0	100.0
520	18.7	73.2	561	0.0	100.0
521	0.0	73.2	562	0.0	100.0
522	0.0	73.2	563	0.0	100.0
523	17.2	90.4	564	0.0	100.0
524	0.0	90.4	565	0.0	100.0
525	7.9	98.3	566	0.0	100.0
526	0.0	98.3	567	0.0	100.0
527	1.7	100.0	568	0.0	100.0
528	0.0	100.0	569	0.0	100.0
529	0.0	100.0	570	0.0	100.0
530	0.0	100.0	571	0.0	100.0
531	0.0	100.0	572	0.0	100.0
532	0.0	100.0	573	0.0	100.0
533	0.0	100.0	574	0.0	100.0
534	0.0	100.0	575	0.0	100.0
535	0.0	100.0	576	0.0	100.0
536	0.0	100.0	577	0.0	100.0
537	0.0	100.0	578	0.0	100.0
538	0.0	100.0	579	0.0	100.0
539	0.0	100.0	580	0.0	100.0
540	0.0	100.0			

Table M-16. 2008–09 NECAP: Scaled Score Distribution—Writing Grade 8

Scale Score	*Percentage*	*Cumulative Percentage*	*Scale Score*	*Percentage*	*Cumulative Percentage*
800	2.4	2.4	841	0.0	100.0
801	1.5	3.8	842	0.0	100.0
802	0.0	3.8	843	0.0	100.0
803	0.0	3.8	844	0.0	100.0
804	0.0	3.8	845	0.0	100.0
805	0.0	3.8	846	0.0	100.0
806	2.5	6.3	847	0.0	100.0
807	0.0	6.3	848	0.0	100.0
808	0.0	6.3	849	0.0	100.0
809	0.0	6.3	850	0.0	100.0
810	3.9	10.2	851	0.0	100.0
811	0.0	10.2	852	0.0	100.0
812	5.9	16.1	853	0.0	100.0
813	0.0	16.1	854	0.0	100.0
814	0.0	16.1	855	0.0	100.0
815	8.9	25.0	856	0.0	100.0
816	0.0	25.0	857	0.0	100.0
817	12.0	37.0	858	0.0	100.0
818	0.0	37.0	859	0.0	100.0
819	15.3	52.3	860	0.0	100.0
820	0.0	52.3	861	0.0	100.0
821	17.3	69.6	862	0.0	100.0
822	0.0	69.6	863	0.0	100.0
823	16.6	86.2	864	0.0	100.0
824	0.0	86.2	865	0.0	100.0
825	10.7	96.9	866	0.0	100.0
826	3.1	100.0	867	0.0	100.0
827	0.0	100.0	868	0.0	100.0
828	0.0	100.0	869	0.0	100.0
829	0.0	100.0	870	0.0	100.0
830	0.0	100.0	871	0.0	100.0
831	0.0	100.0	872	0.0	100.0
832	0.0	100.0	873	0.0	100.0
833	0.0	100.0	874	0.0	100.0
834	0.0	100.0	875	0.0	100.0
835	0.0	100.0	876	0.0	100.0
836	0.0	100.0	877	0.0	100.0
837	0.0	100.0	878	0.0	100.0
838	0.0	100.0	879	0.0	100.0
839	0.0	100.0	880	0.0	100.0
840	0.0	100.0			

Appendix N—Detailed Alpha Coefficient Results

Table N-1. 2008–09 NECAP: Subgroup Reliability by Grade and Content

Grade	*Content*	*Subgroup*	*N*	*(α)*
3	Mathematics	White	25823	0.92
		Native Hawaiian or Pacific Islander	11	0.75
		Hispanic or Latino	2339	0.93
		Black or African American	1239	0.93
		Asian	776	0.93
		American Indian or Alaskan Native	123	0.94
		LEP	1408	0.94
		IEP	4171	0.94
		Low SES	9163	0.93
	Reading	White	25820	0.89
		Native Hawaiian or Pacific Islander	11	0.64
		Hispanic or Latino	2271	0.89
		Black or African American	1221	0.89
		Asian	766	0.88
		American Indian or Alaskan Native	122	0.89
		LEP	1301	0.90
		IEP	4170	0.90
		Low SES	9113	0.90
4	Mathematics	White	26940	0.92
		Native Hawaiian or Pacific Islander	10	0.95
		Hispanic or Latino	2787	0.92
		Black or African American	1401	0.93
		Asian	782	0.94
		American Indian or Alaskan Native	230	0.93
		LEP	1524	0.93
		IEP	4724	0.93
		Low SES	10004	0.93
	Reading	White	26935	0.86
		Native Hawaiian or Pacific Islander	10	0.83
		Hispanic or Latino	2717	0.88
		Black or African American	1389	0.88
		Asian	762	0.85
		American Indian or Alaskan Native	231	0.88
		LEP	1408	0.88
		IEP	4724	0.88
		Low SES	9941	0.88
5	Mathematics	White	27352	0.91
		Native Hawaiian or Pacific Islander	11	0.86
		Hispanic or Latino	2518	0.89
		Black or African American	1370	0.90
		Asian	836	0.92
		American Indian or Alaskan Native	214	0.91
		LEP	1356	0.91
		IEP	5289	0.90
		Low SES	9638	0.90
	Reading	White	27353	0.88
		Native Hawaiian or Pacific Islander	11	0.65
		Hispanic or Latino	2467	0.87
		Black or African American	1354	0.88
		Asian	818	0.87

(cont'd)

Grade	Content	Subgroup	N	(α)
5	Reading	American Indian or Alaskan Native	213	0.88
		LEP	1265	0.88
		IEP	5288	0.88
		Low SES	9596	0.88
	Writing	White	27290	0.74
		Native Hawaiian or Pacific Islander	11	0.53
		Hispanic or Latino	2465	0.76
		Black or African American	1347	0.76
		Asian	819	0.73
		American Indian or Alaskan Native	213	0.76
		LEP	1263	0.78
		IEP	5253	0.77
		Low SES	9553	0.76
6	Mathematics	White	27921	0.92
		Native Hawaiian or Pacific Islander	9	0.95
		Hispanic or Latino	2476	0.91
		Black or African American	1374	0.91
		Asian	794	0.93
		American Indian or Alaskan Native	222	0.93
		LEP	1196	0.91
		IEP	5377	0.89
		Low SES	9596	0.91
	Reading	White	27921	0.87
		Native Hawaiian or Pacific Islander	9	0.92
		Hispanic or Latino	2421	0.87
		Black or African American	1358	0.88
		Asian	786	0.87
		American Indian or Alaskan Native	223	0.91
		LEP	1100	0.87
		IEP	5388	0.87
		Low SES	9550	0.88
7	Mathematics	White	28954	0.92
		Native Hawaiian or Pacific Islander	10	0.89
		Hispanic or Latino	2542	0.89
		Black or African American	1413	0.90
		Asian	753	0.93
		American Indian or Alaskan Native	150	0.89
		LEP	1002	0.91
		IEP	5709	0.89
		Low SES	9699	0.90
	Reading	White	28972	0.88
		Native Hawaiian or Pacific Islander	10	0.78
		Hispanic or Latino	2486	0.88
		Black or African American	1398	0.88
		Asian	734	0.88
		American Indian or Alaskan Native	150	0.91
		LEP	901	0.87
		IEP	5717	0.88
		Low SES	9658	0.88

(cont'd)

Grade	Content	Subgroup	N	(α)
8	Mathematics	White	29907	0.92
		Native Hawaiian or Pacific Islander	16	0.90
		Hispanic or Latino	2706	0.89
		Black or African American	1407	0.89
		Asian	790	0.93
		American Indian or Alaskan Native	131	0.92
		LEP	921	0.90
		IEP	5655	0.87
		Low SES	9521	0.90
	Reading	White	29901	0.89
		Native Hawaiian or Pacific Islander	16	0.84
		Hispanic or Latino	2667	0.90
		Black or African American	1406	0.91
		Asian	778	0.90
		American Indian or Alaskan Native	132	0.89
		LEP	840	0.91
		IEP	5673	0.90
		Low SES	9484	0.90
	Writing	White	29818	0.74
		Native Hawaiian or Pacific Islander	16	0.77
		Hispanic or Latino	2643	0.76
		Black or African American	1393	0.76
		Asian	777	0.75
		American Indian or Alaskan Native	131	0.75
		LEP	832	0.77
		IEP	5619	0.74
		Low SES	9422	0.75
11	Mathematics	White	29562	0.91
		Native Hawaiian or Pacific Islander	15	0.90
		Hispanic or Latino	2207	0.86
		Black or African American	1231	0.87
		Asian	669	0.93
		American Indian or Alaskan Native	148	0.88
		LEP	692	0.88
		IEP	4926	0.83
		Low SES	6762	0.88
	Reading	White	29691	0.89
		Native Hawaiian or Pacific Islander	15	0.63
		Hispanic or Latino	2171	0.87
		Black or African American	1231	0.89
		Asian	661	0.90
		American Indian or Alaskan Native	150	0.90
		LEP	639	0.85
		IEP	4970	0.88
		Low SES	6771	0.89

[1]Only subgroups with sample size ≥10 reported

Table N-2. 2008–09 NECAP: Reliability by Grade, Content, Item Type, and Form

Grade	*Content*	*Stat*	*Form1*	*Form2*	*Form3*	*Form4*	*Form5*	*Form6*	*Form7*	*Form8*	*Form9*
3	Mathematics	Form Alpha	0.94	0.94	0.95	0.95	0.94	0.94	0.94	0.94	0.94
		MC Alpha	0.63	0.60	0.68	0.71	0.61	0.67	0.62	0.59	0.67
		CR Alpha	0.47	0.55	0.56	0.55	0.52	0.51	0.42	0.49	0.49
		Item Type Alpha	0.98	0.98	0.99	0.99	0.98	0.98	0.99	0.99	0.99
		Common Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
		Matrix Alpha	0.71	0.74	0.76	0.77	0.72	0.74	0.70	0.72	0.75
		Admin Type Alpha	0.94	0.94	0.94	0.94	0.93	0.94	0.94	0.94	0.94
	Reading	Form Alpha	0.92	0.93	0.93	0.90	0.88	0.91	0.90	0.90	0.91
		MC Alpha	0.76	0.79	0.81	0.77	0.75	0.80	0.73	0.79	0.79
		CR Alpha	0.61	0.62	0.61	0.04	0.09	0.07	0.07	0.10	0.18
		Item Type Alpha	0.98	0.98	0.98	0.93	0.91	0.97	0.96	0.94	0.96
		Common Alpha	0.89	0.89	0.89	0.88	0.88	0.88	0.88	0.89	0.89
		Matrix Alpha	0.79	0.82	0.82	0.61	0.55	0.74	0.66	0.65	0.70
		Admin Type Alpha	0.92	0.93	0.93	0.81	0.79	0.85	0.84	0.83	0.84
4	Mathematics	Form Alpha	0.94	0.94	0.94	0.95	0.94	0.94	0.95	0.94	0.94
		MC Alpha	0.64	0.66	0.51	0.67	0.61	0.62	0.63	0.64	0.50
		CR Alpha	0.51	0.47	0.45	0.50	0.44	0.45	0.51	0.46	0.48
		Item Type Alpha	0.99	0.98	0.98	0.99	0.99	0.98	0.99	0.99	0.99
		Common Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.94	0.93	0.93
		Matrix Alpha	0.73	0.72	0.65	0.75	0.68	0.69	0.74	0.72	0.67
		Admin Type Alpha	0.94	0.94	0.93	0.94	0.94	0.93	0.94	0.94	0.94
	Reading	Form Alpha	0.91	0.90	0.91	0.87	0.87	0.89	0.89	0.85	0.88
		MC Alpha	0.78	0.76	0.76	0.77	0.76	0.75	0.77	0.75	0.74
		CR Alpha	0.57	0.55	0.62	0.14	0.06	0.11	0.10	0.02	0.15
		Item Type Alpha	0.98	0.98	0.97	0.90	0.90	0.95	0.96	0.89	0.95
		Common Alpha	0.86	0.85	0.85	0.85	0.85	0.86	0.86	0.85	0.86
		Matrix Alpha	0.80	0.79	0.79	0.56	0.54	0.65	0.70	0.45	0.63
		Admin Type Alpha	0.91	0.90	0.91	0.76	0.75	0.80	0.82	0.74	0.81
5	Mathematics	Form Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
		MC Alpha	0.60	0.59	0.60	0.65	0.63	0.59	0.56	0.58	0.61
		OR Alpha	0.63	0.69	0.56	0.46	0.56	0.61	0.59	0.61	0.52
		Item Type Alpha	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.99
		Common Alpha	0.92	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
		Matrix Alpha	0.74	0.78	0.71	0.68	0.72	0.73	0.73	0.75	0.72
		Admin Type Alpha	0.92	0.93	0.92	0.92	0.92	0.92	0.93	0.92	0.93

(cont'd)

Grade	Content	Stat	Form1	Form2	Form3	Form4	Form5	Form6	Form7	Form8	Form9
5	Reading	Form Alpha	0.92	0.92	0.92	0.91	0.90	0.90	0.89	0.89	0.89
		MC Alpha	0.79	0.76	0.75	0.81	0.74	0.78	0.68	0.75	0.69
		CR Alpha	0.75	0.74	0.70	0.19	0.21	0.11	0.09	0.10	0.19
		Item Type Alpha	0.98	0.98	0.98	0.94	0.94	0.95	0.94	0.93	0.94
		Common Alpha	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.87	0.87
		Matrix Alpha	0.82	0.80	0.79	0.69	0.64	0.69	0.59	0.64	0.61
		Admin Type Alpha	0.92	0.92	0.92	0.83	0.82	0.82	0.82	0.81	0.81
	Writing[1]	Form Alpha	0.72								
		MC Alpha	0.70								
		OR Alpha	0.59								
		Item Type Alpha	0.73								
		Common Alpha	0.72								
6	Mathematics	Form Alpha	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
		MC Alpha	0.62	0.63	0.63	0.62	0.65	0.70	0.61	0.62	0.63
		OR Alpha	0.63	0.66	0.55	0.66	0.56	0.70	0.52	0.63	0.40
		Item Type Alpha	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.99
		Common Alpha	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.92	0.93
		Matrix Alpha	0.77	0.77	0.74	0.77	0.73	0.81	0.73	0.77	0.71
		Admin Type Alpha	0.93	0.93	0.93	0.93	0.93	0.94	0.93	0.94	0.93
	Reading	Form Alpha	0.93	0.93	0.92	0.91	0.91	0.91	0.91	0.91	0.91
		MC Alpha	0.77	0.78	0.67	0.79	0.77	0.78	0.78	0.80	0.76
		CR Alpha	0.74	0.77	0.76	0.17	0.15	0.13	0.11	0.24	0.13
		Item Type Alpha	0.98	0.98	0.98	0.94	0.94	0.95	0.95	0.94	0.94
		Common Alpha	0.90	0.90	0.90	0.90	0.89	0.89	0.90	0.90	0.89
		Matrix Alpha	0.80	0.81	0.75	0.70	0.66	0.68	0.69	0.68	0.65
		Admin Type Alpha	0.93	0.93	0.92	0.84	0.84	0.84	0.84	0.83	0.83
7	Mathematics	Form Alpha	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
		MC Alpha	0.65	0.61	0.62	0.61	0.53	0.62	0.64	0.64	0.60
		OR Alpha	0.67	0.57	0.65	0.66	0.70	0.61	0.60	0.49	0.58
		Item Type Alpha	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.98	0.99
		Common Alpha	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
		Matrix Alpha	0.79	0.73	0.77	0.77	0.77	0.75	0.77	0.72	0.75
		Admin Type Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
	Reading	Form Alpha	0.92	0.92	0.92	0.90	0.89	0.89	0.90	0.89	0.90
		MC Alpha	0.76	0.74	0.75	0.77	0.66	0.69	0.75	0.75	0.76
		CR Alpha	0.79	0.77	0.77	0.21	0.24	0.09	0.07	0.19	0.19
		Item Type Alpha	0.98	0.98	0.98	0.93	0.92	0.93	0.94	0.92	0.92
		Common Alpha	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
		Matrix Alpha	0.81	0.79	0.79	0.67	0.59	0.62	0.65	0.60	0.66
		Admin Type Alpha	0.92	0.92	0.92	0.81	0.79	0.81	0.81	0.79	0.80

(cont'd)

Grade	Content	Stat	Form1	Form2	Form3	Form4	Form5	Form6	Form7	Form8	Form9
8	Mathematics	Form Alpha	0.95	0.94	0.94	0.94	0.94	0.95	0.94	0.94	0.94
		MC Alpha	0.57	0.65	0.54	0.62	0.65	0.62	0.56	0.64	0.54
		OR Alpha	0.71	0.67	0.64	0.60	0.58	0.73	0.66	0.59	0.59
		Item Type Alpha	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.99	0.98
		Common Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
		Matrix Alpha	0.79	0.79	0.74	0.76	0.76	0.81	0.77	0.77	0.72
		Admin Type Alpha	0.94	0.94	0.94	0.94	0.93	0.94	0.94	0.94	0.94
	Reading	Form Alpha	0.93	0.92	0.92	0.89	0.89	0.89	0.89	0.89	0.89
		MC Alpha	0.78	0.76	0.76	0.72	0.72	0.68	0.71	0.69	0.74
		CR Alpha	0.78	0.80	0.79	0.19	0.22	0.10	0.07	0.20	0.22
		Item Type Alpha	0.98	0.98	0.98	0.92	0.92	0.92	0.93	0.92	0.91
		Common Alpha	0.89	0.89	0.88	0.88	0.88	0.88	0.88	0.88	0.88
		Matrix Alpha	0.82	0.81	0.82	0.60	0.63	0.57	0.62	0.60	0.64
		Admin Type Alpha	0.93	0.92	0.92	0.80	0.79	0.80	0.80	0.79	0.78
	Writing[1]	Form Alpha	0.76								
		MC Alpha	0.63								
		OR Alpha	0.70								
		Item Type Alpha	0.78								
		Common Alpha	0.76								
11	Mathematics	Form Alpha	0.94	0.94	0.94	0.94	0.93	0.94	0.94	0.94	
		MC Alpha	0.61	0.51	0.59	0.54	0.37	0.55	0.63	0.49	
		OR Alpha	0.65	0.59	0.53	0.61	0.61	0.56	0.56	0.47	
		Item Type Alpha	0.98	0.98	0.99	0.99	0.98	0.99	0.99	0.99	
		Common Alpha	0.93	0.93	0.93	0.92	0.92	0.93	0.93	0.93	
		Matrix Alpha	0.78	0.72	0.72	0.74	0.67	0.73	0.75	0.67	
		Admin Type Alpha	0.93	0.93	0.93	0.93	0.93	0.93	0.94	0.93	
	Reading	Form Alpha	0.93	0.93	0.91	0.90	0.91	0.91	0.91	0.91	
		MC Alpha	0.78	0.79	0.78	0.72	0.79	0.78	0.77	0.75	
		CR Alpha	0.84	0.85	0.18	0.19	0.17	0.11	0.12	0.13	
		Item Type Alpha	0.98	0.98	0.94	0.94	0.94	0.95	0.95	0.95	
		Common Alpha	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	
		Matrix Alpha	0.84	0.84	0.69	0.65	0.70	0.71	0.70	0.70	
		Admin Type Alpha	0.93	0.93	0.83	0.82	0.83	0.84	0.84	0.84	

MC = multiple-choice; OR = open response; CR = constructed response; All = MC, OR, and CR All α = common and matrix items; MC α = MC items only; CR α = CR items only; OR = OR α items only; Frmt Strat = stratified by MC/OR; Com alpha = common items only1Writing tests had only one form

Table N-3. 2008–09 NECAP: Common Item Reporting Category Alpha Reliabilities, with Points Possible, by Grade and Content

Grade	*Content*	*Reporting Category*	*Points Possible*	α
3	Mathematics	Number & Operations	35	0.89
		Geometry & Measurement	10	0.62
		Functions & Algebra	10	0.62
		Data, Statistics, & Probability	10	0.74
	Reading	Word ID/Vocabulary	20	0.76
		Literary	16	0.71
		Informational	16	0.69
		Initial Understanding	19	0.78
		Analysis & Interpretation	13	0.56
4	Mathematics	Number & Operations	32	0.87
		Geometry & Measurement	13	0.69
		Functions & Algebra	10	0.76
		Data, Statistics, & Probability	10	0.72
	Reading	Word ID/Vocabulary	19	0.68
		Literary	16	0.65
		Informational	17	0.67
		Initial Understanding	19	0.68
		Analysis & Interpretation	14	0.63
5	Mathematics	Number & Operations	30	0.83
		Geometry & Measurement	13	0.63
		Functions & Algebra	13	0.63
		Data, Statistics, & Probability	10	0.66
	Reading	Word ID/Vocabulary	10	0.71
		Literary	21	0.75
		Informational	21	0.71
		Initial Understanding	18	0.73
		Analysis & Interpretation	24	0.73
	Writing	Structures of Language & Writing Conventions	10	0.70
		Short Responses	12	0.72
		Extended Responses	15	0.06
6	Mathematics	Number & Operations	26	0.86
		Geometry & Measurement	17	0.73
		Functions & Algebra	13	0.66
		Data, Statistics, & Probability	10	0.63
6	Reading	Word ID/Vocabulary	10	0.72
		Literary	21	0.79
		Informational	21	0.75
		Initial Understanding	19	0.76
		Analysis & Interpretation	23	0.78
7	Mathematics	Number & Operations	20	0.82
		Geometry & Measurement	16	0.66
		Functions & Algebra	20	0.80
		Data, Statistics, & Probability	10	0.61
	Reading	Word ID/Vocabulary	10	0.63
		Literary	21	0.75
		Informational	21	0.75
		Initial Understanding	19	0.72
		Analysis & Interpretation	23	0.78

(cont'd)

Grade	Content	Reporting Category	Points Possible	α
8	Mathematics	Number & Operations	13	0.69
		Geometry & Measurement	17	0.75
		Functions & Algebra	26	0.86
		Data, Statistics, & Probability	10	0.63
	Reading	Word ID/Vocabulary	10	0.64
		Literary	21	0.79
		Informational	21	0.71
		Initial Understanding	18	0.71
		Analysis & Interpretation	24	0.78
	Writing	Structures of Language & Writing Conventions	10	0.63
		Short Responses	12	0.76
		Extended Responses	15	0.24
11	Mathematics	Number & Operations	10	0.63
		Geometry & Measurement	19	0.79
		Functions & Algebra	26	0.85
		Data, Statistics, & Probability	9	0.65
	Reading	Word ID/Vocabulary	10	0.62
		Literary	21	0.77
		Informational	21	0.78
		Initial Understanding	18	0.74
		Analysis & Interpretation	24	0.80

Appendix O—Decision Accuracy and Consistency Results

Table O-1. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 3

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.113	0.020	0.000	0.000	0.133
PP	0.021	0.144	0.037	0.000	0.201
P	0.000	0.032	0.391	0.052	0.475
PWD	0.000	0.000	0.027	0.164	0.191
Total	0.134	0.195	0.455	0.216	1.000
Overall Accuracy (sum of diagonal) = 0.812					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-2. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 3

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.106	0.028	0.001	0.000	0.134
PP	0.028	0.120	0.047	0.000	0.195
P	0.001	0.047	0.353	0.055	0.455
PWD	0.000	0.000	0.055	0.162	0.216
Total	0.134	0.195	0.455	0.216	1.000
Overall Consistency (sum of diagonal) = 0.741					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-3. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 3

Accuracy	0.812
Consistency	0.740
Kappa (k)	0.623

Table O-4. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 3

Achievement Level	*Accuracy*	*Consistency*
SBP	0.851	0.789
PP	0.713	0.616
P	0.824	0.775
PWD	0.859	0.748

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-5. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 3

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.960	0.020	0.021	0.943
PP:P	0.931	0.038	0.032	0.904
P:PWD	0.921	0.052	0.027	0.891

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-6. 2008–09 NECAP: Decision Accuracy Cross Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 4

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.105	0.019	0.000	0.000	0.124
PP	0.020	0.147	0.037	0.000	0.204
P	0.000	0.031	0.375	0.044	0.450
PWD	0.000	0.000	0.022	0.201	0.223
Total	0.125	0.197	0.434	0.245	1.000
Overall Accuracy (sum of diagonal) = 0.828					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-7. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 4

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.098	0.027	0.001	0.000	0.125
PP	0.027	0.124	0.046	0.000	0.197
P	0.001	0.046	0.341	0.046	0.434
PWD	0.000	0.000	0.046	0.199	0.245
Total	0.125	0.197	0.434	0.245	1.000
Overall Consistency (sum of diagonal) = 0.762					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-8. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 4

Accuracy	0.828
Consistency	0.762
Kappa (k)	0.658

Table O-9. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 4

Achievement Level	*Accuracy*	*Consistency*
SBP	0.848	0.782
PP	0.723	0.630
P	0.834	0.787
PWD	0.901	0.813

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-10. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 4

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.961	0.019	0.020	0.946
PP:P	0.933	0.037	0.031	0.906
P:PWD	0.934	0.044	0.022	0.909

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-11. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 5

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.134	0.027	0.001	0.000	0.162
PP	0.026	0.089	0.039	0.000	0.155
P	0.001	0.032	0.415	0.046	0.494
PWD	0.000	0.000	0.023	0.166	0.189
Total	0.161	0.149	0.478	0.212	1.000
Overall Accuracy (sum of diagonal) = 0.804					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-12. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 5

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.123	0.034	0.005	0.000	0.161
PP	0.034	0.069	0.046	0.000	0.149
P	0.005	0.046	0.379	0.048	0.478
PWD	0.000	0.000	0.048	0.164	0.212
Total	0.161	0.149	0.478	0.212	1.000
Overall Consistency (sum of diagonal) = 0.735					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-13. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 5

