New England Common Assessment Program 2012–13 Technical Report

June 2013

100 EDUCATION WAY, DOVER, NH 03820 (800) 431-8901
WWW.MEASUREDPROGRESS.ORG

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
CHAPTER 2 TEST DESIGN AND DEVELOPMENT 3
2.1 TEST SPECIFICATIONS 3
2.1.1 Criterion-Referenced Test 3
2.1.2 Item Types 3
2.1.3 Description of Test Design 4
2.2 READING TEST SPECIFICATIONS 4
2.2.1 Standards 4
2.2.2 Item Types 4
2.2.3 Test Design 5
2.2.4 Blueprints 6
2.2.5 Depth of Knowledge 6
2.2.6 Passage Types 7
2.3 MATHEMATICS TEST SPECIFICATIONS 8
2.3.1 Standards 8
2.3.2 Item Types 8
2.3.3 Test Design 8
2.3.4 Blueprints 9
2.3.5 Depth of Knowledge 10
2.3.6 Use of Calculators and Reference Sheets 10
2.4 WRITING TEST SPECIFICATIONS 11
2.4.1 Standards 11
2.4.2 Item Types 12
2.4.3 Test Design 12
2.4.4 Blueprints 13
2.5 TEST DEVELOPMENT PROCESS 15
2.5.1 Item Development 15
2.5.2 Item Reviews at Measured Progress 16
2.5.3 Item Reviews at State Level 17
2.5.4 Bias and Sensitivity Review 18
2.5.5 Reviewing and Refining 18
2.5.6 Item Editing 18
2.5.7 Item Selection and Operational Test Assembly 19
2.5.8 Operational Test Draft Review 19
2.5.9 Alternative Presentations 20
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 22
3.5 DOCUMENTATION OF ACCOMMODATIONS 22
3.6 TEST SECURITY 23
3.7 TEST AND ADMINISTRATION IRREGULARITIES 24
3.8 TEST ADMINISTRATION WINDOW 24
3.9 NECAP SERVICE CENTER 24
CHAPTER 4 SCORING 25
4.1 SCORING OF STANDARD TEST ITEMS 25
4.1.1 Machine-Scored Items 25
4.1.2 Person-Scored Items 25
4.1.2.1 Scoring Location and Staff 26
4.1.2.2 Benchmarking Meetings with the NECAP State Specialists 27
4.1.2.3 Reader Recruitment and Qualifications 28
4.1.2.4 Methodology for Scoring Polytomous Items 28
4.1.2.5 Reader Training 29
4.1.2.6 Senior Quality Assurance Coordinator and Senior Reader Training 31
4.1.2.7 Monitoring of Scoring Quality Control and Consistency 31
CHAPTER 5 CLASSICAL ITEM ANALYSIS 35
5.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES 35
5.2 DIFFERENTIAL ITEM FUNCTIONING 38
5.3 DIMENSIONALITY ANALYSIS 39
CHAPTER 6 ITEM RESPONSE THEORY SCALING AND EQUATING 44
6.1 ITEM RESPONSE THEORY 46
6.2 ITEM RESPONSE THEORY RESULTS 48
6.3 EQUATING 49
6.4 EQUATING RESULTS 50
6.5 ACHIEVEMENT STANDARDS 51
6.6 REPORTED SCALED SCORES 52
CHAPTER 7 RELIABILITY 55
7.1 RELIABILITY AND STANDARD ERRORS OF MEASUREMENT 56
7.2 SUBGROUP RELIABILITY 57
7.3 REPORTING SUBCATEGORY RELIABILITY 57
7.4 INTERRATER CONSISTENCY 57
7.5 RELIABILITY OF ACHIEVEMENT LEVEL CATEGORIZATION 59
7.5.1 Accuracy and Consistency Results 60
CHAPTER 8 SCORE REPORTING 62
8.1 TEACHING YEARS VERSUS TESTING YEAR REPORTING 62
8.2 PRIMARY REPORTING DELIVERABLES 62
8.3 STUDENT REPORTS 63
8.4 ITEM ANALYSIS REPORTS 64