Accuracy	0.804
Consistency	0.734
Kappa (k)	0.608

Table O-14. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 5

Achievement Level	*Accuracy*	*Consistency*
SBP	0.826	0.760
PP	0.577	0.464
P	0.839	0.792
PWD	0.877	0.773

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-15. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 5

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.945	0.028	0.027	0.923
PP:P	0.927	0.040	0.033	0.898
P:PWD	0.931	0.046	0.023	0.904

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-16. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 6

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.152	0.032	0.001	0.000	0.185
PP	0.029	0.110	0.041	0.000	0.179
P	0.001	0.032	0.369	0.038	0.440
PWD	0.000	0.000	0.021	0.175	0.196
Total	0.181	0.174	0.432	0.213	1.000
Overall Accuracy (sum of diagonal) = 0.806					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-17. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 6

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.138	0.039	0.004	0.000	0.181
PP	0.039	0.087	0.048	0.000	0.174
P	0.004	0.048	0.339	0.041	0.432
PWD	0.000	0.000	0.041	0.172	0.213
Total	0.181	0.174	0.432	0.213	1.000
Overall Consistency (sum of diagonal) = 0.736					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-18. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 6

Accuracy	0.806
Consistency	0.736
Kappa (k)	0.625

Table O-19. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 6

Achievement Level	*Accuracy*	*Consistency*
SBP	0.823	0.761
PP	0.615	0.501
P	0.839	0.785
PWD	0.892	0.806

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-20. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 6

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.938	0.033	0.029	0.914
PP:P	0.926	0.041	0.033	0.897
P:PWD	0.941	0.038	0.021	0.917

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-21. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 7

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.172	0.035	0.001	0.000	0.208
PP	0.030	0.113	0.042	0.000	0.186
P	0.001	0.033	0.358	0.036	0.428
PWD	0.000	0.000	0.019	0.160	0.179
Total	0.203	0.180	0.421	0.196	1.000
Overall Accuracy (sum of diagonal) = 0.803					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-22. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 7

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.157	0.042	0.005	0.000	0.203
PP	0.042	0.089	0.049	0.000	0.180
P	0.005	0.049	0.329	0.039	0.421
PWD	0.000	0.000	0.039	0.157	0.196
Total	0.203	0.180	0.421	0.196	1.000
Overall Consistency (sum of diagonal) = 0.732					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-23. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 7

Accuracy	0.803
Consistency	0.732
Kappa (k)	0.624

Table O-24. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 7

Achievement Level	*Accuracy*	*Consistency*
SBP	0.830	0.772
PP	0.609	0.496
P	0.838	0.782
PWD	0.893	0.803

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-25. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 7

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.934	0.035	0.031	0.907
PP:P	0.924	0.043	0.033	0.893
P:PWD	0.945	0.036	0.019	0.923

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-26. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 8

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.162	0.032	0.000	0.000	0.194
PP	0.029	0.131	0.042	0.000	0.202
P	0.000	0.033	0.371	0.037	0.441
PWD	0.000	0.000	0.019	0.144	0.163
Total	0.191	0.196	0.433	0.181	1.000
Overall Accuracy (sum of diagonal) = 0.808					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-27. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 8

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.148	0.040	0.003	0.000	0.191
PP	0.040	0.106	0.050	0.000	0.196
P	0.003	0.050	0.341	0.038	0.433
PWD	0.000	0.000	0.038	0.142	0.181
Total	0.191	0.196	0.433	0.181	1.000
Overall Consistency (sum of diagonal) = 0.737					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-28. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 8

Accuracy	0.808
Consistency	0.737
Kappa (k)	0.627

Table O-29. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 8

Achievement Level	*Accuracy*	*Consistency*
SBP	0.834	0.775
PP	0.648	0.539
P	0.842	0.789
PWD	0.885	0.787

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-30. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 8

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.939	0.032	0.029	0.914
PP:P	0.924	0.043	0.033	0.894
P:PWD	0.945	0.037	0.019	0.923

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-31. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Mathematics Grade 11

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.350	0.047	0.000	0.000	0.397
PP	0.034	0.227	0.040	0.000	0.301
P	0.000	0.025	0.257	0.010	0.292
PWD	0.000	0.000	0.002	0.008	0.011
Total	0.385	0.298	0.299	0.018	1.000
Overall Accuracy (sum of diagonal) = 0.842					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-32. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Mathematics Grade 11

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.328	0.056	0.001	0.000	0.385
PP	0.056	0.197	0.045	0.000	0.298
P	0.001	0.045	0.245	0.008	0.299
PWD	0.000	0.000	0.008	0.010	0.018
Total	0.385	0.298	0.299	0.018	1.000
Overall Consistency (sum of diagonal) = 0.780					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-33. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Mathematics Grade 11

Accuracy	0.842
Consistency	0.780
Kappa (k)	0.674

Table O-34. 2008–09 NECAP: Indices Conditional On Achievement Level—Mathematics Grade 11

Achievement Level	*Accuracy*	*Consistency*
SBP	0.882	0.853
PP	0.754	0.661
P	0.881	0.820
PWD	0.782	0.535

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-35. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Mathematics Grade 11

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.919	0.047	0.034	0.887
PP:P	0.935	0.040	0.025	0.909
P:PWD	0.988	0.010	0.002	0.983

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-36. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 3

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.078	0.017	0.000	0.000	0.095
PP	0.019	0.112	0.037	0.000	0.167
P	0.000	0.033	0.433	0.079	0.545
PWD	0.000	0.000	0.042	0.151	0.193
Total	0.097	0.161	0.512	0.230	1.000
Overall Accuracy (sum of diagonal) = 0.774					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table K-8b. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 3

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.072	0.024	0.002	0.000	0.097
PP	0.024	0.090	0.047	0.000	0.161
P	0.002	0.047	0.382	0.082	0.512
PWD	0.000	0.000	0.082	0.148	0.230
Total	0.097	0.161	0.512	0.230	1.000
Overall Consistency (sum of diagonal) = 0.692					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-37. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 3

Accuracy	0.773
Consistency	0.691
Kappa (k)	0.524

Table O-38. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 3

Achievement Level	*Accuracy*	*Consistency*
SBP	0.820	0.739
PP	0.667	0.559
P	0.794	0.745
PWD	0.780	0.643

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-39. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 3

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.964	0.017	0.019	0.950
PP:P	0.930	0.037	0.033	0.902
P:PWD	0.878	0.079	0.042	0.836

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-40. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 4

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.062	0.019	0.000	0.000	0.081
PP	0.025	0.146	0.053	0.000	0.224
P	0.000	0.045	0.368	0.069	0.483
PWD	0.000	0.000	0.029	0.183	0.212
Total	0.088	0.210	0.450	0.252	1.000
Overall Accuracy (sum of diagonal) = 0.759					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-41. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 4

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.056	0.029	0.003	0.000	0.088
PP	0.029	0.115	0.065	0.001	0.210
P	0.003	0.065	0.316	0.067	0.450
PWD	0.000	0.001	0.067	0.185	0.252
Total	0.088	0.210	0.450	0.252	1.000
Overall Consistency (sum of diagonal) = 0.672					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-42. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 4

Accuracy	0.759
Consistency	0.672
Kappa (k)	0.519

Table O-43. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 4

Achievement Level	*Accuracy*	*Consistency*
SBP	0.763	0.639
PP	0.651	0.549
P	0.762	0.701
PWD	0.865	0.733

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-44. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 4

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.955	0.019	0.025	0.937
PP:P	0.901	0.053	0.046	0.863
P:PWD	0.902	0.070	0.029	0.865

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-45. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 5

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.044	0.013	0.000	0.000	0.057
PP	0.019	0.187	0.048	0.000	0.255
P	0.000	0.044	0.428	0.050	0.523
PWD	0.000	0.000	0.024	0.142	0.166
Total	0.063	0.245	0.499	0.193	1.000
Overall Accuracy (sum of diagonal) = 0.801					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-46. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 5

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.040	0.023	0.001	0.000	0.063
PP	0.023	0.158	0.064	0.000	0.245
P	0.001	0.064	0.384	0.052	0.499
PWD	0.000	0.000	0.052	0.141	0.193
Total	0.063	0.245	0.499	0.193	1.000
Overall Consistency (sum of diagonal) = 0.723					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-47. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 5

Accuracy	0.801
Consistency	0.723
Kappa (k)	0.574

Table O-48. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 5

Achievement Level	*Accuracy*	*Consistency*
SBP	0.770	0.634
PP	0.736	0.647
P	0.819	0.768
PWD	0.858	0.732

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-49. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 5

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.968	0.013	0.019	0.954
PP:P	0.908	0.048	0.044	0.871
P:PWD	0.926	0.050	0.024	0.897

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-50. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 6

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.058	0.014	0.000	0.000	0.072
PP	0.018	0.207	0.043	0.000	0.268
P	0.000	0.038	0.416	0.043	0.497
PWD	0.000	0.000	0.021	0.142	0.163
Total	0.076	0.259	0.480	0.186	1.000
Overall Accuracy (sum of diagonal) = 0.823					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-51. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 6

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.053	0.022	0.000	0.000	0.076
PP	0.022	0.180	0.057	0.000	0.259
P	0.000	0.057	0.379	0.044	0.480
PWD	0.000	0.000	0.044	0.141	0.186
Total	0.076	0.259	0.480	0.186	1.000
Overall Consistency (sum of diagonal) = 0.753					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-52. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 6

Accuracy	0.823
Consistency	0.753
Kappa (k)	0.627

Table O-53. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 6

Achievement Level	*Accuracy*	*Consistency*
SBP	0.806	0.703
PP	0.771	0.695
P	0.837	0.789
PWD	0.874	0.761

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-54. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 6

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.968	0.014	0.018	0.955
PP:P	0.919	0.043	0.038	0.887
P:PWD	0.936	0.043	0.021	0.911

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-55. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 7

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.044	0.012	0.000	0.000	0.056
PP	0.017	0.165	0.043	0.000	0.225
P	0.000	0.040	0.446	0.053	0.539
PWD	0.000	0.000	0.024	0.157	0.181
Total	0.061	0.217	0.513	0.209	1.000
Overall Accuracy (sum of diagonal) = 0.812					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-56. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 7

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.040	0.020	0.001	0.000	0.061
PP	0.020	0.139	0.057	0.000	0.217
P	0.001	0.057	0.402	0.053	0.513
PWD	0.000	0.000	0.053	0.156	0.209
Total	0.061	0.217	0.513	0.209	1.000
Overall Consistency (sum of diagonal) = 0.737					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-57. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 7

Accuracy	0.811
Consistency	0.737
Kappa (k)	0.590

Table O-58. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 7

Achievement Level	*Accuracy*	*Consistency*
SBP	0.782	0.656
PP	0.732	0.641
P	0.828	0.784
PWD	0.866	0.745

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-59. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 7

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.971	0.012	0.017	0.958
PP:P	0.917	0.043	0.040	0.884
P:PWD	0.923	0.053	0.024	0.894

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-60. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 8

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.065	0.016	0.000	0.000	0.080
PP	0.021	0.199	0.046	0.000	0.266
P	0.000	0.041	0.405	0.046	0.492
PWD	0.000	0.000	0.022	0.140	0.161
Total	0.085	0.256	0.473	0.186	1.000
Overall Accuracy (sum of diagonal) = 0.809					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-61. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 8

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.060	0.025	0.000	0.000	0.085
PP	0.025	0.170	0.060	0.000	0.256
P	0.000	0.060	0.366	0.047	0.473
PWD	0.000	0.000	0.047	0.139	0.186
Total	0.085	0.256	0.473	0.186	1.000
Overall Consistency (sum of diagonal) = 0.735					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-62. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade

Accuracy	0.809
Consistency	0.734
Kappa (k)	0.603

Table O-63. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 8

Achievement Level	*Accuracy*	*Consistency*
SBP	0.803	0.698
PP	0.749	0.666
P	0.824	0.773
PWD	0.866	0.746

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-64. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 8

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.964	0.016	0.021	0.949
PP:P	0.913	0.046	0.041	0.879
P:PWD	0.932	0.046	0.022	0.906

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-65. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Reading Grade 11

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.072	0.016	0.000	0.000	0.088
PP	0.020	0.180	0.043	0.000	0.243
P	0.000	0.037	0.376	0.048	0.462
PWD	0.000	0.000	0.024	0.183	0.208
Total	0.092	0.233	0.443	0.232	1.000
Overall Accuracy (sum of diagonal) = 0.811					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-66. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Reading Grade 11

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.066	0.025	0.000	0.000	0.092
PP	0.025	0.153	0.055	0.000	0.233
P	0.000	0.055	0.338	0.050	0.443
PWD	0.000	0.000	0.050	0.181	0.232
Total	0.092	0.233	0.443	0.232	1.000
Overall Consistency (sum of diagonal) = 0.738					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-67. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Reading Grade 11

Accuracy	0.811
Consistency	0.738
Kappa (k)	0.618

Table O-68. 2008–09 NECAP: Indices Conditional On Achievement Level—Reading Grade 11

Achievement Level	*Accuracy*	*Consistency*
SBP	0.815	0.721
PP	0.742	0.656
P	0.815	0.761
PWD	0.883	0.782

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-69. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Reading Grade 11

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.964	0.016	0.020	0.949
PP:P	0.920	0.043	0.037	0.889
P:PWD	0.927	0.048	0.024	0.899

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-70. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Writing Grade 5

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.083	0.034	0.003	0.000	0.120
PP	0.052	0.158	0.083	0.014	0.307
P	0.004	0.065	0.146	0.104	0.320
PWD	0.000	0.004	0.041	0.209	0.254
Total	0.139	0.261	0.273	0.327	1.000
Overall Accuracy (sum of diagonal) = 0596					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-71. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Writing Grade 5

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.073	0.049	0.014	0.002	0.139
PP	0.049	0.111	0.074	0.027	0.261
P	0.014	0.074	0.104	0.082	0.273
PWD	0.002	0.027	0.082	0.217	0.327
Total	0.139	0.261	0.273	0.327	1.000
Overall Consistency (sum of diagonal) = 0.505					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-72. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Writing Grade 5

Accuracy	0.595
Consistency	0.504
Kappa (k)	0.321

Table O-73. 2008–09 NECAP: Indices Conditional On Achievement Level—Writing Grade 5

Achievement Level	*Accuracy*	*Consistency*
SBP	0.690	0.528
PP	0.514	0.424
P	0.458	0.379
PWD	0.821	0.662

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-74. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Writing Grade 5

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.907	0.037	0.056	0.869
PP:P	0.827	0.100	0.073	0.765
P:PWD	0.836	0.118	0.045	0.779

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Table O-75. 2008–09 NECAP: Decision Accuracy Cross-Tabulation of True and Observed Achievement Level Proportions—Writing Grade 8

Observed Achievement Level	*True Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.098	0.037	0.001	0.000	0.136
PP	0.051	0.235	0.099	0.002	0.386
P	0.001	0.063	0.265	0.073	0.402
PWD	0.000	0.000	0.015	0.061	0.076
Total	0.149	0.335	0.381	0.135	1.000
Overall Accuracy (sum of diagonal) = 0.659					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-76. 2008–09 NECAP: Decision Consistency Cross-Tabulation of Observed Achievement Level Proportions for Two Parallel Forms—Writing Grade 8

Form 2 Achievement Level	*Form 1 Achievement Level*				
	SBP	PP	P	PWD	Total
SBP	0.086	0.055	0.008	0.000	0.149
PP	0.055	0.176	0.099	0.006	0.335
P	0.008	0.099	0.218	0.057	0.381
PWD	0.000	0.006	0.057	0.073	0.135
Total	0.149	0.335	0.381	0.135	1.000
Overall Consistency (sum of diagonal) = 0.553					

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-77. 2008–09 NECAP: Summary of Overall Accuracy and Consistency Indices—Writing Grade 8

Accuracy	0.659
Consistency	0.552
Kappa (k)	0.362

Table O-78. 2008–09 NECAP: Indices Conditional On Achievement Level—Writing Grade 8

Achievement Level	*Accuracy*	*Consistency*
SBP	0.719	0.579
PP	0.609	0.524
P	0.661	0.572
PWD	0.796	0.538

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction

Table O-79. 2008–09 NECAP: Accuracy and Consistency Indices at Cut points—Writing Grade 8

Cut point	*Accuracy*	*False Positive*	*False Negative*	*Consistency*
SBP:PP	0.911	0.038	0.051	0.875
PP:P	0.835	0.102	0.064	0.776
P:PWD	0.910	0.074	0.016	0.875

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction
False Positive = proportion of students with observed score above cut point and true score below cut point
False Negative = proportion of students with observed score below cut point and true score above cut point

Appendix P—STUDENT QUESTIONNAIRE DATA

Table P-1. 2008–09 NECAP: Average Scaled Score, Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 3

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3869	13	344	673	666	1838	692	17	17	48	18
	A	8061	27	343	1299	1572	4131	1059	16	20	51	13
	B	12931	43	350	700	1566	7225	3440	5	12	56	27
	C	5526	18	347	620	837	2905	1164	11	15	53	21
2	(blank)	3916	13	344	680	689	1850	697	17	18	47	18
	A	7718	25	346	882	1382	4105	1349	11	18	53	17
	B	11098	37	349	798	1377	6102	2821	7	12	55	25
	C	4001	13	348	356	530	2159	956	9	13	54	24
	D	3654	12	344	576	663	1883	532	16	18	52	15
3	(blank)	3985	13	344	708	695	1886	696	18	17	47	17
	A	17513	58	347	1743	2621	9504	3645	10	15	54	21
	B	7973	26	348	639	1128	4320	1886	8	14	54	24
	C	916	3	341	202	197	389	128	22	22	42	14
4	(blank)	4070	13	344	725	708	1926	711	18	17	47	17
	A	7186	24	342	1211	1499	3733	743	17	21	52	10
	B	11576	38	349	715	1434	6552	2875	6	12	57	25
	C	7555	25	349	641	1000	3888	2026	8	13	51	27
5	(blank)	3932	13	344	700	683	1851	698	18	17	47	18
	A	21425	71	348	1839	2995	11582	5009	9	14	54	23
	B	2750	9	345	363	496	1486	405	13	18	54	15
	C	1716	6	344	246	328	931	211	14	19	54	12
	D	564	2	339	144	139	249	32	26	25	44	6
6	(blank)	3971	13	344	689	690	1890	702	17	17	48	18
	A	16260	54	347	1606	2394	8612	3648	10	15	53	22
	B	6624	22	347	594	991	3709	1330	9	15	56	20
	C	1305	4	346	159	203	662	281	12	16	51	22
	D	2227	7	346	244	363	1226	394	11	16	55	18
7	(blank)	3964	13	344	705	691	1871	697	18	17	47	18
	A	16962	56	349	1156	2199	9247	4360	7	13	55	26
	B	6064	20	344	829	1085	3289	861	14	18	54	14
	C	2911	10	344	445	539	1516	411	15	19	52	14
	D	486	2	336	157	127	176	26	32	26	36	5
8	(blank)	3953	13	344	694	702	1863	694	18	18	47	18
	A	15094	50	348	1288	2095	8113	3598	9	14	54	24
	B	7339	24	348	619	1045	4097	1578	8	14	56	22
	C	1648	5	344	255	284	847	262	15	17	51	16
	D	2353	8	342	436	515	1179	223	19	22	50	9
9	(blank)	4095	13	344	707	702	1953	733	17	17	48	18
	A	9343	31	346	1012	1509	5067	1755	11	16	54	19
	B	8211	27	349	590	977	4420	2224	7	12	54	27
	C	4571	15	347	425	729	2389	1028	9	16	52	22
	D	4167	14	345	558	724	2270	615	13	17	54	15

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-2. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 4

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3399	11	443	517	685	1581	616	15	20	47	18
	A	6726	22	443	1104	1427	3191	1004	16	21	47	15
	B	15793	52	448	1043	2577	8233	3940	7	16	52	25
	C	4517	15	446	455	806	2272	984	10	18	50	22
2	(blank)	3408	11	443	514	687	1590	617	15	20	47	18
	A	6497	21	444	800	1334	3139	1224	12	21	48	19
	B	12736	42	448	891	2022	6642	3181	7	16	52	25
	C	4734	16	447	397	771	2489	1077	8	16	53	23
	D	3060	10	442	517	681	1417	445	17	22	46	15
3	(blank)	3536	12	443	534	717	1634	651	15	20	46	18
	A	17202	57	446	1618	3151	8761	3672	9	18	51	21
	B	9102	30	447	800	1494	4647	2161	9	16	51	24
	C	595	2	438	167	133	235	60	28	22	39	10
4	(blank)	3514	12	443	551	709	1626	628	16	20	46	18
	A	5735	19	440	1076	1485	2611	563	19	26	46	10
	B	13693	45	447	940	2240	7373	3140	7	16	54	23
	C	7493	25	449	552	1061	3667	2213	7	14	49	30
5	(blank)	3434	11	443	525	686	1602	621	15	20	47	18
	A	23121	76	447	1901	3881	11872	5467	8	17	51	24
	B	2246	7	442	347	532	1081	286	15	24	48	13
	C	1261	4	441	232	300	587	142	18	24	47	11
	D	373	1	436	114	96	135	28	31	26	36	8
6	(blank)	3497	11	443	524	713	1631	629	15	20	47	18
	A	16508	54	447	1472	2811	8291	3934	9	17	50	24
	B	6773	22	446	651	1228	3527	1367	10	18	52	20
	C	1377	5	445	160	257	700	260	12	19	51	19
	D	2280	7	443	312	486	1128	354	14	21	49	16
7	(blank)	3446	11	443	519	696	1596	635	15	20	46	18
	A	17818	59	448	1184	2774	9345	4515	7	16	52	25
	B	5890	19	443	901	1246	2800	943	15	21	48	16
	C	2997	10	443	416	701	1437	443	14	23	48	15
	D	284	1	435	99	78	99	8	35	27	35	3
8	(blank)	3487	11	443	525	710	1618	634	15	20	46	18
	A	14432	47	448	1137	2316	7206	3773	8	16	50	26
	B	8740	29	446	812	1609	4613	1706	9	18	53	20
	C	1733	6	444	222	352	911	248	13	20	53	14
	D	2043	7	440	423	508	929	183	21	25	45	9
9	(blank)	3613	12	444	535	705	1699	674	15	20	47	19
	A	10089	33	446	966	1854	5192	2077	10	18	51	21
	B	8855	29	448	649	1379	4485	2342	7	16	51	26
	C	4260	14	446	447	798	2064	951	10	19	48	22
	D	3618	12	443	522	759	1837	500	14	21	51	14

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-3. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 5

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3485	11	544	483	754	1752	496	14	22	50	14
	A	7504	23	544	899	1712	3893	1000	12	23	52	13
	B	17746	55	547	1049	3265	10350	3082	6	18	58	17
	C	3395	11	546	290	649	1939	517	9	19	57	15
2	(blank)	3473	11	544	476	757	1743	497	14	22	50	14
	A	5112	16	544	617	1213	2654	628	12	24	52	12
	B	14221	44	547	840	2569	8229	2583	6	18	58	18
	C	6130	19	547	368	1120	3566	1076	6	18	58	18
	D	3194	10	543	420	721	1742	311	13	23	55	10
3	(blank)	3568	11	544	493	766	1798	511	14	21	50	14
	A	17512	55	546	1341	3571	9831	2769	8	20	56	16
	B	10564	33	547	763	1900	6110	1791	7	18	58	17
	C	486	2	538	124	143	195	24	26	29	40	5
4	(blank)	3636	11	544	519	787	1820	510	14	22	50	14
	A	4971	15	540	887	1520	2249	315	18	31	45	6
	B	15965	50	547	949	3098	9395	2523	6	19	59	16
	C	7558	24	549	366	975	4470	1747	5	13	59	23
5	(blank)	3500	11	544	477	755	1766	502	14	22	50	14
	A	24108	75	547	1671	4597	13816	4024	7	19	57	17
	B	2744	9	544	311	619	1463	351	11	23	53	13
	C	1239	4	543	173	278	635	153	14	22	51	12
	D	539	2	542	89	131	254	65	17	24	47	12
6	(blank)	3559	11	544	480	769	1796	514	13	22	50	14
	A	15924	50	547	1134	2860	8995	2935	7	18	56	18
	B	8146	25	546	655	1687	4660	1144	8	21	57	14
	C	1960	6	545	161	411	1131	257	8	21	58	13
	D	2541	8	543	291	653	1352	245	11	26	53	10
7	(blank)	3516	11	544	478	755	1780	503	14	21	51	14
	A	19504	61	548	1030	3267	11542	3665	5	17	59	19
	B	5951	19	543	796	1491	3030	634	13	25	51	11
	C	2890	9	543	324	789	1487	290	11	27	51	10
	D	269	1	534	93	78	95	3	35	29	35	1
8	(blank)	3535	11	544	485	765	1780	505	14	22	50	14
	A	14373	45	548	876	2366	8195	2936	6	16	57	20
	B	9999	31	546	755	2138	5762	1344	8	21	58	13
	C	2054	6	544	221	466	1159	208	11	23	56	10
	D	2169	7	540	384	645	1038	102	18	30	48	5
9	(blank)	3696	12	544	500	791	1861	544	14	21	50	15
	A	12609	39	546	923	2435	7229	2022	7	19	57	16
	B	8367	26	547	511	1498	4777	1581	6	18	57	19
	C	4052	13	546	382	808	2214	648	9	20	55	16
	D	3406	11	543	405	848	1853	300	12	25	54	9

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-4. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 6

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3802	12	643	579	873	1883	467	15	23	50	12
	A	6630	20	644	953	1477	3463	737	14	22	52	11
	B	18384	57	648	1153	3404	11084	2743	6	19	60	15
	C	3563	11	647	312	638	2131	482	9	18	60	14
2	(blank)	3798	12	643	590	886	1858	464	16	23	49	12
	A	3491	11	644	446	788	1874	383	13	23	54	11
	B	13836	43	648	802	2323	8454	2257	6	17	61	16
	C	7346	23	647	555	1357	4388	1046	8	18	60	14
	D	3908	12	642	604	1038	1987	279	15	27	51	7
3	(blank)	3865	12	643	595	900	1895	475	15	23	49	12
	A	15948	49	646	1383	3189	9155	2221	9	20	57	14
	B	11968	37	647	856	2145	7262	1705	7	18	61	14
	C	598	2	637	163	158	249	28	27	26	42	5
4	(blank)	3920	12	643	616	913	1914	477	16	23	49	12
	A	4131	13	639	892	1187	1797	255	22	29	44	6
	B	16629	51	647	1072	3237	10038	2282	6	19	60	14
	C	7699	24	649	417	1055	4812	1415	5	14	63	18
5	(blank)	3813	12	643	585	878	1880	470	15	23	49	12
	A	24583	76	647	1791	4598	14711	3483	7	19	60	14
	B	2613	8	644	337	587	1371	318	13	22	52	12
	C	875	3	642	164	211	403	97	19	24	46	11
	D	495	2	641	120	118	196	61	24	24	40	12
6	(blank)	3859	12	643	588	892	1909	470	15	23	49	12
	A	15411	48	648	1109	2745	9173	2384	7	18	60	15
	B	10164	31	646	813	2067	5964	1320	8	20	59	13
	C	1274	4	645	158	253	719	144	12	20	56	11
	D	1671	5	641	329	435	796	111	20	26	48	7
7	(blank)	3849	12	643	584	888	1909	468	15	23	50	12
	A	19866	61	648	1095	3391	12201	3179	6	17	61	16
	B	5235	16	643	779	1245	2668	543	15	24	51	10
	C	3167	10	643	434	796	1705	232	14	25	54	7
	D	262	1	633	105	72	78	7	40	27	30	3
8	(blank)	3876	12	643	591	903	1915	467	15	23	49	12
	A	11228	35	650	685	1582	6701	2260	6	14	60	20
	B	10968	34	646	826	2253	6575	1314	8	21	60	12
	C	3208	10	644	311	733	1905	259	10	23	59	8
	D	3099	10	640	584	921	1465	129	19	30	47	4
9	(blank)	4096	13	644	602	929	2043	522	15	23	50	13
	A	14732	45	647	1087	2863	8763	2019	7	19	59	14
	B	6743	21	648	523	1221	3895	1104	8	18	58	16
	C	3512	11	647	362	576	2036	538	10	16	58	15
	D	3296	10	643	423	803	1824	246	13	24	55	7

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-5. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 7

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3771	11	744	530	813	1935	493	14	22	51	13
	A	6327	19	745	795	1349	3396	787	13	21	54	12
	B	19088	58	749	916	3121	11492	3559	5	16	60	19
	C	3729	11	749	222	547	2238	722	6	15	60	19
2	(blank)	3792	12	744	539	823	1942	488	14	22	51	13
	A	2005	6	744	287	434	1050	234	14	22	52	12
	B	11276	34	749	599	1803	6672	2202	5	16	59	20
	C	9834	30	750	456	1485	5949	1944	5	15	60	20
	D	6008	18	745	582	1285	3448	693	10	21	57	12
3	(blank)	3845	12	744	552	836	1960	497	14	22	51	13
	A	15169	46	748	995	2654	8859	2661	7	17	58	18
	B	13054	40	749	716	2125	7880	2333	5	16	60	18
	C	847	3	740	200	215	362	70	24	25	43	8
4	(blank)	3972	12	744	581	859	2023	509	15	22	51	13
	A	3895	12	741	690	1039	1877	289	18	27	48	7
	B	17271	52	749	859	3047	10431	2934	5	18	60	17
	C	7777	24	751	333	885	4730	1829	4	11	61	24
5	(blank)	3778	11	744	540	813	1936	489	14	22	51	13
	A	25239	77	749	1354	4182	15256	4447	5	17	60	18
	B	2634	8	746	300	569	1332	433	11	22	51	16
	C	868	3	744	176	164	389	139	20	19	45	16
	D	396	1	741	93	102	148	53	23	26	37	13
6	(blank)	3798	12	744	534	829	1948	487	14	22	51	13
	A	13443	41	749	745	2058	8094	2546	6	15	60	19
	B	12418	38	749	733	2136	7350	2199	6	17	59	18
	C	1500	5	746	152	299	836	213	10	20	56	14
	D	1756	5	741	299	508	833	116	17	29	47	7
7	(blank)	3814	12	744	532	828	1958	496	14	22	51	13
	A	20287	62	750	862	2935	12373	4117	4	14	61	20
	B	4833	15	745	587	1079	2594	573	12	22	54	12
	C	3618	11	744	365	885	2002	366	10	24	55	10
	D	363	1	735	117	103	134	9	32	28	37	2
8	(blank)	3808	12	744	538	831	1943	496	14	22	51	13
	A	9464	29	752	411	1014	5500	2539	4	11	58	27
	B	10850	33	748	628	1828	6595	1799	6	17	61	17
	C	4173	13	746	295	901	2532	445	7	22	61	11
	D	4620	14	742	591	1256	2491	282	13	27	54	6
9	(blank)	4138	13	745	552	890	2153	543	13	22	52	13
	A	17228	52	749	911	2884	10347	3086	5	17	60	18
	B	5246	16	749	384	859	2942	1061	7	16	56	20
	C	3100	9	748	273	525	1759	543	9	17	57	18
	D	3203	10	745	343	672	1860	328	11	21	58	10

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-6. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–9—Reading Grade 8

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	3300	10	841	661	802	1419	418	20	24	43	13
	A	5364	16	842	851	1451	2561	501	16	27	48	9
	B	20089	60	847	1334	4377	11290	3088	7	22	56	15
	C	4945	15	848	290	908	2841	906	6	18	57	18
2	(blank)	3363	10	841	669	831	1454	409	20	25	43	12
	A	1605	5	842	266	451	735	153	17	28	46	10
	B	10644	32	847	786	2221	5923	1714	7	21	56	16
	C	11276	33	848	639	2231	6434	1972	6	20	57	17
	D	6810	20	843	776	1804	3565	665	11	26	52	10
3	(blank)	3442	10	841	704	842	1487	409	20	24	43	12
	A	14483	43	846	1196	3306	7871	2110	8	23	54	15
	B	14559	43	847	1000	3023	8209	2327	7	21	56	16
	C	1214	4	839	236	367	544	67	19	30	45	6
4	(blank)	3605	11	841	725	893	1548	439	20	25	43	12
	A	3198	9	838	750	1044	1252	152	23	33	39	5
	B	17566	52	846	1209	4141	9752	2464	7	24	56	14
	C	9329	28	849	452	1460	5559	1858	5	16	60	20
5	(blank)	3437	10	841	679	845	1488	425	20	25	43	12
	A	26775	79	847	1875	5817	15096	3987	7	22	56	15
	B	2265	7	844	309	558	1071	327	14	25	47	14
	C	820	2	843	157	207	327	129	19	25	40	16
	D	401	1	838	116	111	129	45	29	28	32	11
6	(blank)	3436	10	841	684	842	1489	421	20	25	43	12
	A	12687	38	847	932	2673	7094	1988	7	21	56	16
	B	13879	41	847	982	3055	7733	2109	7	22	56	15
	C	1862	6	845	198	411	989	264	11	22	53	14
	D	1834	5	840	340	557	806	131	19	30	44	7
7	(blank)	3491	10	841	676	862	1521	432	19	25	44	12
	A	21658	64	848	1141	4199	12553	3765	5	19	58	17
	B	4626	14	842	719	1316	2153	438	16	28	47	9
	C	3504	10	842	465	1024	1747	268	13	29	50	8
	D	419	1	834	135	137	137	10	32	33	33	2
8	(blank)	3442	10	841	683	864	1480	415	20	25	43	12
	A	8865	26	851	378	1187	5092	2208	4	13	57	25
	B	10271	30	847	739	2128	5901	1503	7	21	57	15
	C	5039	15	844	460	1297	2774	508	9	26	55	10
	D	6081	18	840	876	2062	2864	279	14	34	47	5
9	(blank)	3712	11	842	701	911	1650	450	19	25	44	12
	A	19848	59	847	1360	4297	11131	3060	7	22	56	15
	B	4330	13	847	415	887	2290	738	10	20	53	17
	C	2816	8	846	283	652	1464	417	10	23	52	15
	D	2992	9	843	377	791	1576	248	13	26	53	8

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-7. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 13–23—Reading Grade 11

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
13	(blank)	8022	24	1143	1225	1590	3741	1466	15	20	47	18
	A	4122	12	1141	766	1050	1839	467	19	25	45	11
	B	14221	42	1145	1049	2784	8052	2336	7	20	57	16
	C	7220	21	1149	361	811	3955	2093	5	11	55	29
14	(blank)	7636	23	1144	1137	1497	3571	1431	15	20	47	19
	A	1840	5	1141	327	411	875	227	18	22	48	12
	B	9977	30	1147	648	1524	5557	2248	6	15	56	23
	C	8988	27	1147	570	1493	5062	1863	6	17	56	21
	D	5144	15	1142	719	1310	2522	593	14	25	49	12
15	(blank)	8032	24	1143	1252	1606	3723	1451	16	20	46	18
	A	6281	19	1144	677	1304	3363	937	11	21	54	15
	B	15644	47	1147	930	2510	8706	3498	6	16	56	22
	C	3628	11	1142	542	815	1795	476	15	22	49	13
16	(blank)	8110	24	1143	1243	1609	3792	1466	15	20	47	18
	A	2795	8	1138	661	794	1155	185	24	28	41	7
	B	13221	39	1145	1075	2736	7516	1894	8	21	57	14
	C	9459	28	1149	422	1096	5124	2817	4	12	54	30
17	(blank)	7682	23	1144	1162	1494	3588	1438	15	19	47	19
	A	17260	51	1146	1200	3129	9786	3145	7	18	57	18
	B	5019	15	1145	510	929	2563	1017	10	19	51	20
	C	2536	8	1145	322	453	1195	566	13	18	47	22
	D	1088	3	1142	207	230	455	196	19	21	42	18
18	(blank)	7743	23	1144	1161	1513	3629	1440	15	20	47	19
	A	11173	33	1148	662	1575	6124	2812	6	14	55	25
	B	9824	29	1146	741	1851	5536	1696	8	19	56	17
	C	2526	8	1142	353	579	1306	288	14	23	52	11
	D	2319	7	1138	484	717	992	126	21	31	43	5
19	(blank)	7684	23	1144	1156	1499	3592	1437	15	20	47	19
	A	5696	17	1147	497	904	2853	1442	9	16	50	25
	B	12151	36	1147	756	2023	6870	2502	6	17	57	21
	C	4842	14	1144	477	955	2703	707	10	20	56	15
	D	3212	10	1141	515	854	1569	274	16	27	49	9
20	(blank)	7751	23	1144	1160	1501	3642	1448	15	19	47	19
	A	11851	35	1147	728	1914	6673	2536	6	16	56	21
	B	8810	26	1145	880	1776	4584	1570	10	20	52	18
	C	3428	10	1145	377	604	1822	625	11	18	53	18
	D	1745	5	1142	256	440	866	183	15	25	50	10
21	(blank)	7779	23	1144	1152	1533	3643	1451	15	20	47	19
	A	5805	17	1147	406	830	3137	1432	7	14	54	25
	B	6807	20	1147	566	1132	3572	1537	8	17	52	23
	C	6011	18	1146	486	1045	3326	1154	8	17	55	19
	D	7183	21	1143	791	1695	3909	788	11	24	54	11
22	(blank)	8212	24	1144	1191	1634	3879	1508	15	20	47	18
	A	16502	49	1147	1077	2657	9284	3484	7	16	56	21
	B	4188	12	1144	511	818	2104	755	12	20	50	18
	C	2414	7	1143	302	551	1218	343	13	23	50	14
	D	2269	7	1142	320	575	1102	272	14	25	49	12
23	(blank)	7921	24	1144	1177	1563	3715	1466	15	20	47	19
	A	8381	25	1151	410	797	4162	3012	5	10	50	36
	B	11231	33	1145	882	2088	6706	1555	8	19	60	14
	C	4245	13	1141	599	1234	2164	248	14	29	51	6
	D	1807	5	1139	333	553	840	81	18	31	46	4

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-8. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 10–20—Mathematics Grade 3

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	4131	14	341	885	856	1727	663	21	21	42	16
	A	8430	28	341	1635	1936	3859	1000	19	23	46	12
	B	11924	39	346	992	1995	6067	2870	8	17	51	24
	C	5966	20	345	792	1054	2740	1380	13	18	46	23
11	(blank)	4136	14	341	875	844	1744	673	21	20	42	16
	A	16771	55	344	2191	3296	8159	3125	13	20	49	19
	B	8423	28	345	971	1498	4022	1932	12	18	48	23
	C	1121	4	341	267	203	468	183	24	18	42	16
12	(blank)	4247	14	342	867	857	1790	733	20	20	42	17
	A	1855	6	338	533	497	686	139	29	27	37	7
	B	3281	11	341	594	793	1491	403	18	24	45	12
	C	11865	39	346	1141	1936	6028	2760	10	16	51	23
	D	9203	30	344	1169	1758	4398	1878	13	19	48	20
13	(blank)	3913	13	341	823	800	1648	642	21	20	42	16
	A	21854	72	345	2614	3891	10607	4742	12	18	49	22
	B	2858	9	342	446	672	1364	376	16	24	48	13
	C	1317	4	340	287	327	580	123	22	25	44	9
	D	509	2	338	134	151	194	30	26	30	38	6
14	(blank)	4404	14	341	926	924	1845	709	21	21	42	16
	A	21180	70	345	2463	3714	10327	4676	12	18	49	22
	B	3056	10	342	490	718	1473	375	16	23	48	12
	C	1284	4	340	300	326	553	105	23	25	43	8
	D	527	2	339	125	159	195	48	24	30	37	9
15	(blank)	4038	13	341	848	828	1695	667	21	21	42	17
	A	6032	20	340	1351	1468	2479	734	22	24	41	12
	B	11094	36	345	1243	2053	5565	2233	11	19	50	20
	C	5031	17	347	426	746	2469	1390	8	15	49	28
	D	4256	14	345	436	746	2185	889	10	18	51	21
16	(blank)	3982	13	341	848	812	1674	648	21	20	42	16
	A	14307	47	345	1803	2643	6728	3133	13	18	47	22
	B	8838	29	344	1044	1711	4459	1624	12	19	50	18
	C	1685	6	344	258	318	797	312	15	19	47	19
	D	1639	5	341	351	357	735	196	21	22	45	12
17	(blank)	4229	14	341	881	868	1782	698	21	21	42	17
	A	5365	18	340	1191	1330	2238	606	22	25	42	11
	B	10103	33	344	1178	1998	4992	1935	12	20	49	19
	C	6784	22	347	542	959	3490	1793	8	14	51	26
	D	3970	13	345	512	686	1891	881	13	17	48	22
18	(blank)	4363	14	341	908	903	1836	716	21	21	42	16
	A	2565	8	338	704	671	984	206	27	26	38	8
	B	8098	27	345	839	1559	4133	1567	10	19	51	19
	C	7006	23	347	641	1056	3475	1834	9	15	50	26
	D	8419	28	344	1212	1652	3965	1590	14	20	47	19
19	(blank)	4369	14	342	886	888	1864	731	20	20	43	17
	A	7464	25	342	1179	1617	3642	1026	16	22	49	14
	B	5736	19	347	485	892	2810	1549	8	16	49	27
	C	1819	6	342	351	379	832	257	19	21	46	14
	D	7177	24	345	785	1225	3503	1664	11	17	49	23
	Ee	3886	13	343	618	840	1742	686	16	22	45	18
20	(blank)	4703	15	341	999	980	1984	740	21	21	42	16
	A	12957	43	344	1836	2558	6131	2432	14	20	47	19
	B	9186	30	346	875	1638	4659	2014	10	18	51	22
	C	2301	8	346	279	352	1067	603	12	15	46	26
	D	1304	4	340	315	313	552	124	24	24	42	10

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-9. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 10–20—Mathematics Grade 4

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	3514	12	442	690	676	1511	637	20	19	43	18
	A	6823	22	441	1370	1506	3051	896	20	22	45	13
	B	14497	48	447	1314	2390	7142	3651	9	16	49	25
	C	5665	19	447	626	837	2600	1602	11	15	46	28
11	(blank)	3665	12	442	705	706	1578	676	19	19	43	18
	A	16477	54	445	2016	3021	7900	3540	12	18	48	21
	B	9426	31	446	1066	1509	4440	2411	11	16	47	26
	C	931	3	441	213	173	386	159	23	19	41	17
12	(blank)	3623	12	443	669	684	1560	710	18	19	43	20
	A	1884	6	439	522	453	694	215	28	24	37	11
	B	4299	14	442	735	1012	1926	626	17	24	45	15
	C	13203	43	447	1243	2028	6519	3413	9	15	49	26
	D	7490	25	446	831	1232	3605	1822	11	16	48	24
13	(blank)	3467	11	442	664	659	1498	646	19	19	43	19
	A	23264	76	446	2566	3870	11113	5715	11	17	48	25
	B	2392	8	441	431	553	1106	302	18	23	46	13
	C	1031	3	439	237	249	453	92	23	24	44	9
	D	345	1	438	102	78	134	31	30	23	39	9
14	(blank)	3657	12	442	695	691	1593	678	19	19	44	19
	A	22901	75	446	2439	3761	11054	5647	11	16	48	25
	B	2593	9	441	496	629	1144	324	19	24	44	12
	C	993	3	438	261	244	386	102	26	25	39	10
	D	355	1	438	109	84	127	35	31	24	36	10
15	(blank)	3594	12	442	683	690	1549	672	19	19	43	19
	A	4590	15	441	1050	1030	1796	714	23	22	39	16
	B	11588	38	445	1283	2045	5663	2597	11	18	49	22
	C	6419	21	448	458	897	3184	1880	7	14	50	29
	D	4308	14	445	526	747	2112	923	12	17	49	21
16	(blank)	3514	12	442	676	685	1503	650	19	19	43	18
	A	16261	53	446	1861	2746	7832	3822	11	17	48	24
	B	8279	27	445	975	1509	3955	1840	12	18	48	22
	C	1411	5	445	213	243	617	338	15	17	44	24
	D	1034	3	439	275	226	397	136	27	22	38	13
17	(blank)	3765	12	442	713	729	1625	698	19	19	43	19
	A	3395	11	440	848	784	1310	453	25	23	39	13
	B	9015	30	444	1208	1705	4299	1803	13	19	48	20
	C	9354	31	448	632	1343	4740	2639	7	14	51	28
	D	4970	16	446	599	848	2330	1193	12	17	47	24
18	(blank)	3803	12	442	733	741	1638	691	19	19	43	18
	A	1601	5	437	531	386	565	119	33	24	35	7
	B	7937	26	444	1049	1583	3842	1463	13	20	48	18
	C	9548	31	448	724	1355	4748	2721	8	14	50	28
	D	7610	25	445	963	1344	3511	1792	13	18	46	24
19	(blank)	3835	13	443	697	725	1672	741	18	19	44	19
	A	7873	26	443	1147	1586	3787	1353	15	20	48	17
	B	6000	20	448	501	873	2903	1723	8	15	48	29
	C	1852	6	441	371	429	799	253	20	23	43	14
	D	7260	24	447	724	1113	3429	1994	10	15	47	27
	E	3679	12	444	560	683	1714	722	15	19	47	20
20	(blank)	4074	13	442	762	796	1751	765	19	20	43	19
	A	13594	45	445	1746	2359	6481	3008	13	17	48	22
	B	9827	32	446	968	1706	4762	2391	10	17	48	24
	C	2122	7	445	273	334	992	523	13	16	47	25
	D	882	3	438	251	214	318	99	28	24	36	11

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-10. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 10–20—Mathematics Grade 5

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	3638	11	542	897	556	1551	634	25	15	43	17
	A	8643	27	542	1940	1521	4059	1123	22	18	47	13
	B	15753	49	545	2136	2260	7951	3406	14	14	50	22
	C	4166	13	546	578	468	1971	1149	14	11	47	28
11	(blank)	3761	12	542	925	561	1618	657	25	15	43	17
	A	17026	53	544	2816	2672	8343	3195	17	16	49	19
	B	10556	33	545	1577	1437	5195	2347	15	14	49	22
	C	857	3	541	233	135	376	113	27	16	44	13
12	(blank)	3771	12	542	882	571	1625	693	23	15	43	18
	A	2276	7	540	604	402	1037	233	27	18	46	10
	B	5592	17	542	1120	949	2731	792	20	17	49	14
	C	15819	49	546	2148	2232	7880	3559	14	14	50	22
	D	4742	15	545	797	651	2259	1035	17	14	48	22
13	(blank)	3561	11	542	842	539	1544	636	24	15	43	18
	A	23633	73	545	3626	3379	11581	5047	15	14	49	21
	B	3242	10	542	650	578	1576	438	20	18	49	14
	C	1148	4	541	282	192	537	137	25	17	47	12
	D	616	2	540	151	117	294	54	25	19	48	9
14	(blank)	3779	12	542	873	578	1654	674	23	15	44	18
	A	23244	72	545	3438	3310	11521	4975	15	14	50	21
	B	3497	11	542	804	629	1605	459	23	18	46	13
	C	1119	3	540	298	204	476	141	27	18	43	13
	D	561	2	541	138	84	276	63	25	15	49	11
15	(blank)	3638	11	542	864	551	1575	648	24	15	43	18
	A	4477	14	541	1163	733	1892	689	26	16	42	15
	B	12893	40	545	1989	1923	6421	2560	15	15	50	20
	C	7079	22	546	793	940	3635	1711	11	13	51	24
	D	4113	13	543	742	658	2009	704	18	16	49	17
16	(blank)	3573	11	542	856	541	1541	635	24	15	43	18
	A	18176	56	545	2662	2612	8995	3907	15	14	49	21
	B	8187	25	544	1473	1258	4019	1437	18	15	49	18
	C	1336	4	543	275	225	614	222	21	17	46	17
	D	928	3	539	285	169	363	111	31	18	39	12
17	(blank)	3797	12	542	888	586	1646	677	23	15	43	18
	A	2281	7	539	722	425	890	244	32	19	39	11
	B	8371	26	543	1610	1377	4006	1378	19	16	48	16
	C	11769	37	546	1334	1534	6086	2815	11	13	52	24
	D	5982	19	544	997	883	2904	1198	17	15	49	20
18	(blank)	3828	12	542	911	591	1647	679	24	15	43	18
	A	2002	6	539	602	337	839	224	30	17	42	11
	B	10096	31	544	1724	1645	4926	1801	17	16	49	18
	C	10521	33	546	1234	1360	5449	2478	12	13	52	24
	D	5753	18	544	1080	872	2671	1130	19	15	46	20
19	(blank)	3832	12	542	884	568	1674	706	23	15	44	18
	A	9074	28	542	1803	1568	4419	1284	20	17	49	14
	B	6372	20	547	763	807	3149	1653	12	13	49	26
	C	2325	7	541	623	377	1026	299	27	16	44	13
	D	7218	22	546	842	943	3638	1795	12	13	50	25
	E	3379	10	543	636	542	1626	575	19	16	48	17
20	(blank)	3653	11	542	885	550	1564	654	24	15	43	18
	A	15409	48	545	2394	2237	7554	3224	16	15	49	21
	B	10324	32	544	1596	1568	5186	1974	15	15	50	19
	C	2144	7	544	407	327	1005	405	19	15	47	19
	D	670	2	536	269	123	223	55	40	18	33	8

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-11. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey Questions 10–20—Mathematics Grade 6

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	3920	12	640	1043	695	1513	669	27	18	39	17
	A	10242	32	641	2261	2087	4521	1373	22	20	44	13
	B	15271	47	645	2296	2596	6840	3539	15	17	45	23
	C	3001	9	646	459	397	1215	930	15	13	40	31
11	(blank)	4010	12	641	1037	710	1571	692	26	18	39	17
	A	15809	49	643	2893	2999	7017	2900	18	19	44	18
	B	11646	36	645	1785	1862	5203	2796	15	16	45	24
	C	969	3	638	344	204	298	123	36	21	31	13
12	(blank)	3898	12	641	996	695	1518	689	26	18	39	18
	A	4358	13	642	832	849	1965	712	19	19	45	16
	B	10532	32	644	1642	1896	4858	2136	16	18	46	20
	C	12129	37	644	2240	2070	5126	2693	18	17	42	22
	D	1517	5	642	349	265	622	281	23	17	41	19
13	(blank)	3835	12	641	979	669	1502	685	26	17	39	18
	A	23478	72	644	3845	4114	10395	5124	16	18	44	22
	B	3503	11	641	798	684	1537	484	23	20	44	14
	C	1030	3	640	288	198	402	142	28	19	39	14
	D	588	2	640	149	110	253	76	25	19	43	13
14	(blank)	3856	12	641	983	683	1512	678	25	18	39	18
	A	23255	72	644	3721	4081	10416	5037	16	18	45	22
	B	3627	11	641	882	701	1513	531	24	19	42	15
	C	1068	3	640	322	182	401	163	30	17	38	15
	D	628	2	641	151	128	247	102	24	20	39	16
15	(blank)	3947	12	641	1011	698	1539	699	26	18	39	18
	A	4566	14	641	1150	840	1795	781	25	18	39	17
	B	12338	38	644	2059	2224	5503	2552	17	18	45	21
	C	7101	22	646	904	1103	3316	1778	13	16	47	25
	D	4482	14	642	935	910	1936	701	21	20	43	16
16	(blank)	3873	12	641	1007	680	1515	671	26	18	39	17
	A	18095	56	644	2847	3127	8152	3969	16	17	45	22
	B	8690	27	643	1644	1653	3775	1618	19	19	43	19
	C	996	3	641	242	180	406	168	24	18	41	17
	D	780	2	636	319	135	241	85	41	17	31	11
17	(blank)	4266	13	641	1061	746	1693	766	25	17	40	18
	A	1412	4	635	564	290	445	113	40	21	32	8
	B	5908	18	641	1426	1231	2410	841	24	21	41	14
	C	11691	36	645	1495	1922	5444	2830	13	16	47	24
	D	9157	28	644	1513	1586	4097	1961	17	17	45	21
18	(blank)	4196	13	641	1063	740	1644	749	25	18	39	18
	A	2428	7	640	649	465	972	342	27	19	40	14
	B	9473	29	643	1791	1776	4143	1763	19	19	44	19
	C	10218	32	645	1389	1693	4703	2433	14	17	46	24
	D	6119	19	643	1167	1101	2627	1224	19	18	43	20
19	(blank)	4183	13	641	1041	729	1654	759	25	17	40	18
	A	10424	32	642	2100	2157	4576	1591	20	21	44	15
	B	6054	19	644	963	978	2649	1464	16	16	44	24
	C	2448	8	641	581	487	1049	331	24	20	43	14
	D	6446	20	646	781	963	2911	1791	12	15	45	28
	E	2879	9	643	593	461	1250	575	21	16	43	20
20	(blank)	4675	14	641	1125	822	1883	845	24	18	40	18
	A	16231	50	644	2599	2805	7232	3595	16	17	45	22
	B	9146	28	643	1691	1670	4069	1716	18	18	44	19
	C	1772	5	642	394	354	715	309	22	20	40	17
	D	610	2	635	250	124	190	46	41	20	31	8