8.5 School and District Results Reports 65
8.6 School and District Summary Reports 69
8.7 School and District Student-Level Data Files 69
8.8 Analysis & Reporting System 70
8.8.1 Interactive Reports 70
8.8.1.1 Item Analysis Report 70
8.8.1.2 Achievement Level Summary 70
8.8.1.3 Released Items Summary Data 71
8.8.1.4 Longitudinal Data 71
8.8.2 User Accounts 71
8.9 Decision Rules 72
8.9.1 Quality Assurance 72
CHAPTER 9 VALIDITY 74
9.1 Questionnaire Data 75
9.1.1 Difficulty of Assessment 75
9.1.1.1 Difficulty: Reading 75
9.1.1.2 Difficulty: Mathematics 76
9.1.1.3 Difficulty: Writing 77
9.1.2 Content 78
9.1.2.1 Content: Reading 78
9.1.2.2 Content: Mathematics 80
9.1.2.3 Content: Writing 82
9.1.3 Homework 83
9.1.3.1 Homework: Reading 83
9.1.3.2 Homework: Mathematics 85
9.1.4 Performance in Courses 87
REFERENCES 89
APPENDICES 91
APPENDIX A COMMITTEE MEMBERSHIP
APPENDIX B PARTICIPATION RATES
APPENDIX C ACCOMMODATION FREQUENCIES BY CONTENT AREA
APPENDIX D TABLE OF STANDARD ACCOMMODATIONS
APPENDIX E ITEM-LEVEL CLASSICAL STATISTICS
APPENDIX F ITEM-LEVEL SCORE POINT DISTRIBUTIONS
APPENDIX G DIFFERENTIAL ITEM FUNCTIONING RESULTS
APPENDIX H ITEM RESPONSE THEORY CALIBRATION RESULTS
APPENDIX I TEST CHARACTERISTIC CURVES AND TEST INFORMATION FUNCTION PLOTS
APPENDIX J DELTA ANALYSES AND RESCORE ANALYSES
APPENDIX K α-PLOTS AND b-PLOTS
APPENDIX L PERFORMANCE LEVEL SCORE DISTRIBUTIONS
APPENDIX M RAW TO SCALED SCORE LOOK-UP TABLES
APPENDIX N SCALED SCORE DISTRIBUTIONS

APPENDIX O CLASSICAL RELIABILITIES
APPENDIX P INTERRATER AGREEMENT
APPENDIX Q DECISION ACCURACY AND CONSISTENCY RESULTS
APPENDIX R SAMPLE REPORTS
APPENDIX S 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 Maine, New Hampshire, Rhode Island, and Vermont 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 to provide: (a) data on student achievement in reading/English language arts and mathematics to meet the requirements of NCLB; (b) information to support program evaluation and improvement; and (c) 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; and
- 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 2012–13 NECAP tests. In October, 2012, students in grades 3 through 8 and 11 participated in the administration of the NECAP tests 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. Chapters 2 through 4 provide a description of the NECAP tests by covering the test design and development process, the administration of the tests, and scoring. Chapters 5 through 7 provide statistical and psychometric summaries, including chapters on item analysis, scaling and equating, and reliability. Chapter 8 is devoted to NECAP score reporting, and Chapter 9 is devoted to discussions on validity. Finally, the references cited throughout the report are provided, followed by the report appendices.

CHAPTER 2 TEST DESIGN AND DEVELOPMENT

2.1 TEST SPECIFICATIONS

2.1.1 Criterion-Referenced Test

Items on the NECAP tests are developed specifically for those states participating in the NECAP and are directly linked to the NECAP Grade Level Expectations/Grade Span Expectations (GLEs/GSEs). These GLEs/GSEs are the basis for the reporting categories developed for each content area and are used to help guide the development of test items. Although each item is designed to measure a specific GLE/GSE, an item may address several GLEs/GSEs within a strand.

2.1.2 Item Types

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

Multiple-choice 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 each requires 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 items were administered in grades 3 through 8 and 11 in 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. Each short-answer item requires 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.

Constructed-response items typically require students to use higher-order thinking skills such as summary, evaluation, and analysis in constructing a satisfactory response. Each constructed-response item requires approximately 5 to 10 minutes for most students to complete. These items were administered in grades 3 through 8 and 11 in reading, in grades 5 through 8 and 11 in mathematics, and in grades 5 and 8 in writing.

A common **writing prompt** was administered in grades 5, 8, and 11. In grade 11, an additional matrix prompt was administered to students. 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.

Approximately 25% of the common NECAP items were released to the public in 2012–13. The released NECAP items are posted on a Web site hosted by Measured Progress and on each Department of Education Web site. 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 tests.