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-12. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 10–20—Mathematics Grade 7

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	3923	12	739	1254	747	1361	561	32	19	35	14
	A	11210	34	739	2931	2527	4609	1143	26	23	41	10
	B	14816	45	744	2392	2466	6628	3330	16	17	45	22
	C	3062	9	747	437	334	1083	1208	14	11	35	39
11	(blank)	3995	12	739	1289	758	1393	555	32	19	35	14
	A	15206	46	742	3154	2903	6520	2629	21	19	43	17
	B	12472	38	744	2107	2126	5348	2891	17	17	43	23
	C	1338	4	738	464	287	420	167	35	21	31	12
12	(blank)	3894	12	739	1214	745	1374	561	31	19	35	14
	A	4421	13	741	970	859	1906	686	22	19	43	16
	B	10721	32	743	1905	1889	4786	2141	18	18	45	20
	C	12839	39	743	2566	2383	5225	2665	20	19	41	21
	D	1136	3	739	359	198	390	189	32	17	34	17
13	(blank)	3829	12	739	1189	742	1351	547	31	19	35	14
	A	23958	73	743	4400	4311	10276	4971	18	18	43	21
	B	3591	11	740	910	711	1461	509	25	20	41	14
	C	1062	3	739	346	201	381	134	33	19	36	13
	D	571	2	739	169	109	212	81	30	19	37	14
14	(blank)	3846	12	739	1203	735	1364	544	31	19	35	14
	A	24028	73	743	4319	4312	10389	5008	18	18	43	21
	B	3525	11	740	945	722	1370	488	27	20	39	14
	C	1036	3	738	366	196	338	136	35	19	33	13
	D	576	2	739	181	109	220	66	31	19	38	11
15	(blank)	3908	12	739	1223	749	1375	561	31	19	35	14
	A	4705	14	742	1148	791	1875	891	24	17	40	19
	B	11794	36	743	2255	2172	5050	2317	19	18	43	20
	C	7555	23	744	1187	1335	3367	1666	16	18	45	22
	D	5049	15	741	1201	1027	2014	807	24	20	40	16
16	(blank)	3862	12	739	1221	744	1350	547	32	19	35	14
		19007	58	744	3095	3370	8358	4184	16	18	44	22
	B	8355	25	741	1940	1622	3465	1328	23	19	41	16
	C	982	3	738	331	199	326	126	34	20	33	13
	D	805	2	732	427	139	182	57	53	17	23	7
17	(blank)	4218	13	739	1298	806	1507	607	31	19	36	14
	A	1296	4	735	599	264	336	97	46	20	26	7
	B	4619	14	740	1348	948	1677	646	29	21	36	14
	C	11128	34	744	1702	1954	5052	2420	15	18	45	22
	D	11750	36	743	2067	2102	5109	2472	18	18	43	21
18	(blank)	4140	13	739	1277	791	1479	593	31	19	36	14
	A	4714	14	742	1093	887	1879	855	23	19	40	18
	B	10085	31	742	2126	1940	4230	1789	21	19	42	18
	C	9142	28	744	1423	1545	4087	2087	16	17	45	23
	D	4930	15	742	1095	911	2006	918	22	18	41	19
19	(blank)	4211	13	739	1249	804	1517	641	30	19	36	15
	A	10739	33	742	2282	2189	4540	1728	21	20	42	16
	B	6027	18	742	1257	1109	2474	1187	21	18	41	20
	C	2740	8	740	737	525	1092	386	27	19	40	14
	D	6343	19	745	859	939	2818	1727	14	15	44	27
	E	2951	9	742	630	508	1240	573	21	17	42	19
20	(blank)	4699	14	740	1349	885	1748	717	29	19	37	15
	A	17060	52	744	2817	2962	7576	3705	17	17	44	22
	B	8671	26	742	1997	1694	3485	1495	23	20	40	17
	C	1880	6	740	526	382	709	263	28	20	38	14
	D	701	2	735	325	151	163	62	46	22	23	9

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-13. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 10–20—Mathematics Grade 8

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
10	(blank)	3534	10	838	1168	710	1216	440	33	20	34	12
	A	9597	28	838	2634	2483	3848	632	27	26	40	7
	B	16158	48	843	2356	3209	7772	2821	15	20	48	17
	C	4482	13	848	428	414	1816	1824	10	9	41	41
11	(blank)	3629	11	838	1204	719	1263	443	33	20	35	12
	A	14236	42	841	2809	3138	6289	2000	20	22	44	14
	B	14273	42	844	2079	2569	6547	3078	15	18	46	22
	C	1633	5	838	494	390	553	196	30	24	34	12
12	(blank)	3532	10	838	1136	704	1241	451	32	20	35	13
	A	4494	13	841	962	927	1977	628	21	21	44	14
	B	11325	34	843	1742	2214	5286	2083	15	20	47	18
	C	12826	38	842	2277	2599	5552	2398	18	20	43	19
	D	1594	5	838	469	372	596	157	29	23	37	10
13	(blank)	3434	10	838	1102	677	1221	434	32	20	36	13
	A	25010	74	843	4114	5008	11306	4582	16	20	45	18
	B	3521	10	840	867	779	1413	462	25	22	40	13
	C	1199	4	839	311	260	463	165	26	22	39	14
	D	607	2	839	192	92	249	74	32	15	41	12
14	(blank)	3491	10	838	1132	689	1225	445	32	20	35	13
	A	25136	74	843	3962	5003	11545	4626	16	20	46	18
	B	3439	10	839	964	766	1292	417	28	22	38	12
	C	1113	3	838	343	253	367	150	31	23	33	13
	D	592	2	838	185	105	223	79	31	18	38	13
15	(blank)	3494	10	838	1142	695	1218	439	33	20	35	13
	A	5245	16	841	1071	1022	2289	863	20	19	44	16
	B	11013	33	842	2055	2394	4792	1772	19	22	44	16
	C	8411	25	843	1269	1588	3937	1617	15	19	47	19
	D	5608	17	842	1049	1117	2416	1026	19	20	43	18
16	(blank)	3509	10	838	1144	691	1231	443	33	20	35	13
	A	20343	60	844	2837	3793	9497	4216	14	19	47	21
	B	8013	24	840	1838	1923	3349	903	23	24	42	11
	C	967	3	837	334	227	331	75	35	23	34	8
	D	939	3	834	433	182	244	80	46	19	26	9
17	(blank)	3705	11	838	1197	742	1302	464	32	20	35	13
	A	1249	4	835	496	309	384	60	40	25	31	5
	B	4099	12	838	1201	952	1591	355	29	23	39	9
	C	10939	32	843	1725	2166	5129	1919	16	20	47	18
	D	13779	41	843	1967	2647	6246	2919	14	19	45	21
18	(blank)	3624	11	838	1180	729	1260	455	33	20	35	13
	A	7181	21	843	1221	1333	3239	1388	17	19	45	19
	B	11239	33	842	2039	2292	4959	1949	18	20	44	17
	C	7812	23	842	1277	1638	3517	1380	16	21	45	18
	D	3915	12	841	869	824	1677	545	22	21	43	14
19	(blank)	3659	11	838	1151	733	1288	487	31	20	35	13
	A	10592	31	842	1952	2188	4584	1868	18	21	43	18
	B	6035	18	841	1226	1250	2626	933	20	21	44	15
	C	3070	9	840	692	745	1270	363	23	24	41	12
	D	7079	21	844	974	1243	3366	1496	14	18	48	21
	E	3336	10	842	591	657	1518	570	18	20	46	17
20	(blank)	3537	10	838	1159	697	1238	443	33	20	35	13
	A	18565	55	843	2684	3567	8715	3599	14	19	47	19
	B	8650	26	841	1823	1889	3629	1309	21	22	42	15
	C	2217	7	839	610	488	824	295	28	22	37	13
	D	802	2	836	310	175	246	71	39	22	31	9

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-14. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey Questions 24–36—Mathematics Grade 11

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
24	(blank)	7637	23	1134	3327	2006	2198	106	44	26	29	1
	A	2675	8	1126	2163	411	99	2	81	15	4	0
	B	4407	13	1130	2837	1185	380	5	64	27	9	0
	C	11994	36	1136	3983	4675	3269	67	33	39	27	1
	D	5928	18	1141	878	1342	3477	231	15	23	59	4
	E	906	3	1141	224	122	455	105	25	13	50	12
25	(blank)	7861	23	1133	3442	2081	2233	105	44	26	28	1
	A	2447	7	1127	1840	437	165	5	75	18	7	0
	B	2577	8	1129	1796	557	222	2	70	22	9	0
	C	3716	11	1130	2381	1050	278	7	64	28	7	0
	D	10842	32	1136	3523	4464	2836	19	32	41	26	0
	E	6104	18	1143	430	1152	4144	378	7	19	68	6
26	(blank)	8068	24	1133	3547	2130	2283	108	44	26	28	1
	A	12988	39	1132	6563	4345	2063	17	51	33	16	0
	B	8221	25	1136	2504	2534	3110	73	30	31	38	1
	C	4270	13	1141	798	732	2422	318	19	17	57	7
27	(blank)	8027	24	1133	3557	2121	2242	107	44	26	28	1
	A	7731	23	1133	3579	2427	1680	45	46	31	22	1
	B	12274	37	1137	3730	3604	4658	282	30	29	38	2
	C	5515	16	1134	2546	1589	1298	82	46	29	24	1
28	(blank)	7688	23	1133	3359	2024	2200	105	44	26	29	1
	A	5325	16	1135	2154	1649	1452	70	40	31	27	1
	B	10802	32	1136	3526	3376	3718	182	33	31	34	2
	C	7779	23	1135	3111	2221	2294	153	40	29	29	2
	D	1953	6	1129	1262	471	214	6	65	24	11	0
29	(blank)	7799	23	1133	3414	2054	2224	107	44	26	29	1
	A	11967	36	1134	5271	3365	3090	241	44	28	26	2
	B	6176	18	1135	2344	1906	1830	96	38	31	30	2
	C	4053	12	1136	1423	1246	1342	42	35	31	33	1
	D	3552	11	1137	960	1170	1392	30	27	33	39	1
30	(blank)	7814	23	1133	3414	2067	2224	109	44	26	28	1
	A	14086	42	1135	5537	4090	4175	284	39	29	30	2
	B	5990	18	1135	2330	1813	1766	81	39	30	29	1
	C	3170	9	1135	1243	972	932	23	39	31	29	1
	D	2487	7	1135	888	799	781	19	36	32	31	1
31	(blank)	7900	24	1134	3451	2078	2261	110	44	26	29	1
	A	4619	14	1135	1941	1331	1270	77	42	29	27	2
	B	8688	26	1135	3367	2667	2526	128	39	31	29	1
	C	6456	19	1136	2240	1903	2192	121	35	29	34	2
	D	5884	18	1135	2413	1762	1629	80	41	30	28	1
32	(blank)	7924	24	1133	3492	2074	2249	109	44	26	28	1
	A	17724	53	1137	5465	5617	6286	356	31	32	35	2
	B	5082	15	1133	2534	1437	1069	42	50	28	21	1
	C	1336	4	1129	865	309	159	3	65	23	12	0
	D	1481	4	1128	1056	304	115	6	71	21	8	0
33	(blank)	7918	24	1133	3475	2083	2252	108	44	26	28	1
	A	13615	41	1137	4391	4096	4842	286	32	30	36	2
	B	7589	23	1134	3158	2338	2002	91	42	31	26	1
	C	2891	9	1133	1460	832	576	23	51	29	20	1
	D	1534	5	1130	928	392	206	8	60	26	13	1

(cont'd.)

Question	Resp	NResp	%Resp	AvgSS	NSBP	NPP	NP	NPWD	%SBP	%PP	%P	%PWD
34	(blank)	7928	24	1133	3469	2092	2259	108	44	26	28	1
	A	13415	40	1136	4938	4181	4074	222	37	31	30	2
	B	7762	23	1135	2991	2265	2383	123	39	29	31	2
	C	3014	9	1135	1231	843	889	51	41	28	29	2
	D	1428	4	1131	783	360	273	12	55	25	19	1
35	(blank)	8999	27	1135	3577	2262	2892	268	40	25	32	3
	A	13735	41	1134	5824	4433	3379	99	42	32	25	1
	B	5886	18	1135	2319	1757	1767	43	39	30	30	1
	C	2519	8	1134	1108	670	708	33	44	27	28	1
	D	2408	7	1138	584	619	1132	73	24	26	47	3
36	(blank)	8037	24	1133	3515	2120	2291	111	44	26	29	1
	A	6983	21	1139	1521	1621	3514	327	22	23	50	5
	B	10178	30	1135	3508	3554	3046	70	34	35	30	1
	C	5597	17	1132	3003	1778	810	6	54	32	14	0
	D	2752	8	1129	1865	668	217	2	68	24	8	0

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-15. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 21–32—Writing Grade 5

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
21	(blank)	3579	11	541	859	922	955	843	24	26	27	24
	A	9779	30	544	1645	2635	2911	2588	17	27	30	26
	B	14496	45	547	1681	3613	4443	4759	12	25	31	33
	C	4058	13	544	713	1127	1202	1016	18	28	30	25
	D	153	0	531	75	33	25	20	49	22	16	13
22	(blank)	3609	11	541	864	926	963	856	24	26	27	24
	A	17040	53	546	2386	4467	5194	4993	14	26	30	29
	B	10543	33	547	1400	2682	3186	3275	13	25	30	31
	C	746	2	535	257	221	180	88	34	30	24	12
	D	127	0	529	66	34	13	14	52	27	10	11
23	(blank)	3610	11	541	871	928	959	852	24	26	27	24
	A	23244	72	546	3051	5985	7161	7047	13	26	31	30
	B	3318	10	543	631	912	961	814	19	27	29	25
	C	1317	4	543	269	374	307	367	20	28	23	28
	D	576	2	542	151	131	148	146	26	23	26	25
24	(blank)	3736	12	541	877	964	1003	892	23	26	27	24
	A	3934	12	541	916	1184	999	835	23	30	25	21
	B	6382	20	544	1060	1756	1869	1697	17	28	29	27
	C	13664	43	548	1555	3285	4307	4517	11	24	32	33
	D	4349	14	546	565	1141	1358	1285	13	26	31	30
25	(blank)	4176	13	541	968	1088	1139	981	23	26	27	23
	A	1613	5	536	535	478	375	225	33	30	23	14
	B	3284	10	542	713	977	879	715	22	30	27	22
	C	10045	31	546	1286	2613	3133	3013	13	26	31	30
	D	12947	40	548	1471	3174	4010	4292	11	25	31	33
26	(blank)	3774	12	541	903	988	1000	883	24	26	26	23
	A	5763	18	543	1048	1635	1648	1432	18	28	29	25
	B	7397	23	546	1113	1932	2182	2170	15	26	29	29
	C	11663	36	547	1345	2846	3676	3796	12	24	32	33
	D	3468	11	545	564	929	1030	945	16	27	30	27
27	(blank)	3803	12	541	907	994	1010	892	24	26	27	23
	A	6504	20	544	1126	1754	1908	1716	17	27	29	26
	B	6120	19	545	1009	1635	1791	1685	16	27	29	28
	C	9277	29	547	1039	2304	2915	3019	11	25	31	33
	D	6361	20	546	892	1643	1912	1914	14	26	30	30
28	(blank)	3835	12	541	892	1015	1023	905	23	26	27	24
	A	12766	40	548	1353	2950	4019	4444	11	23	31	35
	B	7993	25	541	1631	2549	2261	1552	20	32	28	19
	C	2064	6	541	440	644	579	401	21	31	28	19
	D	4439	14	550	432	882	1375	1750	10	20	31	39
	E	968	3	540	225	290	279	174	23	30	29	18
29	(blank)	3840	12	541	913	1017	1022	888	24	26	27	23
	A	9811	31	548	1193	2359	2977	3282	12	24	30	33
	B	6100	19	544	1019	1659	1804	1618	17	27	30	27
	C	6957	22	545	1000	1893	2089	1975	14	27	30	28
	D	5357	17	545	848	1402	1644	1463	16	26	31	27
30	(blank)	3867	12	541	924	1018	1018	907	24	26	26	23
	A	9663	30	546	1451	2503	2884	2825	15	26	30	29
	B	5779	18	545	907	1581	1654	1637	16	27	29	28
	C	6682	21	546	897	1680	2035	2070	13	25	30	31
	D	6074	19	546	794	1548	1945	1787	13	25	32	29

(cont'd)

Question	Resp	NResp	%Resp	AvgSS	NSBP	NPP	NP	NPWD	%SBP	%PP	%P	%PWD
31	(blank)	3936	12	541	923	1036	1043	934	23	26	26	24
	A	3498	11	544	660	908	953	977	19	26	27	28
	B	3980	12	544	745	1052	1051	1132	19	26	26	28
	C	7040	22	547	873	1765	2245	2157	12	25	32	31
	D	13611	42	546	1772	3569	4244	4026	13	26	31	30
32	(blank)	4255	13	542	954	1146	1148	1007	22	27	27	24
	A	6627	21	543	1122	1954	1963	1588	17	29	30	24
	B	2654	8	542	535	818	778	523	20	31	29	20
	C	2922	9	544	503	823	843	753	17	28	29	26
	D	15607	49	548	1859	3589	4804	5355	12	23	31	34

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-16. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey Questions 21–32—Writing Grade 8

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
21	(blank)	3534	11	836	987	1158	1126	263	28	33	32	7
	A	5806	17	838	1214	2109	2108	375	21	36	36	6
	B	18251	54	842	1915	6111	8335	1890	10	33	46	10
	C	5674	17	841	801	1988	2317	568	14	35	41	10
	D	298	1	831	131	94	64	9	44	32	21	3
22	(blank)	3572	11	836	991	1175	1145	261	28	33	32	7
	A	13796	41	841	1814	4842	5947	1193	13	35	43	9
	B	14517	43	842	1652	4806	6464	1595	11	33	45	11
	C	1420	4	834	453	561	356	50	32	40	25	4
	D	258	1	828	138	76	38	6	53	29	15	2
23	(blank)	3610	11	836	1002	1210	1143	255	28	34	32	7
	A	25707	77	842	2933	8784	11445	2545	11	34	45	10
	B	2806	8	838	666	992	945	203	24	35	34	7
	C	1000	3	836	285	336	302	77	29	34	30	8
	D	440	1	834	162	138	115	25	37	31	26	6
24	(blank)	3694	11	836	1022	1225	1175	272	28	33	32	7
	A	3432	10	838	716	1309	1184	223	21	38	34	6
	B	6971	21	841	1001	2333	2954	683	14	33	42	10
	C	15241	45	842	1620	5086	6951	1584	11	33	46	10
	D	4225	13	840	689	1507	1686	343	16	36	40	8
25	(blank)	3716	11	837	1016	1237	1188	275	27	33	32	7
	A	1586	5	834	500	603	416	67	32	38	26	4
	B	3191	10	839	683	1125	1149	234	21	35	36	7
	C	12213	36	842	1409	4112	5414	1278	12	34	44	10
	D	12857	38	842	1440	4383	5783	1251	11	34	45	10
26	(blank)	3708	11	837	1012	1237	1190	269	27	33	32	7
	A	4824	14	839	819	1842	1839	324	17	38	38	7
	B	5749	17	840	1010	2062	2285	392	18	36	40	7
	C	11708	35	842	1341	3961	5197	1209	11	34	44	10
	D	7574	23	843	866	2358	3439	911	11	31	45	12
27	(blank)	3724	11	837	1016	1235	1199	274	27	33	32	7
	A	4620	14	839	858	1709	1734	319	19	37	38	7
	B	6468	19	840	1003	2275	2666	524	16	35	41	8
	C	10991	33	842	1205	3703	4918	1165	11	34	45	11
	D	7760	23	842	966	2538	3433	823	12	33	44	11
28	(blank)	3719	11	837	995	1256	1199	269	27	34	32	7
	A	12179	36	843	1220	4028	5688	1243	10	33	47	10
	B	5797	17	836	1428	2557	1634	178	25	44	28	3
	C	2318	7	838	474	918	819	107	20	40	35	5
	D	8275	25	845	603	2190	4226	1256	7	26	51	15
	E	1275	4	836	328	511	384	52	26	40	30	4
29	(blank)	3826	11	836	1056	1274	1225	271	28	33	32	7
	A	10188	30	843	1014	3134	4805	1235	10	31	47	12
	B	6500	19	841	966	2228	2674	632	15	34	41	10
	C	7315	22	840	1074	2703	2977	561	15	37	41	8
	D	5734	17	840	938	2121	2269	406	16	37	40	7
30	(blank)	3811	11	837	1032	1275	1226	278	27	33	32	7
	A	6684	20	840	1055	2371	2687	571	16	35	40	9
	B	6120	18	841	965	2069	2500	586	16	34	41	10
	C	8167	24	842	998	2770	3595	804	12	34	44	10
	D	8781	26	842	998	2975	3942	866	11	34	45	10

(cont'd)

Question	Resp	NResp	%Resp	AvgSS	NSBP	NPP	NP	NPWD	%SBP	%PP	%P	%PWD
31	(blank)	3942	12	837	1059	1325	1278	280	27	34	32	7
	A	4160	12	841	569	1390	1812	389	14	33	44	9
	B	5071	15	841	758	1711	2097	505	15	34	41	10
	C	8036	24	842	933	2656	3589	858	12	33	45	11
	D	12354	37	841	1729	4378	5174	1073	14	35	42	9
32	(blank)	4283	13	837	1094	1446	1438	305	26	34	34	7
	A	3764	11	837	834	1575	1195	160	22	42	32	4
	B	4973	15	839	878	1901	1829	365	18	38	37	7
	C	3784	11	839	648	1406	1454	276	17	37	38	7
	D	16759	50	843	1594	5132	8034	1999	10	31	48	12

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Table P-17. 2008–09 NECAP: Average Scaled Score and Counts and Percentages within Performance Levels of Responses to Student Survey, Questions 1–12—Writing Grade 11

Question	*Resp*	*NResp*	*%Resp*	*AvgSS*	*NSBP*	*NPP*	*NP*	*NPWD*	*%SBP*	*%PP*	*%P*	*%PWD*
1	(blank)	7416	22	5.8	1052	3507	2569	288	14	47	35	4
	A	4361	13	5	879	2537	903	42	20	58	21	1
	B	15222	45	6.1	1122	8024	5669	407	7	53	37	3
	C	6519	19	6.8	314	2407	3322	476	5	37	51	7
2	(blank)	7449	22	5.8	1071	3527	2566	285	14	47	34	4
	A	6757	20	5.8	656	3684	2274	143	10	55	34	2
	B	15640	47	6.3	1054	7438	6495	653	7	48	42	4
	C	3672	11	5.6	586	1826	1128	132	16	50	31	4
3	(blank)	7357	22	5.8	1045	3487	2544	281	14	47	35	4
	A	19391	58	6.1	1531	9809	7395	656	8	51	38	3
	B	4027	12	6.1	393	1889	1577	168	10	47	39	4
	C	2056	6	6	253	947	773	83	12	46	38	4
	D	687	2	5.3	145	343	174	25	21	50	25	4
4	(blank)	7479	22	5.8	1061	3560	2577	281	14	48	34	4
	A	3253	10	6	398	1523	1198	134	12	47	37	4
	B	6661	20	6.1	597	3229	2575	260	9	48	39	4
	C	12621	38	6.2	865	6332	4969	455	7	50	39	4
	D	3504	10	5.7	446	1831	1144	83	13	52	33	2
5	(blank)	7676	23	5.8	1104	3656	2628	288	14	48	34	4
	A	3245	10	5.7	414	1736	1010	85	13	53	31	3
	B	3554	11	5.7	412	1913	1145	84	12	54	32	2
	C	8332	25	6.2	622	4160	3252	298	7	50	39	4
	D	10711	32	6.3	815	5010	4428	458	8	47	41	4
6	(blank)	8042	24	5.8	1146	3909	2697	290	14	49	34	4
	A	1644	5	5.6	253	849	494	48	15	52	30	3
	B	3666	11	5.9	422	1851	1274	119	12	50	35	3
	C	10013	30	6.3	666	4753	4171	423	7	47	42	4
	D	10153	30	6.1	880	5113	3827	333	9	50	38	3
7	(blank)	7499	22	5.8	1052	3571	2593	283	14	48	35	4
	A	2489	7	5.3	434	1363	637	55	17	55	26	2
	B	4516	13	5.8	549	2398	1435	134	12	53	32	3
	C	9421	28	6.1	721	4721	3663	316	8	50	39	3
	D	9593	29	6.4	611	4422	4135	425	6	46	43	4
8	(blank)	7382	22	5.8	1026	3510	2560	286	14	48	35	4
	A	8360	25	6.1	671	4124	3234	331	8	49	39	4
	B	4491	13	5.3	695	2634	1095	67	15	59	24	1
	C	1610	5	5.5	234	885	468	23	15	55	29	1
	D	10076	30	6.6	471	4457	4667	481	5	44	46	5
	E	1599	5	5.4	270	865	439	25	17	54	27	2
9	(blank)	7493	22	5.8	1070	3558	2583	282	14	47	34	4
	A	7819	23	6.4	491	3574	3396	358	6	46	43	5
	B	5352	16	6.1	421	2714	2013	204	8	51	38	4
	C	6536	19	5.9	635	3437	2290	174	10	53	35	3
	D	6318	19	5.8	750	3192	2181	195	12	51	35	3
10	(blank)	7565	23	5.8	1082	3594	2604	285	14	48	34	4
	A	4902	15	5.9	538	2504	1702	158	11	51	35	3
	B	4616	14	6	473	2340	1643	160	10	51	36	3
	C	6935	21	6.2	527	3452	2712	244	8	50	39	4
	D	9500	28	6.2	747	4585	3802	366	8	48	40	4

(cont'd)

11	(blank)	7654	23	5.8	1100	3644	2624	286	14	48	34	4
	A	2699	8	6.2	243	1220	1109	127	9	45	41	5
	B	3775	11	6.1	327	1881	1418	149	9	50	38	4

Question	Resp	NResp	%Resp	AvgSS	NSBP	NPP	NP	NPWD	%SBP	%PP	%P	%PWD
	C	6345	19	6.2	527	3071	2489	258	8	48	39	4
	D	13045	39	6	1170	6659	4823	393	9	51	37	3
	(blank)	7770	23	5.8	1119	3692	2668	291	14	48	34	4
	A	1343	4	5.2	246	758	317	22	18	56	24	2
12	B	7816	23	6.2	578	3923	3026	289	7	50	39	4
	C	4128	12	5.9	474	2047	1458	149	11	50	35	4
	D	12461	37	6.2	950	6055	4994	462	8	49	40	4

SBP = Substantially Below Proficient; PP = Partially Proficient; P = Proficient; PWD = Proficient with Distinction.

Grades 3–8 NECAP Student Questionnaire

Grades 3 – 8 NECAP Student Questionnaire—October 2008

Reading Questions

1. How difficult was the reading test?
A. harder than my regular reading schoolwork
B. about the same as my regular reading schoolwork
C. easier than my regular reading schoolwork

2. How interesting were the reading passages?
A. All of the passages were interesting to me.
B. Most of the passages were interesting to me.
C. Most of the passages were not interesting to me.
D. None of the passages were interesting to me.

3. How hard did you try on the reading test?
A. I tried harder on this test than I do on my regular reading schoolwork.
B. I tried about the same as I do on my regular reading schoolwork.
C. I did not try as hard on this test as I do on my regular reading schoolwork.

4. How difficult were the reading passages on the test?
A. Most of the passages were more difficult than what I normally read for school.
B. Most of the passages were about the same as what I normally read for school.
C. Most of the passages were easier than what I normally read for school.

5. Did you have enough time to answer all of the questions on the reading test?
A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

6. How often do you have Language Arts/Reading homework?
A. almost every day
B. a few times a week
C. a few times a month
D. I usually don't have homework in Language Arts/Reading.

7. When I am reading and come to a word I do not know, I usually
A. figure it out myself.
B. ask someone what the word is.
C. skip the word.
D. stop reading.

8. How often do you choose to read in your free time?

A. almost every day
B. a few times a week
C. a few times a month
D. I almost never read.

9. How do you most often find information about things that interest you?

A. I use a computer.
B. I look in books, magazines, or newspapers.
C. I ask someone.
D. I watch TV or videos.

Mathematics Questions

10. How difficult was the mathematics test?

A. harder than my regular mathematics schoolwork
B. about the same as my regular mathematics schoolwork
C. easier than my regular mathematics schoolwork

11. How hard did you try on the mathematics test?

A. I tried harder on this test than I do on my regular mathematics schoolwork.
B. I tried about the same as I do on my regular mathematics schoolwork.
C. I did not try as hard on this test as I do on my regular mathematics schoolwork.

12. How much did you use a calculator on the test?

A. If it was allowed, I used it on most questions.
B. If it was allowed, I used it on some questions.
C. I didn't use it on very many questions.
D. I didn't have a calculator.

13. Did you have enough time to answer all of the questions on the non-calculator session on the mathematics test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

14. Did you have enough time to answer all of the questions on the calculator sessions of the mathematics test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

15. How often do you work with other students in small groups on problem-solving tasks in mathematics class?

A. almost every day
B. a few times a week
C. a few times a month
D. never or almost never

16. How often do you have mathematics homework?

A. almost every day
B. a few times a week
C. a few times a month
D. I usually don't have homework in mathematics.

17. How often do you use hands-on materials such as base-ten blocks, cubes, rods, counters, geoboards, and tangrams in mathematics class?

A. almost every day
B. a few times a week
C. a few times a month
D. a few times a year or less

18. How often do you use a calculator in mathematics class?

A. almost every day
B. a few times a week
C. a few times a month
D. a few times a year or less

19. How do you spend most of your time in mathematics class?

A. I listen to my teacher
B. I work by myself.
C. I work in small groups.
D. I do some work myself and some in small groups.
E. The whole class works together.

20. In mathematics class, how often are you asked to explain how you solved a problem?

A. almost every day
B. a few times a week
C. a few times a month
D. a few times a year or less

<u>Writing Questions (Grades 5 and 8 only)</u>

21. How difficult was the writing test?

A. harder than my regular writing schoolwork
B. about the same as my regular writing schoolwork
C. easier than my regular writing schoolwork
D. I did not take the writing test.

22. How hard did you try on the writing test?

A. I tried harder on this test than I do on my regular schoolwork.
B. I tried about the same as I do on my regular schoolwork.
C. I did not try as hard on this test as I do on my regular schoolwork.
D. I did not take the writing test.

23. Did you have enough time to answer all of the questions on the writing test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

24. How often are you asked to write at least one paragraph for Reading/Language Arts class?

A. more than once a day
B. once a day
C. a few times a week
D. less than once a week

25. How often are you asked to write at least one paragraph for Science class?

A. more than once a day
B. once a day
C. a few times a week
D. less than once a week

26. How often are you asked to use writing to explain your mathematical ideas?

A. more than once a day
B. once a day
C. a few times a week
D. less than once a week

27. I choose my own topics for writing

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

28. I know how to revise my writing to improve it

A. on my own.
B. with my teacher's help.
C. with help from my family or friends.
D. by using all of the above.
E. but I rarely revise my writing.

29. I write more than one draft

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

30. I discuss my rough drafts with the teacher

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

31. I discuss my rough drafts with other students

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

32. What kinds of writing do you do most in school?

A. I mostly write stories.
B. I mostly write reports.
C. I mostly write about things I've read.
D. I do all kinds of writing.

Thank you very much for all of your hard work during testing and for answering these questions..

Grade 11 NECAP Student Questionnaire – October 2008

Writing Questions

1. How difficult was the writing test?
- A. harder than my regular writing work
- B. about the same as my regular writing work
- C. easier than my regular writing work

2. How hard did you try on the writing test?
- A. I tried harder on this test than I do on my regular work.
- B. I tried about the same as I do on my regular work.
- C. I did not try as hard on this test as I do on my regular work.

3. Did you have enough time to complete the prompts on the writing test?
- A. I had enough time to complete the prompts and check my work.
- B. I had enough time to complete the prompts, but I did not have time to check my work.
- C. I felt rushed, but I was able to complete the prompts.
- D. I did not have enough time to complete the prompts.

4. How often are you asked to write at least one paragraph in English class?
- A. more than once a day
- B. once a day
- C. a few times a week
- D. less than once a week

5. How often are you asked to use writing to explain your mathematical ideas?
- A. more than once a day
- B. once a day
- C. a few times a week
- D. less than once a week

6. How often are you asked to write at least one paragraph in Science class?
- A. more than once a day
- B. once a day
- C. a few times a week
- D. less than once a week

7. I choose my own topics for writing
- A. almost always.
- B. more than half the time.
- C. about half the time.
- D. less than half the time.

8. I know how to revise my writing to improve it
- A. on my own.
- B. with my teacher's help.
- C. with help from my family or friends.
- D. by using all of the above.
- E. but I rarely revise my writing.

9. I write more than one draft

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

10. I discuss my rough drafts with the teacher

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

11. I discuss my rough drafts with other students

A. almost always.
B. more than half the time.
C. about half the time.
D. less than half the time.

12. What kinds of writing do you do most in school?

A. I mostly write narratives/poems.
B. I mostly write reports/persuasive pieces.
C. I mostly write about things I've read.
D. I do all kinds of writing.

Reading Questions

13. How difficult was the reading test?

A. harder than my regular reading work
B. about the same as my regular reading work
C. easier than my regular reading work

14. How interesting were the reading passages?

A. All of the passages were interesting to me.
B. Most of the passages were interesting to me.
C. Most of the passages were not interesting to me.
D. None of the passages were interesting to me.

15. How hard did you try on the reading test?

A. I tried harder on this test than I do on my regular reading work.
B. I tried about the same as I do on my regular reading work.
C. I did not try as hard on this test as I do on my regular reading work.

16. How difficult were the reading passages on the test?

A. Most of the passages were more difficult than what I normally read for school.
B. Most of the passages were about the same as what I normally read for school.
C. Most of the passages were easier than what I normally read for school.

17. Did you have enough time to answer all of the questions on the reading test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

18. How often do you have reading homework in English class?

A. almost every day
B. a few times a week
C. a few times a month
D. I usually don't have reading homework in English class.

19. How often do you have reading homework in other subject areas?

A. almost every day
B. a few times a week
C. a few times a month
D. I usually don't have reading homework in other subject areas.

20. How do you most often learn new vocabulary words?

A. I am taught new vocabulary words in most of my courses.
B. I am taught new vocabulary words mostly in my English class.
C. I learn new vocabulary words on my own using a dictionary or computer.
D. I rarely learn new vocabulary words.

21. How often do you choose to read in your free time?

A. almost every day
B. a few times a week
C. a few times a month
D. I almost never read.

22. How do you most often find information about things that interest you?

A. I use a computer.
B. I look in books, magazines, or newspapers.
C. I ask someone.
D. I watch TV or videos.

23. What grade did you receive in the last English course you completed?

A. A
B. B
C. C
D. lower than C

Mathematics Questions

24. What best describes the last mathematics course you completed?

A. General Mathematics or pre-Algebra
B. Algebra I or Integrated Mathematics I
C. Geometry or Integrated Mathematics II
D. Algebra II or Integrated Mathematics III
E. Pre-Calculus/Advanced Mathematics or Higher

25. What best describes the mathematics course you are currently taking or will be taking this year?

A. General Mathematics or pre-Algebra
B. Algebra I or Integrated Mathematics I
C. Geometry or Integrated Mathematics II
D. Algebra II or Integrated Mathematics III
E. Pre-Calculus/Advanced Mathematics or Higher

26. How difficult was the mathematics test compared to your current or most recent mathematics class?

A. more difficult
B. about the same
C. less difficult

27. How hard did you try on the mathematics test compared to your current or most recent mathematics class?

A. I tried harder on this test.
B. I tried about the same.
C. I did not try as hard on this test.

28. How much did you use a calculator on the test?

A. When it was allowed, I used it on most questions.
B. When it was allowed, I used it on some questions.
C. I didn't use it on very many questions.
D. I didn't have a calculator.

29. Did you have enough time to answer all of the questions on the non-calculator session of the mathematics test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

30. Did you have enough time to answer all of the questions on the calculator session of the mathematics test?

A. I had enough time to answer all of the questions and check my work.
B. I had enough time to answer all of the questions, but I did not have time to check my work.
C. I felt rushed, but I was able to answer all of the questions.
D. I did not have enough time to answer all of the questions.

31. How often do you work in groups with other students on problem-solving tasks in mathematics?

A. almost every day
B. a few times a week
C. a few times a month
D. never or almost never

32. How often do you have mathematics homework assignments?

A. almost every day
B. a few times a week
C. a few times a month
D. I usually don't have homework in mathematics.

33. How often do you use a calculator in mathematics class?

A. almost every day
B. a few times a week
C. a few times a month
D. a few times a year or less

34. In mathematics class, how often are you asked to explain how you solved a problem?

A. almost every day
B. a few times a week
C. a few times a month
D. a few times a year or less

35. If you left any questions blank on the mathematics test, which statement best describes the reason?

A. I didn't know the mathematics needed to answer the questions.
B. I knew the mathematics, but have never been given questions like these before.
C. Answering the questions would take too much effort.
D. I didn't have enough time to answer the questions.

36. What grade did you receive in the last mathematics course you completed?

A. A
B. B
C. C
D. lower than C

Thank you very much for all of your hard work during testing and for answering these questions

APPENDIX Q—SAMPLE REPORTS

Technical Report — Appendix Q: Sample Reports

Report	Grades Available	Teaching Year & Testing Year	Sample Report Included
Student Report	3-8, 11	No	Grade 5 & 11, testing year
Item Analysis: Reading	3-8, 11	Yes	Grade 11, testing year
Item Analysis: Mathematics	3-8, 11	Yes	Grade 5, testing year
Item Analysis: Writing	5, 8, 11	Yes	Grade 5 & 11, testing year
School Results Report	3-8, 11	Yes	Grade 11, testing year
School Summary Report	One summary of all grades in a school	Yes	All grades, testing year
District Results Report	3-8, 11	Yes	Grade 5, testing year
District Summary Report	One summary of all grades in a school	Yes	All grades, testing year

NECAP Student Report - Fall 2008

This report contains results from the Fall 2008 New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. The NECAP tests are designed to measure student performance on grade level expectations (GLE) developed and adopted by the three states. Specifically, the tests are designed to measure the content and skills that students are expected to have as they begin the current enrolled grade. In other words, content and skills which students have learned through the end of the previous grade.

NECAP test results are used primarily for school improvement and accountability. Achievement level results are used in the state accountability system required under No Child Left Behind. More detailed school and district results are used by schools to help improve curriculum and instruction. Individual student results are used to support information gathered through classroom instruction and assessments. Contact the school for more information on this student's overall achievement.

Achievement Levels and Corresponding Score Ranges

Student performance on the NECAP tests is classified into one of four achievement levels describing students' level of proficiency on the content and skills required through the end of the previous grade. Performance at Proficient or Proficient with Distinction indicates that the student has a level of proficiency necessary to begin working successfully on current grade content and skills. Performance below Proficient suggests that additional instruction and student work may be needed on the previous grade content and skills as the student is introduced to new content and skills at the current grade. Refer to the Achievement Level Descriptions contained in this report for a more detailed description of the achievement levels.

There is a wide range of student proficiency within each achievement level. NECAP test results are also reported as scaled scores to provide additional information about the location of student performance within each achievement level. NECAP scores are reported as three-digit scores in which the first digit represents the grade level. The remaining digits range from 00 to 80. Scores of 40 and higher indicate a level of proficiency at or above the Proficient level. Scores below 40 indicate proficiency below the Proficient level. For example, scores of 340 at grade 3, 540 at grade 5, and 740 at grade 7 each indicate Proficient performance at each grade level.

Comparisons to Other Beginning of Grade Students

The tables in the middle section of the report provide the percentage of students performing at each achievement level in the student's school, district, and statewide. Note that one or two students can have a large impact on percentages in small schools and districts. Results are not reported for schools or districts with nine (9) or fewer students.

Performance in Content Area Subcategories

This section of the report provides information about student performance on sets of items measuring particular content and skills within each test. These results can provide a general idea of relative strengths and weaknesses in comparison to other students. However, results in this section are based on small numbers of test items and should be interpreted cautiously.

Students at Proficient Level

This column shows the average performance on these items of students who performed near the beginning of the Proficient achievement level on the overall test. Students whose performance in a category falls within the range shown performed similarly to those students. This comparison can provide some information about the level of performance needed to perform at the Proficient level.

Comments about this student's writing performance

Students in grades 5 and 8 took the NECAP writing test which included a writing prompt that required students to produce a written response up to three pages long. Student responses were scored independently by two scorers. Each scorer was able to choose up to three comments from a prepared list to provide feedback about each student's performance on the writing prompt. If both scorers selected the same comment, it is listed only once.

Achievement Level Descriptions

Proficient with Distinction (Level 4) - Students performing at this level demonstrate the prerequisite knowledge and skills needed to participate and excel in instructional activities aligned with the GLE at the current grade level. Errors made by these students are few and minor and do not reflect gaps in prerequisite knowledge and skills.

Proficient (Level 3) - Students performing at this level demonstrate minor gaps in the prerequisite knowledge and skills needed to participate and perform successfully in instructional activities aligned with the GLE at the current grade level. It is likely that any gaps in prerequisite knowledge and skills demonstrated by these students can be addressed during the course of typical classroom instruction.

Partially Proficient (Level 2) - Students performing at this level demonstrate gaps in prerequisite knowledge and skills needed to participate and perform successfully in instructional activities aligned with the GLE at the current grade level. Additional instructional support may be necessary for these students to meet grade level expectations.

Substantially Below Proficient (Level 1) - Students performing at this level demonstrate extensive and significant gaps in prerequisite knowledge and skills needed to participate and perform successfully in instructional activities aligned with the GLE at the current grade level. Additional instructional support is necessary for these students to meet grade level expectations.

Student	Grade 05	School	District	State

Fall 2008 - Beginning of Grade 5 NECAP Test Results

Content Area	Achievement Level	Scaled Score	This Student's Achievement Level and Score
Reading			Below (500–530) · Partial (530–540) · Proficient (540–556) · Distinction (556–580)

Content Area	Achievement Level	Scaled Score	This Student's Achievement Level and Score
Mathematics			Below (500–533) · Partial (533–540) · Proficient (540–554) · Distinction (554–580)

Content Area	Achievement Level	Scaled Score	This Student's Achievement Level and Score
Writing			Below (500–528) · Partial (528–540) · Proficient (540–555) · Distinction (555–580)

Interpretation of Graphic Display

The line (▌) represents the student's score. The bar (▬) surrounding the score represents the probable range of scores for the student if he or she were to be tested many times. This statistic is called the standard error of measurement. See the reverse side for the achievement level descriptions.

This Student's Achievement Level Compared to Other Beginning of Grade 5 Students by School, District, and State

	Reading				Mathematics				Writing			
	Student	School	District	State	Student	School	District	State	Student	School	District	State
Proficient with Distinction												
Proficient												
Partially Proficient												
Substantially Below Proficient												

This Student's Performance in Content Area Subcategories

Reading		Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Word ID/ Vocabulary		10					
Type of Text*	Literary	21					
	Informational	21					
Level of Comprehension*	Initial Understanding	18					
	Analysis and Interpretation	24					

Mathematics	Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Numbers and Operations	30					
Geometry and Measurement	13					
Functions and Algebra	13					
Data, Statistics, and Probability	10					

Writing	Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Structures of Language & Writing Conventions	10					
Short Responses	12					
Extended Response	15					

Comments about this student's writing performance:

*With the exception of Word ID/Vocabulary items, reading items are reported in two ways - Type of Text and Level of Comprehension.

NECAP Student Report - Fall 2008

This report contains results from the Fall 2008 New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. The NECAP tests are designed to measure student performance on grade span expectations (GSE) developed and adopted by the three states. Specifically, the tests are designed to measure the content and skills that students are expected to have as they begin the current enrolled grade. In other words, content and skills which students have learned through the end of the previous grade.

NECAP test results are used primarily for school improvement and accountability. Achievement level results are used in the state accountability system required under No Child Left Behind. More detailed school and district results are used by schools to help improve curriculum and instruction. Individual student results are used to support information gathered through classroom instruction and assessments. Contact the school for more information on this student's overall achievement.

Achievement Levels and Corresponding Score Ranges

Student performance on the NECAP tests is classified into one of four achievement levels describing students' level of proficiency on the content and skills required through the end of the previous grade. Performance at Proficient or Proficient with Distinction indicates that the student has a level of proficiency necessary to begin working successfully on current grade content and skills. Performance below Proficient suggests that additional instruction and student work may be needed on the previous grade content and skills as the student is introduced to new content and skills at the current grade. Refer to the Achievement Level Descriptions contained in this report for a more detailed description of the achievement levels.

There is a wide range of student proficiency within each achievement level. NECAP test results are also reported as scaled scores to provide additional information about the location of student performance within each achievement level. Grade 11 NECAP scores are reported as four-digit scores in which the first two digits represent the grade level. The remaining digits range from 00 to 80. Scores of 40 and higher indicate a level of proficiency at or above the Proficient level. Scores below 40 indicate proficiency below the Proficient level. For example, a score of 1140 indicates Proficient performance at this grade level. The writing score is reported as the total points earned on the NECAP scoring rubric for writing. This rubric describes the most important features expected in student writing.

Comparisons to Other Beginning of Grade Students

The tables in the middle section of the report provide the percentage of students performing at each achievement level in the student's school, district, and statewide. Note that one or two students can have a large impact on percentages in small schools and districts. Results are not reported for schools or districts with nine (9) or fewer students.

Performance in Content Area Subcategories

This section of the report provides information about student performance on sets of items measuring particular content and skills within each test. These results can provide a general idea of relative strengths and weaknesses in comparison to other students. However, results in this section are based on small numbers of test items and should be interpreted cautiously.

Students at Proficient Level

This column shows the average performance on these items of students who performed near the beginning of the Proficient achievement level on the overall test. Students whose performance in a category falls within the range shown performed similarly to those students. This comparison can provide some information about the level of performance needed to perform at the Proficient level.

Comments about this student's writing performance

Students in grade 11 took the NECAP writing test which required students to produce a written response up to three pages long. Student responses were scored independently by two scorers. Each scorer was able to choose up to three comments from a prepared list to provide feedback about each student's performance on the writing prompt. If both scorers selected the same comment, it is listed only once.

Achievement Level Descriptions

Proficient with Distinction (Level 4) - Students performing at this level demonstrate the prerequisite knowledge and skills needed to participate and excel in instructional activities aligned with the grade 9-10 GSEs. Errors made by these students are few and minor and do not reflect gaps in prerequisite knowledge and skills.

These students are prepared to perform successfully in classroom instruction aligned with grade 11-12 expectations.

Proficient (Level 3) - Students performing at this level demonstrate minor gaps in the knowledge and skills needed to participate and perform successfully in instructional activities aligned with the grade 9-10 GSEs.

It is likely that any gaps in the prerequisite knowledge and skills demonstrated by these students can be addressed by the classroom teacher during the course of classroom instruction aligned with grade 11-12 expectations.

Partially Proficient (Level 2) - Students performing at this level demonstrate gaps in the knowledge and skills needed to participate and perform successfully in instructional activities aligned with the grade 9-10 GSEs.

Additional instructional support may be necessary for these students to perform successfully in courses aligned with grade 11-12 expectations.

Substantially Below Proficient (Level 1) - Students performing at this level demonstrate extensive and significant gaps in the prerequisite knowledge and skills needed to participate and perform successfully in instructional activities aligned with the grade 9-10 GSEs.

Additional instruction and support is necessary for these students to meet the grade 9-10 GSEs.

Student	Grade 11	School	District	State

Fall 2008 - Beginning of Grade 11 NECAP Test Results

Content Area	Achievement Level	Scaled Score	This Student's Achievement Level and Scaled Score
Reading			Below / Partial / Proficient / Distinction 1100 — 1130 — 1140 — 1154 — 1180

Content Area	Achievement Level	Scaled Score	This Student's Achievement Level and Scaled Score
Mathematics			Below / Partial / Proficient / Distinction 1100 — 1134 — 1140 — 1152 — 1180

Content Area	Achievement Level	Total Points	This Student's Achievement Level and Total Points
Writing			Below / Partial / Proficient / Distinction 2 — 4 — 7 — 10 — 12

Interpretation of Graphic Display

The line (▌) represents the student's score. The bar (▬) surrounding the score represents the probable range of scores for the student if he or she were to be tested many times. This statistic is called the standard error of measurement. See the reverse side for the achievement level descriptions.

This Student's Achievement Level Compared to Other Beginning of Grade 11 Students by School, District, and State

	Reading				Mathematics				Writing			
	Student	School	District	State	Student	School	District	State	Student	School	District	State
Proficient with Distinction												
Proficient												
Partially Proficient												
Substantially Below Proficient												

This Student's Performance in Content Area Subcategories

Reading		Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Word ID/ Vocabulary		10					
Type of Text*	Literary	21					
	Informational	21					
Level of Comprehension*	Initial Understanding	18					
	Analysis and Interpretation	24					

Mathematics	Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Numbers and Operations	10					
Geometry and Measurement	19					
Functions and Algebra	26					
Data, Statistics, and Probability	9					

Comments about this student's writing performance:

Writing	Possible Points	Student	Average Points Earned: School	District	State	Students at Proficient Level
Extended Response	12					

*With the exception of Word ID/Vocabulary items, reading items are reported in two ways - Type of Text and Level of Comprehension.