2.1.3 Description of Test Design

The NECAP test is structured using both common and matrix items. Common items are taken by all students in a given grade level. Student scores are based only on common items. Matrix items are either new items included on the test for field-test purposes or equating items used to link one year's results to those of previous years. In addition, field-test and equating items are divided among the multiple forms of the test for each grade and content area. The number of test forms varies by grade and content area but ranges between eight and nine forms. Each student takes only one form of the test and therefore answers a fraction of the field-test items. Equating and field-test items are not distinguishable to students and have a negligible impact on testing time. Because all students participate in the field test, an adequate sample size is provided to produce reliable data that can be used to inform item selection for future tests.

2.2 READING TEST SPECIFICATIONS

2.2.1 Standards

The test framework for reading in grades 3 through 8 was based on the NECAP GLEs, and each item on the NECAP tests was designed to measure a specific GLE. The test framework for reading in grade 11 was based on the NECAP GSEs, and each item on the NECAP test was designed to measure a specific GSE.

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 multiple-choice items.

2.2.2 Item Types

The NECAP reading tests include multiple-choice and constructed-response items. Multiple-choice items require students to demonstrate a wide range of knowledge and skills, requiring one minute of response time per item. Constructed-response items are more complex, requiring 5 to 10 minutes of response time per item. Each type of item is worth a specific number of points in the student's total reading score, as shown in Table 2-1.

Table 2-1. 2012–13 NECAP: Reading Item Types

Item Type	*Possible Score Points*
MC	0 or 1
CR	0, 1, 2, 3, or 4

MC = multiple-choice; CR = constructed-response

2.2.3 Test Design

Table 2-2 summarizes the numbers and types of items that were used in the 2012–13 NECAP reading tests 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 multiple-choice item was worth one point and each constructed-response item was worth four points.

Table 2-2. 2012–13 NECAP: Item Type and Number of Items—Reading Grades 3–8

	Forms	*Passages*		*Stand-alone*	*Total*	
		Long	*Short*	*MC*	*MC*	*CR*
Common		2	2	4	28	6
Matrix—Equating	1–3	3	3	6	36	9
Matrix—FT	4–7	4	4	8	48	12
	8–9	2	2	4	24	6
Total per Student	1–7	3	3	6	36	9
	8–9	3	3	6	36	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 2-3 summarizes the numbers and types of items that were used in the 2012–13 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 multiple-choice item was worth one point and each constructed-response item was worth four points.

Table 2-3. 2012–13 NECAP: Item Type and Number of Items—Reading Grade 11

	Forms	*Passages*		*Stand-alone*	*Total*	
		Long	*Short*	*MC*	*MC*	*CR*
Common		2	2	4	24	6
Matrix—Equating	1–2	2	2	4	24	6
Matrix—FT	3–8	6	6	12	72	18
Total per Student		3	3	6	36	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.

2.2.4 Blueprints

The distribution of emphasis for reading is shown in Table 2-4.

Table 2-4. 2012–13 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 2-5 shows the content area subcategory 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 2-5. 2012–13 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		18	17	10	10	10	10	9
Type of Text	Literary	18	19	21	21	21	21	21
	Informational	16	16	21	21	21	21	22
Level of Comprehension	Initial Understanding	18	20	20	21	18	19	20
	Analysis and Interpretation	16	15	22	21	24	23	23
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).

2.2.5 Depth of Knowledge

Each item on the NECAP tests in reading is assigned a Depth of Knowledge (DOK) level according to the cognitive demand of the item. DOK is not synonymous with difficulty. The DOK level rates the complexity of the mental processing a student must use to answer the question. Each of the three levels is described in Figure 2-1.

Figure 2-1. 2012–13 NECAP: Depth of Knowledge—Reading

Level 1 (Recall)	This level requires students to receive or recite facts or to use simple skills or abilities. Oral reading that does not include analysis of the text as well as basic comprehension of a text is included. Items require only a shallow understanding of text presented and often consist of verbatim recall from text or simple understanding of a single word or phrase.
Level 2 (Skill/Concept)	This level includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Intersentence analysis of inference is required. Some important concepts are covered but not in a complex way.
Level 3 (Strategic Thinking)	This level requires students to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standards and items involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or application of prior knowledge. Items may also involve more superficial connections between texts.