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Item Analysis Report
Reading

School:
District:
State:
Code:

	Released Items																	Total Test Results							
Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
Content Strand	WV	WV	WV	II	WV	II	II	IA	IA	IA	IA	IA	WV	II	II	II	II	Word ID/ Vocabulary	Literary	Informational	Initial Understanding	Analysis & Interpretation			
GSE Code	10-2	10-3	10-3	10-7	10-3	10-7	10-7	10-8	10-8	10-8	10-8	10-8	10-3	10-7	10-7	10-7	10-7								
Depth of Knowledge Code	1	2	1	2	1	2	2	2	2	2	2	3	1	2	1	1	2								
Item Type	MC	MC	MC	MC	MC	MC	CR	MC	MC	MC	MC	CR	MC	MC	MC	MC	CR								
Correct MC Response	C	D	C	B	A	C		B	D	D	B		A	D	C	A									
Name/Student ID — Total Possible Points	1	1	1	1	1	1	4	1	1	1	1	4	1	1	1	1	4	10	21	21	18	24	52		

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Percent Correct/Average Score: School																	
Percent Correct/Average Score: District																	
Percent Correct/Average Score: State																	

LEGEND FOR THE ITEM ANALYSIS REPORT - GRADE 11 READING

Released Items Section

Released Item Number: This number corresponds to the item number in the released item documents. This report provides complete data on items that are being released, which are approximately 25% of the items used to calculate scores.

Content Strand: The letters indicate the content strand with which the item is aligned: Word ID/Vocabulary (**WV**), Literary/Initial Understanding (**LI**), Literary/Analysis & Interpretation (**LA**), Informational/Initial Understanding (**II**), or Informational/Analysis & Interpretation (**IA**).

GSE Code: The first two digits indicate the grade of the GSE tested. The third digit indicates the GSE measured by the item.

Depth of Knowledge Code: This number indicates the Depth of Knowledge to which the item is coded.

Item Type: This indicates whether the question is multiple choice (**MC**) or constructed response (**CR**).

Correct MC Response: This is the correct letter response for multiple-choice questions.

Total Possible Points: The number indicates the maximum points awarded for the item: 1 point for a multiple-choice question and 4 points for a constructed-response question.

Student Item Results: Each student's name and state assigned student identification number are listed, followed by a score for each released item on the test included in this report.

- For multiple-choice (**MC**) questions only, a plus sign (+) indicates a correct response. If the student answered incorrectly, the letter of his or her response is indicated. An asterisk (*) indicates that the student selected more than one response.
- For all other item types, a number indicates how many points a student earned for that item.
- For all item types, a blank space indicates that the student left the question blank. A dash (–) means that the score was invalidated and that the student received no credit for parts of the test that were administered under non-standard conditions.

Total Test Results Section

Subcategory Points Earned: These columns show the points the student earned in each content strand. The content strand points earned are based on all common items in the test and not just the released items.

Total Points Earned: This column shows the total number of points the student earned on all common items.

Scaled Score: This column shows the scaled score reported as a 4-digit number. The first 2 digits are the grade and the next two digits are a score of 00-80. If the row is blank in this column, it means that the student was classified as Not Tested. (See Achievement Level below.)

Achievement Level: For Tested students, this column shows the achievement level into which the student's scores fall: **4** = Proficient with Distinction, **3** = Proficient, **2** = Partially Proficient, and **1** = Substantially Below Proficient. For Not Tested students, there are six reasons why a student did not participate: **A** = student participated in an alternate assessment in 2007-08, **L** = student is first year LEP, **W** = student withdrew from school after Oct. 1, 2008, **E** = student enrolled in school after Oct. 1, 2008, **S** = state approved special consideration, and **N** = other reason.

School/District/State Percent Correct/Average Score:

- **Released Items:** Percent correct refers to the percent of tested students who answered a multiple-choice item correctly. Average score refers to the average number of points awarded to all tested students for that constructed-response item.
- **Subcategory Points Earned:** Average score refers to the average number of points awarded to all tested students for that subcategory.

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Item Analysis Report
Mathematics

School:
District:
State:
Code:

	Released Items																	Total Test Results							
Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
Content Strand	NO	NO	NO	NO	GM	GM	GM	GM	GM	FA	NO	GM	NO	FA	DP			Numbers & Operations	Geometry & Measurement	Functions & Algebra	Data, Statistics, & Probability				
GLE Code	4-1	4-2	4-2	4-3	4-1	4-3	4-4	4-5	4-7	4-3	4-4	4-1	4-4	4-4	4-4										
Depth of Knowledge Code	2	2	2	2	2	1	1	2	2	1	2	2	2	2	2										
Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	CR										
Correct MC Response	B	C	C	D	D	B	D	B	A	C															
Name/Student ID Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	2	2	4			30	13	13	10		66		

LEGEND FOR THE ITEM ANALYSIS REPORT - MATHEMATICS

Released Items Section

Released Item Number: This number corresponds to the item number in the released item documents. This report provides complete data on items that are being released, which are approximately 25% of the items used to calculate scores.

Content Strand: The letters indicate the content strand with which the item is aligned: Numbers & Operations (**NO**), Geometry & Measurement (**GM**), Functions & Algebra (**FA**), or Data, Statistics, & Probability (**DP**).

GLE Code: The first digit indicates the grade of the GLE tested. The second digit indicates the GLE measured by the item.

Depth of Knowledge Code: This number indicates the Depth of Knowledge to which the item is coded.

Item Type: This indicates whether the question is multiple choice (**MC**), short answer (**SA**), or constructed response (**CR**).

Correct MC Response: This is the correct letter response for multiple-choice questions.

Total Possible Points: The number indicates the maximum points awarded for the item: 1 point for a multiple-choice question; 0-2 points for a short-answer question; and 0-4 points for a constructed-response question (grades 5-8 only).

Student Item Results: Each student's name and state assigned student identification number are listed, followed by a score for each released item on the test included in this report.

- For multiple-choice (**MC**) questions only, a plus sign (+) indicates a correct response. If the student answered incorrectly, the letter of his or her response is indicated. An asterisk (*) indicates that the student selected more than one response.
- For all other item types, a number indicates how many points a student earned for that item.
- For all item types, a blank space indicates that the student left the question blank. A dash (–) means that the score was invalidated and that the student received no credit for parts of the test that were administered under non-standard conditions.

Total Test Results Section

Subcategory Points Earned: These columns show the points the student earned in each content strand. The content strand points earned are based on all common items in the test and not just the released items.

Total Points Earned: This column shows the total number of points the student earned on all common items.

Scaled Score: This column shows the scaled score reported as a 3-digit number. The first digit is the grade and the next two digits are a score of 00-80. If the row is blank in this column, it means that the student was classified as Not Tested. (See Achievement Level below.)

Achievement Level: For Tested students, this column shows the achievement level into which the student's scores fall: **4** = Proficient with Distinction, **3** = Proficient, **2** = Partially Proficient, and **1** = Substantially Below Proficient. For Not Tested students, there are six reasons why a student did not participate: **A** = student participated in an alternate assessment in 2007-08, **L** = student is first year LEP, **W** = student withdrew from school after Oct. 1, 2008, **E** = student enrolled in school after Oct. 1, 2008, **S** = state approved special consideration, and **N** = other reason.

School/District/State Percent Correct/Average Score:

- **Released Items:** Percent correct refers to the percent of tested students who answered a multiple-choice item correctly. Average score refers to the average number of points awarded to all tested students for that short-answer or constructed-response item.
- **Subcategory Points Earned:** Average score refers to the average number of points awarded to all tested students for that subcategory.

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Item Analysis Report
Writing

School:
District:
State:
Code:

	Released Items																	Total Test Results							
Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
Content Strand	SC	SC	SC	SC	SC	SC	SC	SC	SC	SC	IR	LR	RW	NW	NW	NW	NW	Structures of Language & Writing Conventions	Short Responses	Extended Response					
GLE Code	4-1	4-9	4-9	4-9	4-9	4-9	4-9	4-9	4-1	4-1	4-2	4-3	4-6	4-4	4-4	4-4	4-4								
Depth of Knowledge Code	2	1	1	1	1	2	1	1	2	2	2	2	2	3	3	3	3								
Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	CR	CR	CR	SA	SA	SA	ER								
Correct MC Response	C	A	C	C	D	C	D	D	B	B															
Name/Student ID Total Possible Points	1	1	1	1	1	1	1	1	1	1	4	4	4	1	1	1	12	10	12	15			37		

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Percent Correct/Average Score: School																	
Percent Correct/Average Score: District																	
Percent Correct/Average Score: State																	

LEGEND FOR THE ITEM ANALYSIS REPORT - GRADE 5 WRITING

Released Items Section

Released Item Number: This number corresponds to the item number in the released item documents. The complete writing test, which is made up entirely of common items, is being released. This report provides complete data on those items.

Content Strand: The letters indicate the content strand with which the item is aligned: Structures of Language & Writing Conventions (**SC**), Short Responses — Response to Informational Text (**IR**), Response to Literary Text **(LR)**, Report Writing (**RW**), Extended Response — Narrative Writing (**NW**).

GLE Code: The first digit indicates the grade of the GLE tested. The second digit indicates the GLE measured by the item.

Depth of Knowledge Code: This number indicates the Depth of Knowledge to which the item is coded.

Item Type: This indicates whether the question is multiple choice (**MC**), constructed response (**CR**), short answer (**SA**), or extended response (**ER**).

Correct MC Response: This is the correct letter response for multiple-choice questions.

Total Possible Points: The number indicates the maximum points awarded for the item: 1 point for a multiple-choice question, 1 point for a short-answer question, 0-4 points for a constructed-response question, and 0-12 points for the extended response.

Student Item Results: Each student's name and state assigned student identification number are listed, followed by a score for each released item on the test included in this report.

- For multiple-choice (**MC**) questions only, a plus sign (+) indicates a correct response. If the student answered incorrectly, the letter of his or her response is indicated. An asterisk (*) indicates that the student selected more than one response.
- For all other item types, a number indicates how many points a student earned for that item.
- For all item types, a blank space indicates that the student left the question blank. A dash (–) means that the score was invalidated and that the student received no credit for parts of the test that were administered under non-standard conditions.

Total Test Results Section

Subcategory Points Earned: These columns show the points the student earned in each content strand. The content strand points earned are based on all items in the test.

Total Points Earned: This column shows the total number of points the student earned on all common items.

Scaled Score: This column shows the scaled score reported as a 3-digit number. The first digit is the grade and the next two digits are a score of 00-80. If the row is blank in this column, it means that the student was classified as Not Tested. (See Achievement Level below.)

Achievement Level: For Tested students, this column shows the achievement level into which the student's scores fall: **4** = Proficient with Distinction, **3** = Proficient, **2** = Partially Proficient, and **1** = Substantially Below Proficient. For Not Tested students, there are six reasons why a student did not participate: **A** = student participated in an alternate assessment in 2007-08, **L** = student is first year LEP, **W** = student withdrew from school after Oct. 1, 2008, **E** = student enrolled in school after Oct. 1, 2008, **S** = state approved special consideration, and **N** = other reason.

School/District/State Percent Correct/Average Score:

- **Released Items:** Percent correct refers to the percent of tested students who answered a multiple-choice item correctly. Average score refers to the average number of points awarded to all tested students for that short-answer or constructed-response item or the extended response item.
- **Subcategory Points Earned:** Average score refers to the average number of points awarded to all tested students for that subcategory.

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Item Analysis Report
Writing

School:
District:
State:
Code:

Content Strand	GSE Codes	Depth of Knowledge Code	Item Type	Total Possible Points
Informational Writing	10.6, 10.7, 10.8, 10.1, 10.9	3	Extended Response	12

Name/Student ID	Total Test Results	
	Total Points Earned	Achievement Level

LEGEND FOR THE ITEM ANALYSIS REPORT - GRADE 11 WRITING

Released Items Section

Content Strand: This indicates the genre of the extended response item: Informational Writing.

GSE Codes: The first two digits indicate the grade of the GSE tested. The third digit indicates the GSE measured by the item.

Depth of Knowledge Code: This number indicates the Depth of Knowledge to which the item is coded.

Item Type: This indicates the type of question: Extended Response.

Total Possible Points: The number indicates the maximum points awarded for the item: 0-12 points for the extended response.

Total Test Results Section

Total Points Earned: This column shows the total number of points the student earned on the common extended response item.

Achievement Level: For Tested students, this column shows the achievement level into which the student's scores fall: **4** = Proficient with Distinction, **3** = Proficient, **2** = Partially Proficient, and **1** = Substantially Below Proficient. For Not Tested students, there are six reasons why a student did not participate: **A** = student participated in an alternate assessment in 2007-08, **L** = student is first year LEP, **W** = student withdrew from school after Oct. 1, 2008, **E** = student enrolled in school after Oct. 1, 2008, **S** = state approved special consideration, and **N** = other reason.

School/District/State/Average Points:

The numbers in these rows indicate the average number of points earned on the writing test for the school, district, and state.

About The New England Common Assessment Program

This report highlights results from the Fall 2008 New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. NECAP test results are used primarily for school improvement and accountability. Achievement level results are used in the state accountability system required under No Child Left Behind (NCLB). More detailed school and district results are used by schools to help improve curriculum and instruction. Individual student results are used to support information gathered through classroom instruction and assessments.

NECAP tests in reading and mathematics are administered to students in grades 3 through 8 and 11 and writing tests are administered to students in grades 5, 8, and 11. The NECAP grade 11 tests are designed to measure student performance on grade span expectations (GSE) developed and adopted by the three states. Specifically, the tests are designed to measure the content and skills that students are expected to have as they begin the school year in their current grade – in other words, the content and skills which students have learned through the end of the previous grade.

Each test contains a mix of multiple-choice and constructed-response questions. Constructed-response questions require students to develop their own answers to questions. On the mathematics test, students may be required to provide the correct answer to a computation or word problem, draw or interpret a chart or graph, or explain how they solved a problem. On the reading test, students may be required to make a list or write a few paragraphs to answer a question related to a literary or informational passage. On the writing test, students are required to provide two extended responses of 1-3 pages.

This report contains a variety of school- and/or district-, and state-level assessment results for the NECAP tests administered at a grade level. Achievement level distributions and mean scaled scores are provided for all students tested as well as for subgroups of students classified by demographics or program participation. The report also contains comparative information on school and district performance on subtopics within each content area tested.

In addition to this report of grade 11 results, schools and districts will also receive Item Analysis Reports, Released Item support materials, and student-level data files containing NECAP results. Together, these reports and data constitute a rich source of information to support local decisions in curriculum, instruction, assessment, and professional development. Over time, this information can also strengthen school's and district's evaluation of their ongoing improvement efforts.

Fall 2008
Beginning of Grade 11
NECAP Tests

Grade 11 Students in 2008-2009

School Results

School:

District:

Code:

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Grade Level Summary Report

School:
District:
State:
Code:

Schools and districts administered all NECAP tests to every enrolled student with the following exceptions: students who participated in the alternate assessment for the 2007-08 school year, first year LEP students, students who withdrew from the school after October 1, 2008, students who enrolled in the school after October 1, 2008, students for whom a special consideration was granted through the state Department of Education, and other students for reasons not approved. On this page, and throughout this report, results are only reported for groups of students that are larger than nine (9).

PARTICIPATION in NECAP	Number									Percentage								
	School			District			State			School			District			State		
Students enrolled on or after October 1																		
	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing
Students tested																		
Students not tested in NECAP																		
State Approved																		
Alternate Assessment																		
First Year LEP																		
Withdrew After October 1																		
Enrolled After October 1																		
Special Consideration																		
Other																		

NECAP RESULTS

	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
READING																									
MATH																									
WRITING																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

Note: Throughout this report, percentages may not total 100 since each percentage is rounded to the nearest whole number.

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Reading Results

School:
District:
State:
Code:

Proficient with Distinction (Level 4)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student offers insightful observations/assertions that are well supported by references to the text. Student uses range of vocabulary strategies and breadth of vocabulary knowledge to read and comprehend a wide variety of texts.

Proficient (Level 3)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student makes and supports relevant assertions by referencing text. Student uses vocabulary strategies and breadth of vocabulary knowledge to read and comprehend text.

Partially Proficient (Level 2)
Student's performance demonstrates an inconsistent ability to read and comprehend grade-appropriate text. Student attempts to analyze and interpret literary and informational text. Student may make and/or support assertions by referencing text. Student's vocabulary knowledge and use of strategies may be limited and may impact the ability to read and comprehend text.

Substantially Below Proficient (Level 1)
Student's performance demonstrates minimal ability to derive/construct meaning from grade-appropriate text. Student may be able to recognize story elements and text features. Student's limited vocabulary knowledge and use of strategies impacts the ability to read and comprehend text.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL 2007-08													
2008-09													
Cumulative Total													
DISTRICT 2007-08													
2008-09													
Cumulative Total													
STATE 2007-08													
2008-09													
Cumulative Total													

Subtopic	Total Possible Points	Percent of Total Possible Points (0 10 20 30 40 50 60 70 80 90 100)
Word ID/Vocabulary	19	
Type of Text		
Literary	42	
Informational	43	
Level of Comprehension		
Initial Understanding	35	
Analysis & Interpretation	50	

● School
▲ District
◆ State
— Standard Error Bar

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Disaggregated Reading Results

School:
District:
State:
Code:

REPORTING CATEGORIES	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Mathematics Results

School:
District:
State:
Code:

Proficient with Distinction (Level 4)
Student's problem solving demonstrates logical reasoning with strong explanations that include both words and proper mathematical notation. Student's work exhibits a high level of accuracy, effective use of a variety of strategies, and an understanding of mathematical concepts within and across grade level expectations. Student demonstrates the ability to move from concrete to abstract representations.

Proficient (Level 3)
Student's problem solving demonstrates logical reasoning with appropriate explanations that include both words and proper mathematical notation. Student uses a variety of strategies that are often systematic. Computational errors do not interfere with communicating understanding. Student demonstrates conceptual understanding of most aspects of the grade level expectations.

Partially Proficient (Level 2)
Student's problem solving demonstrates logical reasoning and conceptual understanding in some, but not all, aspects of the grade level expectations. Many problems are started correctly, but computational errors may get in the way of completing some aspects of the problem. Student uses some effective strategies. Student's work demonstrates that he or she is generally stronger with concrete than abstract situations.

Substantially Below Proficient (Level 1)
Student's problem solving is often incomplete, lacks logical reasoning and accuracy, and shows little conceptual understanding in most aspects of the grade level expectations. Student is able to start some problems but computational errors and lack of conceptual understanding interfere with solving problems successfully.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL 2007-08													
2008-09													
Cumulative Total													
DISTRICT 2007-08													
2008-09													
Cumulative Total													
STATE 2007-08													
2008-09													
Cumulative Total													

Subtopic	Total Possible Points	Percent of Total Possible Points (0 10 20 30 40 50 60 70 80 90 100)
Numbers and Operations	21	
Geometry and Measurement	41	
Functions and Algebra	55	
Data, Statistics, and Probability	19	

● School
▲ District
◆ State
— Standard Error Bar

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Disaggregated Mathematics Results

School:
District:
State:
Code:

REPORTING CATEGORIES	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Writing Results

School:
District:
State:
Code:

Proficient with Distinction (Level 4)
Student's writing demonstrates an ability to respond to prompt/task with clarity and insight. Focus is well developed and maintained throughout response. Response demonstrates use of strong organizational structures. A variety of elaboration strategies is evident. Sentence structures and language choices are varied and used effectively. Response demonstrates control of conventions; minor errors may occur.

Proficient (Level 3)
Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.

Partially Proficient (Level 2)
Student's writing demonstrates an attempt to respond to prompt/task. Focus may be present but not maintained. Organizational structure is inconsistent with limited use of transitions. Details may be listed and lack elaboration. Sentence structures and language use are unsophisticated and may be repetitive. Response demonstrates inconsistent control of conventions.

Substantially Below Proficient (Level 1)
Student's writing demonstrates a minimal response to prompt/task. Focus is unclear or lacking. Little or no organizational structure is evident. Details are minimal and/or random. Sentence structures and language use are minimal or absent. Frequent errors in conventions may interfere with meaning.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL 2007-08													
2008-09													
Cumulative Total													
DISTRICT 2007-08													
2008-09													
Cumulative Total													
STATE 2007-08													
2008-09													
Cumulative Total													

Strand	Total Possible Points	Percent of Total Possible Points (0 10 20 30 40 50 60 70 80 90 100)	Number of Prompts	Distribution of Score Points Across Prompts	0 %	1 %	2 %	3 %	4 %	5 %	6 %
Writing in Response to Text • Response to Informational Text • Response to Literary Text	12		2	School District State							
Informational Writing • Report • Procedure • Persuasive Essay	18		3	School District State							
Expressive Writing • Reflective Essay	6		1	School District State							

● School ▲ District ◆ State — Standard Error Bar

Fall 2008 - Beginning of Grade 11 NECAP Tests
Grade 11 Students in 2008-2009
Disaggregated Writing Results

School:
District:
State:
Code:

REPORTING CATEGORIES	School													District						State					
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 NECAP Tests
School Summary
2008-2009 Students

School:
District:
State:
Code:

Reading	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Mathematics	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Writing	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

About The New England Common Assessment Program

This report highlights results from the Fall 2008 Beginning of Grade New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in New Hampshire, Rhode Island, and Vermont as part of each state's statewide assessment program. NECAP test results are used primarily for school improvement and accountability. Achievement level results are used in the state accountability system required under No Child Left Behind (NCLB). More detailed school and district results are used by schools to help improve curriculum and instruction. Individual student results are used to support information gathered through classroom instruction and assessments.

NECAP tests in reading and mathematics are administered to students in grades 3 through 8 and writing tests are administered to students in grades 5 and 8. The NECAP tests are designed to measure student performance on grade level expectations (GLE) developed and adopted by the three states. Specifically, the tests are designed to measure the content and skills that students are expected to have as they begin the school year in their current grade – in other words, the content and skills which students have learned through the end of the previous grade.

Each test contains a mix of multiple-choice and constructed-response questions. Constructed-response questions require students to develop their own answers to questions. On the mathematics test, students may be required to provide the correct answer to a computation or word problem, draw or interpret a chart or graph, or explain how they solved a problem. On the reading test, students may be required to make a list or write a few paragraphs to answer a question related to a literary or informational passage. On the writing test, students are required to provide a single extended response of 1-3 pages and three shorter responses to questions measuring different types of writing.

This report contains a variety of school- and/or district-, and state-level assessment results for the NECAP tests administered at a grade level. Achievement level distributions and mean scaled scores are provided for all students tested as well as for subgroups of students classified by demographics or program participation. The report also contains comparative information on school and district performance on subtopics within each content area tested.

In addition to this report of grade level results, schools and districts will also receive Summary Reports, Item Analysis Reports, Released Item support materials, and student-level data files containing NECAP results. Together, these reports and data constitute a rich source of information to support local decisions in curriculum, instruction, assessment, and professional development. Over time, this information can also strengthen school's and district's evaluation of their ongoing improvement efforts.

Fall 2008
Beginning of Grade 5
NECAP Tests

Grade 5 Students in 2008-2009

District Results

District:

Code:

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Grade Level Summary Report

District:
State:
Code:

Schools and districts administered all NECAP tests to every enrolled student with the following exceptions: students who participated in the alternate assessment for the 2007-08 school year, first year LEP students, students who withdrew from the school after October 1, 2008, students who enrolled in the school after October 1, 2008, students for whom a special consideration was granted through the state Department of Education, and other students for reasons not approved. On this page, and throughout this report, results are only reported for groups of students that are larger than nine (9).

PARTICIPATION in NECAP	Number									Percentage								
	School			District			State			School			District			State		
Students enrolled on or after October 1																		
	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing	Reading	Math	Writing
Students tested																		
Students not tested in NECAP																		
State Approved																		
Alternate Assessment																		
First Year LEP																		
Withdrew After October 1																		
Enrolled After October 1																		
Special Consideration																		
Other																		

NECAP RESULTS

	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
READING																									
MATH																									
WRITING																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

Note: Throughout this report, percentages may not total 100 since each percentage is rounded to the nearest whole number.

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Reading Results

District:
State:
Code:

Proficient with Distinction (Level 4)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student offers insightful observations/assertions that are well supported by references to the text. Student uses range of vocabulary strategies and breadth of vocabulary knowledge to read and comprehend a wide variety of texts.

Proficient (Level 3)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student makes and supports relevant assertions by referencing text. Student uses vocabulary strategies and breadth of vocabulary knowledge to read and comprehend text.

Partially Proficient (Level 2)
Student's performance demonstrates an inconsistent ability to read and comprehend grade-appropriate text. Student attempts to analyze and interpret literary and informational text. Student may make and/or support assertions by referencing text. Student's vocabulary knowledge and use of strategies may be limited and may impact the ability to read and comprehend text.

Substantially Below Proficient (Level 1)
Student's performance demonstrates minimal ability to derive/construct meaning from grade-appropriate text. Student may be able to recognize story elements and text features. Student's limited vocabulary knowledge and use of strategies impacts the ability to read and comprehend text.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
DISTRICT													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
STATE													
2006-07													
2007-08													
2008-2009													
Cumulative Total													

Subtopic	Total Possible Points	Percent of Total Possible Points (0 10 20 30 40 50 60 70 80 90 100)
Word ID/Vocabulary	25	
Type of Text		
Literary	56	
Informational	49	
Level of Comprehension		
Initial Understanding	42	
Analysis & Interpretation	63	

● School
▲ District
◆ State
— Standard Error Bar

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Disaggregated Reading Results

District:
State:
Code:

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Mathematics Results

District:
State:
Code:

Proficient with Distinction (Level 4)
Student's problem solving demonstrates logical reasoning with strong explanations that include both words and proper mathematical notation. Student's work exhibits a high level of accuracy, effective use of a variety of strategies, and an understanding of mathematical concepts within and across grade level expectations. Student demonstrates the ability to move from concrete to abstract representations.

Proficient (Level 3)
Student's problem solving demonstrates logical reasoning with appropriate explanations that include both words and proper mathematical notation. Student uses a variety of strategies that are often systematic. Computational errors do not interfere with communicating understanding. Student demonstrates conceptual understanding of most aspects of the grade level expectations.

Partially Proficient (Level 2)
Student's problem solving demonstrates logical reasoning and conceptual understanding in some, but not all, aspects of the grade level expectations. Many problems are started correctly, but computational errors may get in the way of completing some aspects of the problem. Student uses some effective strategies. Student's work demonstrates that he or she is generally stronger with concrete than abstract situations.

Substantially Below Proficient (Level 1)
Student's problem solving is often incomplete, lacks logical reasoning and accuracy, and shows little conceptual understanding in most aspects of the grade level expectations. Student is able to start some problems but computational errors and lack of conceptual understanding interfere with solving problems successfully.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
DISTRICT													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
STATE													
2006-07													
2007-08													
2008-2009													
Cumulative Total													

Subtopic	Total Possible Points	Percent of Total Possible Points (0 10 20 30 40 50 60 70 80 90 100)
Numbers & Operations	69	
Geometry & Measurement	32	
Functions & Algebra	32	
Data, Statistics, & Probability	25	

● School
▲ District
◆ State
— Standard Error Bar

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Disaggregated Mathematics Results

District:
State:
Code:

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Writing Results

District:
State:
Code:

Proficient with Distinction (Level 4)
Student's writing demonstrates an ability to respond to prompt/task with clarity and insight. Focus is well developed and maintained throughout response. Response demonstrates use of strong organizational structures. A variety of elaboration strategies is evident. Sentence structures and language choices are varied and used effectively. Response demonstrates control of conventions; minor errors may occur.

Proficient (Level 3)
Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.

Partially Proficient (Level 2)
Student's writing demonstrates an attempt to respond to prompt/task. Focus may be present but not maintained. Organizational structure is inconsistent with limited use of transitions. Details may be listed and lack elaboration. Sentence structures and language use are unsophisticated and may be repetitive. Response demonstrates inconsistent control of conventions.

Substantially Below Proficient (Level 1)
Student's writing demonstrates a minimal response to prompt/task. Focus is unclear or lacking. Little or no organizational structure is evident. Details are minimal and/or random. Sentence structures and language use are minimal or absent. Frequent errors in conventions may interfere with meaning.

	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	
SCHOOL													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
DISTRICT													
2006-07													
2007-08													
2008-2009													
Cumulative Total													
STATE													
2006-07													
2007-08													
2008-2009													
Cumulative Total													

Subtopic	Total Possible Points
Structures of Language & Writing Conventions	10
Short Responses	12
Extended Response	15

Percent of Total Possible Points: 0 10 20 30 40 50 60 70 80 90 100

● School
▲ District
◆ State
— Standard Error Bar

Fall 2008 - Beginning of Grade 5 NECAP Tests
Grade 5 Students in 2008-2009
Disaggregated Writing Results

District:
State:
Code:

REPORTING CATEGORIES	District													State											
	Enrolled	NT Approved	NT Other	Tested	Level 4		Level 3		Level 2		Level 1		Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score	Tested	Level 4	Level 3	Level 2	Level 1	Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%		N	%	%	%	%		N	%	%	%	%	
All Students																									
Gender																									
Male																									
Female																									
Not Reported																									
Primary Race/Ethnicity																									
American Indian or Alaskan Native																									
Asian																									
Black or African American																									
Hispanic or Latino																									
Native Hawaiian or Pacific Islander																									
White (non-Hispanic)																									
No Primary Race/Ethnicity Reported																									
LEP Status																									
Currently receiving LEP services																									
Former LEP student - monitoring year 1																									
Former LEP student - monitoring year 2																									
All Other Students																									
IEP																									
Students with an IEP																									
All Other Students																									
SES																									
Economically Disadvantaged Students																									
All Other Students																									
Migrant																									
Migrant Students																									
All Other Students																									
Title I																									
Students Receiving Title I Services																									
All Other Students																									

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

NOTE: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2008 NECAP Tests

District Summary

2008-2009 Students

District:
State:
Code:

Reading	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Mathematics	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Writing	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
					Level 4		Level 3		Level 2		Level 1		Mean Scaled Score
	N	N	N	N	N	%	N	%	N	%	N	%	

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

Appendix R—Decision Rules

ANALYSIS AND REPORTING DECISION RULES

NECAP
Fall 07-08 Grades 03-08 Administration

This document details rules for analysis and reporting. The final student level data set used for analysis and reporting is described in the "Data Processing Specifications." This document is considered a draft until the NECAP State Department of Education (DOE) signs off. If there are rules that need to be added or modified after said sign-off, DOE sign off will be obtained for each rule. Details of these additions and modifications will be in the Addendum section.

I. General Information

A. *Tests administered:*

Grade	Subject	Test items used for Scaling	IREF Reporting Categories (Subtopic and Subcategory IREF Source)
03	Reading	Common	Cat2
03	Math	Common	Cat1
04	Reading	Common	Cat2
04	Math	Common	Cat1
05	Reading	Common	Cat2
05	Math	Common	Cat1
05	Writing	Common	type
06	Reading	Common	Cat2
06	Math	Common	Cat1
07	Reading	Common	Cat2
07	Math	Common	Cat1
08	Reading	Common	Cat2
08	Math	Common	Cat1
08	Writing	Common	type

B. *Reports Produced*:

1. Student Report
   a. Testing School District
2. School Item Analysis Report by Grade and Subject
   a. Testing School District
   b. Teaching School District
3. Grade Level School/District/State Results
   a. Testing School District
   b. Teaching School District – District and School Levels only
4. School/District/State Summary
   a. Testing School District
   b. Teaching School District – District and School Levels only

C. *Files Produced:*

1. State Student Cleanup Data
2. Preliminary State Results
3. State Student Released Item Data
4. State Student Raw Data
5. State Student Scored Data
6. District Student Data
7. Item Information
8. Grade Level Results Report Disaggregated and Historical Data
9. Grade Level Results Report Participation Category Data
10. Grade Level Results Report Subtopic Data
11. Summary Results Data
12. Released Item Percent Responses Data
13. Invalidated Students Original Score
14. Multiple Choice Response Distribution Data Grades 05-08
15. Block Blank Response Distribution Data Grades 03 & 04

D. *School Type:*

SchType	**Source: ICORE SubTypeID**	**Description**
PUB	1,12,13	Public School
PRI	3	Private School
OOD	4	Out-of-District Private Providers
OUT	8	Out Placement
CHA	11	Charter School
INS	7	Institution
OTH	9	Other

School Type Impact on Data Analysis and Reporting				
Level	**Testing**		**Teaching**	
	Impact on Analysis	**Impact on Reporting**	**Impact on Analysis**	**Impact on Reporting**
Student	n/a	Report students based on testing discode and schcode. District data will be blank for students tested at PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	n/a	n/a
School	Include all non-home school students using testing school code for aggregations	Generate a report for each school with at least one student enrolled using the tested school aggregate denominator. District data will be blank for PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	Include all non-home school students using the teaching school code. Exclude students who do not have a teaching school code.	Generate a report for each school with at least one student enrolled using the teaching school aggregate denominator. District data will be blank for PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.
District	For OUT and OOD schools, aggregate using the sending district. If OUT or OOD student does not have a sending district, do not include in aggregations. Do not include students tested at PRI, INS, or OTH schools Do not include home school students.	Generate a report for each district with at least one student enrolled using the tested district aggregate denominator. Always report tested year state data.	Do not include students taught at PRI, OOD, OUT, INS, or OTH schools. Do not include students who do not have a teaching district code. Do not include home school students.	Generate a report for each district with at least one student enrolled using the teaching district aggregate denominator. Always report tested year state data.
State	Do not include students tested at PRI schools for NH and RI. Include all students for VT. Do not include home school students.	Always report testing year state data.	n/a	n/a

E. *Requirements To Report Aggregate Data(Minimum N)*

Calculation Description	Rule
Number and Percent at each achievement level, mean score by disaggregated category and aggregate level	If the number of tested students included in the denominator is less than 10, then do not report.
Content Area Subcategories Average Points Earned based on common items only by aggregate level	If the number of tested students included in the denominator is less than 10, then do not report.
Aggregate data on Item Analysis report	No required minimum number of students
Number and Percent of students in a participation category by aggregate level	No required minimum number of students
Content Area Subtopic Percent of Total Possible Points and Standard Error Bar	If any item was not administered to at least one tested student included in the denominator or the number of tested students included in the denominator is less than 10, then do not report

F. *Special Forms:*

1. Form 00 is created for students whose matrix scores will be ignored for analysis. Such students include Braille or administration issues resolved by program management.

G. *Other Information*

1. Home school students are excluded from all school, district, and state level aggregations. Home school students receive a parent letter based on the testing school. Print aggregate data based on the testing school. Print tested year state data. Home school students are not listed on the item analysis report.
2. Plan504 data not available for NH and VT; therefore 504 Plan section will be suppressed for NH and VT.
3. To calculate Title1 data for writing using Title1rea variable.
4. Title 1 data are not available for VT; therefore Title 1 section will be suppressed for VT.
5. Only students with a testing year school type of OUT or OOD are allowed to have a sending district code. Non-public sending district codes will be ignored. For RI, senddiscode of 88 is ignored. For NH, senddiscode of 000 is ignored.
6. Several reports and data files are provided by testing and teaching school district levels. Testing level is defined to be the school and district where the student tested (discode and schcode). Teaching level is defined to be where the student was enrolled last year (sprdiscode and sprschcode). Every student will have testing district and school codes. Some students will have a teaching school code. Some students will have a teaching district code.

II. Student Participation / Exclusions

A. *Test Attempt Rules by content area*

1. A content area was attempted if any multiple choice item or non-field test open response item has been answered. (Use original item responses – see special circumstances section II.F)
2. A multiple choice item has been answered by a student if the response is A, B, C, D, or * (*=multiple responses)
3. An open response item has been answered if it is not scored blank ‘B’

B. *Session Attempt Rules by content area*

1. A session was attempted if any multiple choice item or non-field test open response item has been answered in the session. (Use original item responses – see special circumstances section II.F)

C. *Not Tested Reasons by content area*

1. Not Tested State Approved Alternate Assessment

   a. If content area "Alternate Assessment blank or partially blank reason" is marked, then student is identified as "Not Tested State Approved Alternate Assessment".

2. Not Tested State Approved First Year LEP (reading and writing only)

   a. If content area "First Year LEP blank or partially blank reason" is marked, then student is identified as "Not Tested State Approved First Year LEP".

3. Not Tested State Approved Special Consideration

   a If content area "Special Consideration blank or partially blank reason" is marked, the student is identified as "Not Tested State Approved Special Consideration".

4. Not Tested State Approved Withdrew After October 1

   a. If content area "Withdrew After October 1 blank or partially blank reason" is marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Withdrew After October 1"

5. Not Tested State Approved Enrolled After October 1

   a. If content area "Enrolled After October 1 blank or partially blank reason" is marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Enrolled After October 1".

6. Not Tested Other

   a. If content area test was not attempted, the student is identified as "Not Tested Other".

D. *Not Tested Reasons Hierarchy by content area: if more than one reason for not testing at a content area is identified then select the first category indicated in the order of the list below.*

1. Not Tested State Approved Alternate Assessment
2. Not Tested State Approved First Year LEP (reading and writing only)
3. Not Tested State Approved Special Consideration
4. Not Tested State Approved Withdrew After October 1
5. Not Tested State Approved Enrolled After October 1
6. Not Tested Other

E. *Student Participation Status by content area*

1. Tested

   a. If the student does not have any content area not tested reasons identified, then the student is considered Tested for the content area.

2. Not Tested: State Approved Alternate Assessment
3. Not Tested: State Approved First Year LEP (reading and writing only)
4. Not Tested: State Approved Special Consideration
5. Not Tested: State Approved Withdrew After October 1
6. Not Tested: State Approved Enrolled After October 1
7. Not Tested: Other

F. *Special Circumstances by content area*

1. Students identified as content area tested and did not attempt all sessions in the test are considered to be "Tested Incomplete."
2. Students identified as content area tested and have at least one of the content area invalidation session flags marked will be treated as "Tested with Non-Standard Accommodations". Math accommodation F01 also identifies non-standard accommodations for Math.
3. For students identified as "Tested with Non-Standard Accommodations" the content area sessions item responses which are marked for invalidation will be treated as a non-response. For the students with math accommodations F01 marked, the non-calculator session 1 math items will be treated as a non-response.
4. Students identified as tested in a content area will receive released item scores, scaled score, scale score bounds, achievement level, raw total score, subcategory scores, and writing annotations (where applicable).
5. Students identified as not tested in a content area will not receive a scaled score, scaled score bounds, achievement level, writing annotations (where applicable). They will receive released item scores, raw total score, and subcategory scores.

G. *Student Participation Summary*

Participation Status	Description	Raw Score(*)	Scaled Score	Ach. Level	Student Report Ach. Level Text	Roster Ach. Level Text
1	Tested	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
2	Not Tested State Approved Alternate Assessment	✓			Alternate Assessment	A
3	Not Tested State Approved First Year LEP	✓			First Year LEP	L
4	Not Tested State Approved Enrolled After October 1	✓			Enrolled After October 1	E
5	Not Tested State Approved Withdrew After October 1	✓			Withdrew After October 1	W
6	Not Tested State Approved Special Consideration	✓			Special Consideration	S
7	Not Tested Other	✓			Not Tested	N

(*) Raw scores are not printed on student report for students with a not tested status.

III. Calculations

A. *Rounding*

1. All percents are rounded to the nearest whole number
2. All mean scaled scores are rounded to the nearest whole number
3. Content Area Subcategories: Average Points Earned (student report): round to the nearest tenth.
4. Round non-multiple choice average item scores to the nearest tenth.

B. *Students included in calculations based on participation status*

1. For number and percent of students enrolled, tested, and not tested categories include all students not excluded by other decision rules.
2. For number and percent at each achievement level, average scaled score, subtopic percent of total possible points and standard error, subcategories average points earned, percent/correct average score for each released item include all tested students not excluded by other decision rules.

C. *Raw scores*

1. For all analysis, non-response for an item by a tested student is treated as a score of 0.
2. Content Area Total Points: Sum the pointes earned by the student for the common items.

D. *Item Scores*

1. For all analysis, non-response for an item by a tested student is treated as a score of 0.
2. For multiple choice released item data store a '+' for correct response, or A,B,C,D,* or blank
3. For open response released items, store the student score. If the score is not numeric ('B'), then store it as blank.
4. For students identified as content area tested with non-standard accommodations, then store the released item score as '-' for invalidated items.

E. *Scaling*

Scaling is done using a look-up table provided by psychometrics and the student's raw score.

F. *SubTopic Item Scores*

1. Identify the Subtopic
   a. The excel file IREF_ReportingCategories.xls outlines the IREF variables and values for identifying the Content Strand, GLE code, Depth of Knowledge code, subtopics, and subcategories. The variable type in IREF is the source for the Item Type, except the writing prompt item type is reported as "ER".
2. Student Content Area Subcategories (student report): Subtopic item scores at the student level is the sum of the points earned by the student for the common items in the subtopic.
3. Content Area Subtopic (grade level results report): Subtopic scores are based on all unique common and matrix items. The itemnumber identifies each unique item.
   a. Percent of Total Possible Points:
      I. For each unique common and matrix item calculate the average student score as follows: (sum student item score/number of tested students administered the item).
      II. 100 * (Sum the average score for items in the subtopic)/(Total Possible Points for the subtopic) rounded to the nearest whole number.
   b. Standard Error Bar: Before multiplying by 100 and rounding the Percent of Total Possible points (ppe) calculate standard error for school,district and state: 100* (square root (((ppe)*(1-ppe)/number of tested students))) rounded to the nearest whole number
      Percent of Total Possible Points +/- Standard Error

G. *Cumulative Total*

1. Include the yearly results where the number tested is greater than or equal to 10
2. Cumulative total N (Enrolled, Not Tested Approved, Not Tested Other, Tested, at each achievement level) is the sum of the yearly results for each category where the number tested is greater than or equal to 10.

3. Cumulative percent for each achievement level is 100*(Number of students at the achievement level cumulative total / number of students tested cumulative total) rounded to the nearest whole number.

4. Cumulative mean scaled score is a weighted average. For years where the number tested is greater than or equal to 10, (sum of (yearly number tested * yearly mean scaled score)) / (sum of yearly number tested) rounded to the nearest whole number.

H. *Average Points Earned Students at Proficient Level (Range)*

1. Select all students across the states with Y40 scaled score, where Y=grade. Average the content area subcategories across the students and round to the nearest tenth. Add and subtract one standard error of measurement to get the range.

I. Writing Annotations

1. Students with a writing prompt score of 2-12 receive at least one, but up to five statements based on decision rules for annotations as outlined in Final Statements & Decision Rules for NECAP Writing Annotations.doc

IV. Report Specific Rules

A. *Student Report*

1. Student header Information

a. If "FNAME" or "LNAME" is not missing then print "FNAME MI LNAME". Otherwise, print "No Name Provided".

b. Print the student's tested grade

c. For school and district name, print the abbreviated tested school and district ICORE name based on school type decision rules.

d. Print "NH","RI", or "VT" for state.

2. Test Results by content area

a. For students identified as "Not Tested", print the not tested reason in the achievement level, leave scaled score and graphic display blank.

b. For students identified as tested for the content area then do the following

I. Print the complete achievement level name the student earned

II. Print the scaled score the student earned

III. Print a vertical black bar for the student scaled score with gray horizontal bounds in the graphic display

IV. For students identified as "Tested with a non-standard accommodation" for a content area, print '**' after the content area earned achievement level and after student points earned for each subcategory.

V. For students identified as "Tested Incomplete" for a content area, place a section symbol after content area earned scaled score.

3. Exclude students based on school type and participation status decision rules for aggregations.

4. This Student's Achievement Compared to Other Students by content area

a. For tested students, print a check mark in the appropriate achievement level in the content area student column. For not tested students leave blank

b. For percent of students with achievement level by school, district and state print aggregate data based on school type and minimum N rules

5. This Student's Performance in Content Area Subcategories by content area

a. Always print total possible points and students at proficient average points earned range.

b. For students identified as not tested then leave student scores blank

c. For students identified as tested do the following

I. Always print student subcategory scores

II. If the student is identified as tested with a non-standard accommodation for the content area then place '**" after the student points earned for each subcategory.

d. Print aggregate data based on school type and minimum N-size rules.

5. Writing Annotations (Grades 05 and 08 only)

a. For students with writing prompt score of 2-12 print at least one, but up to five annotation statements.

B. *School Item Analysis Report by Grade and Subject*

1. Reports are created for testing school and teaching school independently.

2. School Header Information

a. Use abbreviated ICORE school and district name based on school type decision rules

b. Print "New Hampshire", "Rhode Island", or "Vermont" for State.

c. For NH, the code should print SAU code – district code – school code. For RI and VT, the code should print district code – school code.

3. For multiple choice items, print '+' for correct response, or A,B,C,D,* or blank

4. For open response items, print the student score. If the score is not numeric ('B'), then leave blank.

5. For students identified as content area tested with non-standard accommodations, print '-' for invalidated items.

6. All students receive subcategory points earned and total points earned.

7. Leave scaled score blank for not tested students and print the not tested reason in the achievement level column.

8. Exclude students based on school type and participation status decision rules for aggregations.

9. Always print aggregated data regardless of N-size based on school type decision rules.

10. For students identified as not tested for the content area print a cross symbol next to students' name.

11. For students identified as tested incomplete for the content area print a section symbol next to the scaled score.

12. Home school student are not listed on the report.

C. Grade Level School/District/State Results

1. Reports are run by testing state, testing district, testing school, teaching district, and teaching school.

2. Exclude students based on school type and participation status decision rules for aggregations.

3. Report Header Information

a. Use abbreviated school and district name from ICORE based on school type decision rules.

b. Print "New Hampshire", "Rhode Island", or "Vermont" to reference the state. The state graphic is printed on the first page.

4. Report Section: Participation in NECAP

a. For testing level reports always print number and percent based on school type decision rules.

b. For the teaching level reports leave the section blank.

5. Report Section: NECAP Results by content area

a. For the testing level report always print based on minimum N-size and school type decision rules.

b. For the teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

6. Report Section: Historical NECAP Results by content area

a. For teaching level report always print current year, prior years, and cumulative total results based on minimum N-size and school type decision rules.

b. For teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

7. Report Section: Subtopic Results by content area

a. For testing and teaching level reports always print based on minimum N-size and school type decision rules

8. Report Section: Disaggregated Results by content area

a. For testing level report always print based on minimum N-size and school type decision rules.

b. For teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

D. *School/District/State Summary*

1. Reports are run by testing state, testing district, testing school, teaching district, and teaching school

2. Exclude students based on school type and participation status decision rules for aggregations.

3. For testing level report print entire aggregate group across grades tested and list grades tested results based on minimum N-size and school type decision rules. Mean scaled score across the grades is not calculated.

4. For the teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules. Mean scaled score across the grades is not calculated.

V. Data File Rules

In the file names GG refers to the two digit grade (03-08) , YYYY refers to the year 0708, DDDDD refers to the district code, and SS refers to two letter state code.

A. State Student Cleanup Data

1. One CSV file per grade and state will be created based on the file layout NECAPYYYYF Gr 03-08 11 Student Demographic Cleanup File Layout.xls.

2. Refer to NECAPYYYYF Gr 03-08 11 Student Demographic Cleanup Description.doc

3. Session Invalidation Flags are marked as follows.

a. If reaaccF02 or reaaccF03 is marked, then mark reaInvSes1, reaInvSes2, and reaInvSes3

b. If mataccF03 is marked, then mark matInvSes1, matInvSes2, and matInvSes03. MataccF01 is left as marked on booklet.

c. If wriaccF03 is marked, then mark wriInvSes1 and wriInvSes2

B. Preliminary State Results

1. A PDF file will be created for each state containing preliminary state results for each grade and subject and will list historical state data for comparison.
2. The file name will be SSPreliminaryResultsDATE.pdf

C. State Student Released Item Data

1. Students who tested at a private school are excluded from NH and RI student data files.
2. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 03-08 State Student Data Released Item Layout.xls
3. The CSV file name will be NECAP YYYY Fall State Student Data Released Item Gr GG.csv.

D. State Student Raw Data

1. Students who tested at a private school are excluded from NH and RI student data files.
2. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 03-08 State Student Raw Data File Layout.xls
3. The CSV file name will be NECAP YYYY Fall State Student Raw Data File Gr GG.csv.

E. State Student Scored Data

1. Students who tested at a private school are excluded from NH and RI student data files.
2. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 03-08 State Student Scored Data File Layout.xls
3. The CSV file name will be NECAP YYYY Fall State Student Scored Data File Gr GG.csv.

F. District Student Data

1. Students with the Discode or SendDiscode will be in the district grade specific CSV file for the testing year.
2. Students with a sprDiscode will be in the district grade specific CSV file for the teaching year.
3. Home school students are excluded from district student data files. For NH and RI only public school districts will receive district data files. (Districts with at least one school with schoolsubtypeID=1 in ICORE)
4. Testing and teaching CSV files will be created for each state and grade and district following the layout NECAP YYYY Fall Gr 03-08 District Student Data Layout.xls
5. The testing CSV file name will be NECAP YYYY Fall Testing District Slice Gr GG_DDDDD.csv. The teaching CSV file name will be NECAP YYYY Fall Teaching District Slice Gr GG_DDDDD.csv.

G. Item Information

1. An excel file will be created containing item information for common items: grade, subject, raw data item name, item type, key, and point value.
2. The file name will be NECAP YYYY Fall Gr 03-08 Item Information.xls

H. Grade Level Results Report Disaggregated and Historical Data

1. Teaching and testing CSV files will be created for each state and grade containing the grade level results disaggregated and historical data following the layout NECAP YYYY Fall Gr 03-08 Results Report Disaggregated and Historical Data Layout.xls.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. The testing file name will be NECAP YYYY Fall Testing Results Report Disaggregated and Historical Data Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Results Report Disaggregated and Historical Data Gr GG.csv.

I. Grade Level Results Report Participation Category Data

1. Teaching CSV file will be created for each state and grade containing the grade level results participation data following the layout NECAP YYYY Fall Gr 03-08 Results Report Participation Category Data Layout.xls.
2. The testing file name will be NECAP YYYY Fall Testing Results Report Participation Category Data Gr GG.csv

J. Grade Level Results Report Subtopic Data

1. Teaching and testing CSV files will be created for each state and grade containing the grade level results subtopic data following the layout NECAP YYYY Fall Gr 03-08 Results Report Subtopic Data Layout.xls.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. The testing file name will be NECAP YYYY Fall Testing Results Report Subtopic Data Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Results Report Subtopic Data Gr GG.csv.

K. Summary Results Data

1. Teaching and testing CSV files will be created for each state and grade containing the summary report data following the layout NECAP YYYY Fall Gr 03-08 Summary Results Layout.xls.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. The testing file name will be NECAP YYYY Fall Testing Summary Results Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Summary Results Gr GG.csv.

L. Released Item Percent Responses Data

1. The CSV files will only contain state level aggregation for released items.
2. Teaching and testing CSV files will be created for each state and grade containing the released item analysis report state data following the layout NECAP YYYY Fall Gr 03-08 Released Item Percent Responses Layout.xls.
3. The testing file name will be NECAP YYYY Fall Testing Released Item Percent Responses.csv . The teaching file name will be NECAP YYYY Fall Teaching Released Item Percent Responses.csv.

M. Invalidated Students Original Score

1. Original raw scores for students whose responses were invalidated for reporting will be provided.
2. Students who tested at a private school are excluded from NH and RI student data files.
3. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 03-08 State Invalidated Student Original Scored Data File Layout.xls.
4. The CSV file name will be NECAP YYYY Fall State Student Scored Data File Gr GG OriScInvStu.csv.

N. Multiple Choice Response Distribution Data Grades 05-08

1. One CSV file will be created containing the frequency of multiple responses (*) for multiple choice items.
2. All students are included in the frequencies.
3. The file will follow the layout NECAP YYYY Fall Multiple MC Responses Freq Layout.xls and will be named NECAP YYYY Fall Multiple MC Responses Freq.xls.