Table 2-6 lists the percentage of actual score points assigned to each DOK level in reading.

Table 2-6. 2012–13 NECAP: Depth of Knowledge in Terms of Percentage of Test by Grade—Reading Grades 3–8 and 11

DOK	*Grade*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Level 1	62%	54%	15%	10%	15%	11%	25%
Level 2	38%	38%	85%	90%	85%	81%	67%
Level 3	0%	8%	0%	0%	0%	8%	8%
TOTAL	100%	100%	100%	100%	100%	100%	100%

2.2.6 Passage Types

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; dramas; 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.

2.3 MATHEMATICS TEST SPECIFICATIONS

2.3.1 Standards

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

- 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.

2.3.2 Item Types

The NECAP mathematics tests include multiple-choice, short-answer, and constructed-response items. Short-answer items require students to perform a computation or solve a problem, requiring two to five minutes of response time per item. Constructed-response items are more complex, requiring 8 to 10 minutes of response time per item. Each type of item is worth a specific number of points in the student's total mathematics score, as shown in Table 2-7.

Table 2-7. 2012–13 NECAP: Mathematics Item Types

Item Type	*Possible Score Points*
MC	0 or 1
SA	0, 1, or 2
CR	0, 1, 2, 3, or 4

MC = multiple-choice; SA = short-answer; CR = constructed-response

2.3.3 Test Design

Table 2-8 summarizes the numbers and types of items that were used in the 2012–13 NECAP mathematics tests for grades 3 and 4, 5 through 8, and 11, respectively. Note that all students received the common items plus equating and field-test items in their forms. Each multiple-choice item was worth one

point, each short-answer item was worth either one or two points, and each constructed-response item was worth four points. Except for grade 11, score points within a grade level were evenly divided so that multiple-choice items represented approximately 50% of the possible score points and short-answer and constructed-response items together represented approximately 50% of the possible score points. In grade 11, multiple-choice items represented approximately 40% of the possible score points and short-answer and constructed-response represented approximately 60% of the possible score points.

Table 2-8. 2012–13 NECAP: Item Type and Number of Items—Mathematics

Content Area and Grade	*Common*				*Matrix–equating*				*Matrix–FT*				*Total per student*			
	MC	*SA1*	*SA2*	*CR*	*MC*	*SA1*	*SA2*	*CR*	*MC*	*SA1*	*SA2*	*CR*	*MC*	*SA1*	*SA2*	*CR*
3–4	35	10	10		6	2	2		3	1	1		44	13	13	
5–8	32	6	6	4	6	2	2	1	3	1	1	1	41	9	9	6
11	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; CR = constructed-response
For grades 3 and 4 and 5–8, total of nine forms per grade; six contained unique matrix-equating items.
For grade 11, total of eight forms; six contained unique matrix-equating items.

2.3.4 Blueprints

The distribution of emphasis for NECAP content strands for mathematics is shown in Table 2-9.

Table 2-9. 2012–13 NECAP: Distribution of Emphasis in Terms of Target Percentage of Test by Grade—Mathematics 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)*
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. 2012–13 NECAP: Reporting Subcategories and Possible Raw Score Points by Grade—Mathematics Grades—3–8 and 11

Subcategory	*Grade tested*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Numbers and Operations	35	32	30	26	19	13	10
Geometry and Measurement	10	13	13	17	17	17	19
Functions and Algebra	10	10	13	13	20	26	25
Data, Statistics, and Probability	10	10	10	10	10	10	10
TOTAL	65	65	66	66	66	66	64

2.3.5 Depth of Knowledge

Each item on the NECAP test in mathematics is assigned a DOK level according to the cognitive demand of the item. DOK is not synonymous with difficulty. The DOK level rates the complexity of the mental processing a student must use to solve a problem. Each of the three levels is described in Figure 2-2.

Figure 2-2. 2012–13 NECAP: Depth of Knowledge—Mathematics

Level 1 (Recalling Information and Carrying Out Simple Procedures)	This level requires the recall of a fact, definition, term, or simple procedure; the application of a formula; or the performance of a straight algorithmic procedure. Items at this level may require students to demonstrate a rote response.
Level 2 (Skill/Concept)	This level requires mental processing beyond that of a simple habitual response. These items often require students to make some decisions about how to approach a problem.
Level 3 (Strategic Thinking, Reasoning, Planning, Drawing Conclusions, and Using Concepts and Evidence)	This level requires students to develop a plan or sequence of steps. These items are more complex and abstract than the items at the previous two levels. These items may also have more than one possible answer and may require students to use evidence, make conjectures, or justify their answers.