O. Block Blank Response Distribution Data Grades 03 & 04

Addenda

1. 01/04/2008: Grade Level School/District/State Results – Cumulative Total

   - Suppress cumulative total data if at least one reported year has fewer than 10 tested students.

Analysis and Reporting Decision Rules
NECAP
Fall 07-08 Grade 11 Administration

This document details rules for analysis and reporting. The final student level data set used for analysis and reporting is described in the "Data Processing Specifications." This document is considered a draft until the NECAP State Department of Education (DOE) signs off. If there are rules that need to be added or modified after said sign-off, DOE sign off will be obtained for each rule. Details of these additions and modifications will be in the Addendum section.

VI. General Information

A. *Tests administered:*

Grade	Subject	Test items used for Scaling	IREF Reporting Categories (Subtopic and Subcategory IREF Source)
11	Reading	Common	Cat2
11	Math	Common	Cat1
11	Writing	Common	form

B. *Reports Produced*:

1. Student Report
   b. Testing School District
2. School Item Analysis Report by Grade and Subject
   c. Testing School District
   d. Teaching School District
3. Grade Level School/District/State Results
   c. Testing School District
   d. Teaching School District – District and School Levels only

C. *Files Produced:*

1. State Student Cleanup Data
2. Preliminary State Results
3. State Student Released Item Data
4. State Student Raw Data
5. State Student Scored Data
6. District Student Data
7. Item Information
8. Grade Level Results Report Disaggregated and Historical Data
9. Grade Level Results Report Participation Category Data
10. Grade Level Results Report Subtopic Data
11. Released Item Percent Responses Data
12. Invalidated Students Original Score
13. Multiple Choice Response Distribution Data Grades 11

D. *School Type:*

SchType	Source: ICORE SubTypeID	Description
PUB	1,12,13	Public School
PRI	3	Private School
OOD	4	Out-of-District Private Providers
OUT	8	Out Placement
CHA	11	Charter School
INS	7	Institution
OTH	9	Other

School Type Impact on Data Analysis and Reporting				
Level	**Testing**		**Teaching**	
	Impact on Analysis	**Impact on Reporting**	**Impact on Analysis**	**Impact on Reporting**
Student	n/a	Report students based on testing discode and schcode. District data will be blank for students tested at PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	n/a	n/a
School	Include all non-home school students using testing school code for aggregations	Generate a report for each school with at least one student enrolled using the tested school aggregate denominator. District data will be blank for PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	Include all non-home school students using the teaching school code. Exclude students who do not have a teaching school code.	Generate a report for each school with at least one student enrolled using the teaching school aggregate denominator. District data will be blank for PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.
District	For OUT and OOD schools, aggregate using the sending district. If OUT or OOD student does not have a sending district, do not include in aggregations. Do not include students tested at PRI, INS, or OTH schools Do not include home school students.	Generate a report for each district with at least one student enrolled using the tested district aggregate denominator. Always report tested year state data.	Do not include students taught at PRI, OOD, OUT, INS, or OTH schools. Do not include students who do not have a teaching district code. Do not include home school students.	Generate a report for each district with at least one student enrolled using the teaching district aggregate denominator. Always report tested year state data.
State	Do not include students tested at PRI schools for NH and RI. Include all students for VT. Do not include home school students.	Always report testing year state data.	n/a	n/a

E. *Requirements To Report Aggregate Data(Minimum N)*

Calculation Description	Rule
Number and Percent at each achievement level, mean score by disaggregated category and aggregate level	If the number of tested students included in the denominator is less than 10, then do not report.
Content Area Subcategories Average Points Earned based on common items only by aggregate level	If the number of tested students included in the denominator is less than 10, then do not report.
Aggregate data on Item Analysis report	No required minimum number of students
Number and Percent of students in a participation category by aggregate level	No required minimum number of students
Content Area Subtopic Percent of Total Possible Points and Standard Error Bar and Grade 11 Writing Distribution of Score Points Across Prompts	If any item was not administered to at least one tested student included in the denominator or the number of tested students included in the denominator is less than 10, then do not report

F. *Special Forms:*

1. Form 00 is created for students whose matrix scores will be ignored for analysis. Such students include Braille or administration issues resolved by program management.

G. *Other Information*

1. Home school students are excluded from all school, district, and state level aggregations. Home school students receive a parent letter based on the testing school. Print aggregate data based on the testing school. Print tested year state data. Home school students are not listed on the item analysis report.
2. Plan504 data not available for NH and VT; therefore 504 Plan section will be suppressed for NH and VT.
3. To calculate Title1 data for writing using Title1rea variable.
4. Title 1 data are not available for VT; therefore Title 1 section will be suppressed for VT.
5. Only students with a testing year school type of OUT or OOD are allowed to have a sending district code. Non-public sending district codes will be ignored. For RI, senddiscode of 88 is ignored. For NH, senddiscode of 000 is ignored.
6. Several reports and data files are provided by testing and teaching school district levels. Testing level is defined to be the school and district where the student tested (discode and schcode). Teaching level is defined to be where the student was enrolled last year (sprdiscode and sprschcode). Every student will have testing district and school codes. Some students will have a teaching school code. Some students will have a teaching district code.

VII. Student Participation / Exclusions

A. *Test Attempt Rules by content area*

1. Grade 11 writing was attempted if the common writing prompt is not scored blank ‘B’. For all other grades and content areas test attempt can be determined as follows. A content area was attempted if any multiple choice item or non-field test open response item has been answered. (Use original item responses – see special circumstances section II.F)
2. A multiple choice item has been answered by a student if the response is A, B, C, D, or * (*=multiple responses)
3. An open response item has been answered if it is not scored blank ‘B’

B. *Session Attempt Rules by content area*

1. A session was attempted if any multiple choice item or non-field test open response item has been answered in the session. (Use original item responses – see special circumstances section II.F)

2. Because of the test design for grade 11 writing, only determine if session 1 was attempted. Session 2 is ignored.

C. *Not Tested Reasons by content area*

1. Not Tested State Approved Alternate Assessment

   b. If content area "Alternate Assessment blank or partially blank reason" is marked, then student is identified as "Not Tested State Approved Alternate Assessment".

2. Not Tested State Approved First Year LEP (reading and writing only)

   a. If content area "First Year LEP blank or partially blank reason" is marked, then student is identified as "Not Tested State Approved First Year LEP".

3. Not Tested State Approved Special Consideration

   a. If content area "Special Consideration blank or partially blank reason" is marked, the student is identified as "Not Tested State Approved Special Consideration".

4. Not Tested State Approved Withdrew After October 1

   b. If content area "Withdrew After October 1 blank or partially blank reason" is marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Withdrew After October 1". For grade 11 writing, only use session 1 attempt status.

5. Not Tested State Approved Enrolled After October 1

   - If content area "Enrolled After October 1 blank or partially blank reason" is marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Enrolled After October 1". For grade 11 writing, only use session 1 attempt status.

6. Not Tested Other

   a. If content area test was not attempted, the student is identified as "Not Tested Other".

D. *Not Tested Reasons Hierarchy by content area: if more than one reason for not testing at a content area is identified then select the first category indicated in the order of the list below.*

7. Not Tested State Approved Alternate Assessment
8. Not Tested State Approved First Year LEP (reading and writing only)
9. Not Tested State Approved Special Consideration
10. Not Tested State Approved Withdrew After October 1
11. Not Tested State Approved Enrolled After October 1
12. Not Tested Other

E. *Student Participation Status by content area*

1. Tested

   a. If the student does not have any content area not tested reasons identified, then the student is considered Tested for the content area.

8. Not Tested: State Approved Alternate Assessment
9. Not Tested: State Approved First Year LEP (reading and writing only)
10. Not Tested: State Approved Special Consideration

11. Not Tested: State Approved Withdrew After October 1

12. Not Tested: State Approved Enrolled After October 1

13. Not Tested: Other

F. *Special Circumstances by content area*

6. Students identified as content area tested and did not attempt all sessions in the test are considered to be "Tested Incomplete." Not applicable at grade 11 writing.

7. Students identified as content area tested and have at least one of the content area invalidation session flags marked will be treated as "Tested with Non-Standard Accommodations". Math accommodation F01 also identifies non-standard accommodations for Math.

8. For students identified as "Tested with Non-Standard Accommodations" the content area sessions item responses which are marked for invalidation will be treated as a non-response. For the students with math accommodations F01 marked, the non-calculator session 1 math items will be treated as a non-response.

9. Students identified as tested in a content area will receive released item scores, scaled score, scale score bounds, achievement level, raw total score, subcategory scores, and writing annotations (where applicable).

10. Students identified as not tested in a content area will not receive a scaled score, scaled score bounds, achievement level, writing annotations (where applicable). They will receive released item scores, raw total score, and subcategory scores.

G. *Student Participation Summary*

Participation Status	Description	Raw Score(*)	Scaled Score (**)	Ach. Level	Student Report Ach. Level Text	Roster Ach. Level Text
1	Tested	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
2	Not Tested State Approved Alternate Assessment	✓			Alternate Assessment	A
3	Not Tested State Approved First Year LEP	✓			First Year LEP	L
4	Not Tested State Approved Enrolled After October 1	✓			Enrolled After October 1	E
5	Not Tested State Approved Withdrew After October 1	✓			Withdrew After October 1	W
6	Not Tested State Approved Special Consideration	✓			Special Consideration	S
7	Not Tested Other	✓			Not Tested	N

(*) Raw scores are not printed on student report for students with a not tested status.

(**) Grade 11 writing students do not receive a scaled score. The writing achievement level is determined by the total common writing prompt score.

VIII. Calculations

A. *Rounding*

5. All percents are rounded to the nearest whole number
6. All mean scaled scores are rounded to the nearest whole number
7. Grade 11 writing mean (raw) score is rounded to the nearest tenth.
8. Content Area Subcategories: Average Points Earned (student report): round to the nearest tenth.
9. Round non-multiple choice average item scores to the nearest tenth.

B. *Students included in calculations based on participation status*

3. For number and percent of students enrolled, tested, and not tested categories include all students not excluded by other decision rules.
4. For number and percent at each achievement level, average scaled score, subtopic percent of total possible points and standard error, subtopic distribution across writing prompts, subcategories average points earned, percent/correct average score for each released item include all tested students not excluded by other decision rules.

C. *Raw scores*

1. For all analysis, non-response for an item by a tested student is treated as a score of 0.
2. Content Area Total Points: Sum the pointes earned by the student for the common items.

D. *Item Scores*

1. For all analysis, non-response for an item by a tested student is treated as a score of 0.
2. For multiple choice released item data store a '+' for correct response, or A,B,C,D,* or blank
3. For open response released items, store the student score. If the score is not numeric ('B'), then store it as blank.
4. For students identified as content area tested with non-standard accommodations, then store the released item score as '-' for invalidated items.
5. For common writing prompt score, the final score of record is the sum of scorer 1 and scorer 2. If both scorers give the student a B(F), then the final score is B(F).
6. For matrix writing prompt score, the final score of record is scorer 1.

E. *Scaling*

Scaling is done using a look-up table provided by psychometrics and the student's raw score.

F. *SubTopic Item Scores*

4. Identify the Subtopic
   a. The excel file IREF_ReportingCategories.xls outlines the IREF variables and values for identifying the Content Strand, GLE code, Depth of Knowledge code, subtopics, and subcategories. The variable type in IREF is the source for the Item Type, except the writing prompt item type is reported as "ER".
5. Student Content Area Subcategories (student report): Subtopic item scores at the student level is the sum of the points earned by the student for the common items in the subtopic. For grade 11 writing, the subtopic score is the final score of record for the common writing prompt.
6. Content Area Subtopic (grade level results report): Subtopic scores are based on all unique common and matrix items. The itemnumber identifies each unique item.
   a. Percent of Total Possible Points:

I. For each unique common and matrix item calculate the average student score as follows: (sum student item score/number of tested students administered the item).

II. 100 * (Sum the average score for items in the subtopic)/(Total Possible Points for the subtopic) rounded to the nearest whole number.

b. Standard Error Bar: Before multiplying by 100 and rounding the Percent of Total Possible points (ppe) calculate standard error for school,district and state: 100* (square root (((ppe)*(1-ppe)/number of tested students))) rounded to the nearest whole number
Percent of Total Possible Points +/- Standard Error

G. *Grade 11 Writing: Distribution of Score Points Across Prompts*.

1. Each prompt is assigned a subtopic based on information provided by program management.

2. The set of items used to calculate the percent at each score point is defined as follows: scorer 1 common prompt score, scorer 2 common prompt score, scorer 1of each matrix prompt. (Note: scores of 'B' and 'F' are treated as a 0 score for tested students.)

3. Using the set of items do the following to calculate the percent at each score point.

- Step1 A: For each item, calculate the number of students at each score point. Adjust the common item counts by multiplying the common items' number of students at each score point by 0.5.

- Step 1 B: Calculate the total number of scores by summing up the number of students at each score point across the items in the subtopic

- Step 2: For each score point, sum up the (adjusted) number of students at the score point across the items in the subtopic. Divide the sum by total number of scores for the subtopic. Multiply that by 100 and round to the nearest whole number.

4. *Example*

	Common Prompt		Matrix Prompt 1	Matrix Prompt 2	Matrix Prompt 3	Matrix Prompt 4	Matrix Prompt 5
Item	**C1**	**C2**	**M1**	**M2**	**M3**	**M4**	**M5**
Subtopic	**1**	**1**	**1**	**2**	**2**	**2**	**3**
Student	**Student Item Score**						
A	3	4	2				
B	4	4					
C	2	1	3				
D	5	2		4			
E	3	2		1			
F	0	0			2		
G	1	2	1				
H	6	5	5				
I	2	2				1	
J	3	2				2	
K	5	4					4

Score Point	**Step 1 Number at each score point**						
Item	**C1**	**C2**	**M1**	**M2**	**M3**	**M4**	**M5**
Subtopic	***1***	***1***	***1***	***2***	***2***	***2***	***3***
0	0.5	0.5	0	0	0	0	0
1	0.5	0.5	1	1	0	1	0
2	1	2.5	1	0	1	1	0
3	1.5	0	1	0	0	0	0
4	0.5	1.5	0	1	0	0	1
5	1	0.5	1	0	0	0	0
6	0.5	0	0	0	0	0	0
Total	15			5			1

Score Point	**Step 2 Percent at each score point**		
Subtopic	**1**	**2**	**3**
0	7	0	0
1	13	40	0
2	30	40	0
3	17	0	0
4	13	20	100
5	17	0	0
6	3	0	0

Cumulative Total

5. Include the yearly results where the number tested is greater than or equal to 10
6. Cumulative total N (Enrolled, Not Tested Approved, Not Tested Other, Tested, at each achievement level) is the sum of the yearly results for each category where the number tested is greater than or equal to 10.
7. Cumulative percent for each achievement level is 100*(Number of students at the achievement level cumulative total / number of students tested cumulative total) rounded to the nearest whole number.
8. Cumulative mean scaled score is a weighted average. For years where the number tested is greater than or equal to 10, (sum of (yearly number tested * yearly mean scaled score)) / (sum of yearly number tested) rounded to the nearest whole number.

H. *Average Points Earned Students at Proficient Level (Range)*

2. Select all students across the states with Y40 scaled score, where Y=grade. Average the content area subcategories across the students and round to the nearest tenth. Add and subtract one standard error of measurement to get the range.

I. Writing Annotations

2. Students with a writing prompt score of 2-12 receive at least one, but up to five statements based on decision rules for annotations as outlined in Final Statements & Decision Rules for NECAP Writing Annotations.doc. Grade 11 students with the common writing prompt score of F or 0 will also receive annotations.

IX. Report Specific Rules

A. *Student Report*

1. Student header Information

a. If "FNAME" or "LNAME" is not missing then print "FNAME MI LNAME". Otherwise, print "No Name Provided".

b. Print the student's tested grade

c. For school and district name, print the abbreviated tested school and district ICORE name based on school type decision rules.

d. Print "NH","RI", or "VT" for state.

2. Test Results by content area

c. For students identified as "Not Tested", print the not tested reason in the achievement level, leave scaled score and graphic display blank.

d. For students identified as tested for the content area then do the following

VI. Print the complete achievement level name the student earned

VII. Print the scaled score the student earned

VIII. Print a vertical black bar for the student scaled score with gray horizontal bounds in the graphic display

IX. For students identified as "Tested with a non-standard accommodation" for a content area, print '**' after the content area earned achievement level and after student points earned for each subcategory.

X. For students identified as "Tested Incomplete" for a content area, place a section symbol after content area earned scaled score.

3. Grade 11 writing graphic display will not have standard error bars. Also, if a student's total points earned is 0 for writing, do not print the graphic display.

4. Exclude students based on school type and participation status decision rules for aggregations.

5. This Student's Achievement Compared to Other Students by content area

c. For tested students, print a check mark in the appropriate achievement level in the content area student column. For not tested students leave blank

d. For percent of students with achievement level by school, district and state print aggregate data based on school type and minimum N rules

6. This Student's Performance in Content Area Subcategories by content area

b. Always print total possible points and students at proficient average points earned range.

c. For students identified as not tested then leave student scores blank

d. For students identified as tested do the following

I. Always print student subcategory scores

II. If the student is identified as tested with a non-standard accommodation for the content area then place '**" after the student points earned for each subcategory.

e. Print aggregate data based on school type and minimum N-size rules.

5. Writing Annotations

a. For students with writing prompt score of 2-12 print at least one, but up to five annotation statements. Grade 11 students with the common writing prompt score of F or 0 will also receive annotations.

B. *School Item Analysis Report by Grade and Subject*

13. Reports are created for testing school and teaching school independently.

14. School Header Information

d. Use abbreviated ICORE school and district name based on school type decision rules

e. Print "New Hampshire", "Rhode Island", or "Vermont" for State.

f. For NH, the code should print SAU code – district code – school code. For RI and VT, the code should print district code – school code.

15. For multiple choice items, print '+' for correct response, or A,B,C,D,* or blank

16. For open response items, print the student score. If the score is not numeric ('B'), then leave blank.

17. For students identified as content area tested with non-standard accommodations, print '-' for invalidated items.

18. All students receive subcategory points earned and total points earned, including grade 11 writing.

19. Leave scaled score blank for not tested students and print the not tested reason in the achievement level column.

20. Exclude students based on school type and participation status decision rules for aggregations.

21. Always print aggregated data regardless of N-size based on school type decision rules.

22. For students identified as not tested for the content area print a cross symbol next to students' name.

23. For students identified as tested incomplete for the content area print a section symbol next to the scaled score.

24. Home school student are not listed on the report.

C. Grade Level School/District/State Results

9. Reports are run by testing state, testing district, testing school, teaching district, and teaching school using the aggregate school and district codes described in the school type table.

10. Exclude students based on school type and participation status decision rules for aggregations.

11. Report Header Information

c. Use abbreviated school and district name from ICORE based on school type decision rules.

d. Print "New Hampshire", "Rhode Island", or "Vermont" to reference the state. The state graphic is printed on the first page.

12. Report Section: Participation in NECAP

c. For testing level reports always print number and percent based on school type decision rules.

d. For the teaching level reports leave the section blank.

13. Report Section: NECAP Results by content area

    c. For the testing level report always print based on minimum N-size and school type decision rules.

    d. For the teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

14. Report Section: Historical NECAP Results by content area

    c. For teaching level report always print current year, prior years, and cumulative total results based on minimum N-size and school type decision rules.

    d. For teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

15. Report Section: Subtopic Results by content area

    b. For testing and teaching level reports always print based on minimum N-size and school type decision rules

16. Report Section: Disaggregated Results by content area

    c. For testing level report always print based on minimum N-size and school type decision rules.

    d. For teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules.

D. *School/District/State Summary*

1. Reports are run by testing state, testing district, testing school, teaching district, and teaching school using the aggregate school and district codes described in the school type table.
2. Exclude students based on school type and participation status decision rules for aggregations.
3. For testing level report print entire aggregate group across grades tested and list grades tested results based on minimum N-size and school type decision rules. Mean scaled score across the grades is not calculated.
4. For the teaching level report leave Enrolled, NT Approved, and NT Other blank. Print Tested, number and percent at each achievement level, mean scaled score based on minimum N-size and school type decision rules. Mean scaled score across the grades is not calculated.

X. Data File Rules

In the file names GG refers to the two digit grade (11) , YYYY refers to the year 0708, DDDDD refers to the district code, and SS refers to two letter state code.

A. State Student Cleanup Data

4. One CSV file per grade and state will be created based on the file layout NECAPYYYYF Gr 03-08 11 Student Demographic Cleanup File Layout.xls.
5. Refer to NECAPYYYYF Gr 03-08 11 Student Demographic Cleanup Description.doc
6. Session Invalidation Flags are marked as follows.

    a. If reaaccF02 or reaaccF03 is marked, then mark reaInvSes1, reaInvSes2, and reaInvSes3

    b. If mataccF03 is marked, then mark matInvSes1, matInvSes2, and matInvSes03. MataccF01 is left as marked on booklet.

    c. If wriaccF03 is marked, then mark wriInvSes1 and wriInvSes2

B. Preliminary State Results

3. A PDF file will be created for each state containing preliminary state results for each grade and subject and will list historical state data for comparison.

4. The file name will be SSPreliminaryResultsDATE.pdf

C. State Student Released Item Data

4. Students who tested at a private school are excluded from NH and RI student data files.

5. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 11 State Student Data Released Item Layout.xls

6. The CSV file name will be NECAP YYYY Fall State Student Data Released Item Gr GG.csv.

D. State Student Raw Data

1. Students who tested at a private school are excluded from NH and RI student data files.

2. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 11 State Student Raw Data File Layout.xls

3. The CSV file name will be NECAP YYYY Fall State Student Raw Data File Gr GG.csv.

E. State Student Scored Data

1. Students who tested at a private school are excluded from NH and RI student data files.

2. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 11 State Student Scored Data File Layout.xls

3. The CSV file name will be NECAP YYYY Fall State Student Scored Data File Gr GG.csv.

F. District Student Data

1. Students with the Discode or SendDiscode will be in the district grade specific CSV file for the testing year.

2. Students with a sprDiscode will be in the district grade specific CSV file for the teaching year.

3. Home school students are excluded from district student data files. For NH and RI only public school districts will receive district data files. (Districts with at least one school with schoolsubtypeID=1 in ICORE)

4. Testing and teaching CSV files will be created for each state and grade and district following the layout NECAP YYYY Fall Gr 11 District Student Data Layout.xls

5. The testing CSV file name will be NECAP YYYY Fall Testing District Slice Gr GG_DDDDD.csv. The teaching CSV file name will be NECAP YYYY Fall Teaching District Slice Gr GG_DDDDD.csv.

G. Item Information

1. An excel file will be created containing item information for common items: grade, subject, raw data item name, item type, key, and point value.

2. The file name will be NECAP YYYY Fall Gr 11 Item Information.xls

H. Grade Level Results Report Disaggregated and Historical Data

1. Teaching and testing CSV files will be created for each state and grade containing the grade level results disaggregated and historical data following the layout NECAP YYYY Fall Gr 11 Results Report Disaggregated and Historical Data Layout.xls.

2. Data will be suppressed based on minimum N-size and report type decision rules.

3. The testing file name will be NECAP YYYY Fall Testing Results Report Disaggregated and Historical Data Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Results Report Disaggregated and Historical Data Gr GG.csv.

I. Grade Level Results Report Participation Category Data

1. Teaching CSV file will be created for each state and grade containing the grade level results participation data following the layout NECAP YYYY Fall Gr 11 Results Report Participation Category Data Layout.xls.
2. The testing file name will be NECAP YYYY Fall Testing Results Report Participation Category Data Gr GG.csv

J. Grade Level Results Report Subtopic Data

1. Teaching and testing CSV files will be created for each state and grade containing the grade level results subtopic data following the layout NECAP YYYY Fall Gr 11 Results Report Subtopic Data Layout.xls.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. The testing file name will be NECAP YYYY Fall Testing Results Report Subtopic Data Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Results Report Subtopic Data Gr GG.csv.

K. Released Item Percent Responses Data

1. The CSV files will only contain state level aggregation for released items.
2. Teaching and testing CSV files will be created for each state and grade containing the released item analysis report state data following the layout NECAP YYYY Fall Gr 11 Released Item Percent Responses Layout.xls.
3. The testing file name will be NECAP YYYY Fall Testing Released Item Percent Responses.csv . The teaching file name will be NECAP YYYY Fall Teaching Released Item Percent Responses.csv.

L. Invalidated Students Original Score

1. Original raw scores for students whose responses were invalidated for reporting will be provided.
2. Students who tested at a private school are excluded from NH and RI student data files.
3. A CSV file will be created for each state and grade following the layout NECAP YYYY Fall Gr 11 State Invalidated Student Original Scored Data File Layout.xls.
4. The CSV file name will be NECAP YYYY Fall State Student Scored Data File Gr GG OriScInvStu.csv.

M. Multiple Choice Response Distribution Data Grades 11

1. One CSV file will be created containing the frequency of multiple responses (*) for multiple choice items.
2. All students are included in the frequencies.
3. The file will follow the layout NECAP YYYY Fall Multiple MC Responses Freq Layout.xls and will be named NECAP YYYY Fall Multiple MC Responses Freq.xls.

Addenda

2/4/2008: The writing student's at proficient extended response range on the student report will be '7'.

2/6/2008: Summary results data files will be created as follows:

1. Teaching and testing CSV files will be created for each state and grade containing the summary report data following the layout NECAP YYYY Fall Gr 11 Summary Results Layout.xls.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. The testing file name will be NECAP YYYY Fall Testing Summary Results Gr GG.csv . The teaching file name will be NECAP YYYY Fall Teaching Summary Results Gr GG.csv.