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

Table 2-11. 2012–13 NECAP: Depth of Knowledge in Terms of Percentage of Test by Grade—Mathematics Grades 3–8 and 11

DOK	*Grade*						
	3	*4*	*5*	*6*	*7*	*8*	*11*
Level 1	34	18	29	24	33	23	31
Level 2	54	74	64	64	61	59	66
Level 3	12	8	7	12	6	18	3
TOTAL	100	100	100	100	100	100	100

2.3.6 Use of Calculators and Reference Sheets

The mathematics specialists from the New Hampshire, Rhode Island, Maine, and Vermont 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 tests and permitted in the remaining two sessions. It was decided that, at grade 11, calculators should be prohibited in the first of the two sessions and permitted in the second session.

Reference sheets are provided to students at grades 5–8 and 11. These sheets contain information, such as formulas, that students may need to answer certain test items. The reference sheets are published each year with the released items and have remained the same for several years over the various test administrations. Toolkits are provided to students at grades 3–5. These toolkits contain manipulatives to answer specific questions. The toolkits are designed for specific items and therefore change annually. They are published with the released items. All students in grades 3–8 and 11 receive rulers for use on the mathematics test. Students in grades 5–7 receive protractors for use on the mathematics test. Students may keep the rulers and protractors after test administration.

2.4 Writing Test Specifications

2.4.1 Standards

Grades 5 and 8

The test framework for grades 5 and 8 writing was based on the NECAP GLEs, and each item on the NECAP tests was 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)

Grade 11

The test framework for grade 11 writing was based on the NECAP GSEs, and each item on the NECAP test was designed to measure a specific GSE. The content standards for grade 11 writing identify six genres:

- writing in response to literary text

- writing in response to informational text
- report writing
- procedural writing
- persuasive writing
- reflective writing

2.4.2 Item Types

The NECAP writing tests include multiple-choice items, constructed-response items, and extended-response writing prompts. At grades 5 and 8, multiple-choice items provide breadth of coverage of the GLEs/GSEs, requiring approximately one minute for most students to answer each item. Constructed-response items are more complex, requiring 5 to 10 minutes of response time per item. At grades 5, 8, and 11, students are required to respond to a writing prompt, receiving 45 minutes (plus additional time if necessary) to develop a response. Each type of item is worth a specific number of points in the student's total writing score, as shown in Table 2-12.

Table 2-12. 2012–13 NECAP: Writing Item Types

Item Type	*Possible Score Points*
MC	0 or 1
CR	0, 1, 2, 3, or 4
ER	0, 2–12

MC = multiple-choice; CR=constructed-response; ER = extended-response writing prompt

2.4.3 Test Design

All items on the grades 5 and 8 writing tests were common. Each multiple-choice item was worth one point, each constructed-response item was worth four points, and the extended-response writing prompt was worth 12 points.

Table 2-13 summarizes the test design used in the 2012–13 NECAP writing test for grade 11. There were a total of eight forms: five equating forms and three field-test forms. Each grade 11 student responded to two different extended-response writing prompts, one common and either one matrix-equating or one field-test prompt depending on the form. The common prompt was worth 12 points.

Table 2-13. 2012–13 NECAP: Number of Items by Item Type—Writing Grade 11

Common	*Matrix–Equating (5 Forms)*	*Matrix–Field Test (3 Forms)*
1 ER	1 ER per form	1 ER per form

ER = extended-response writing prompt

2.4.4 Blueprints

Grades 5 and 8

The writing prompt and the three constructed-response items each address a different genre. In addition, structures of language and writing conventions are assessed through multiple-choice items but are also assessed through the extended-response writing prompts and the constructed-response items. The extended-response writing prompts and constructed-response 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 write 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 Response" lists the total raw score points from the three constructed-response items; the reporting category "Extended Response" lists the total raw score points from the writing prompt.

Table 2-14. 2012–13 NECAP: Reporting Subcategory and Possible Raw Score Points Possible by Grade—Writing Grades 5 and 8

Subcategory	*Grade Tested*	
	5	*8*
Structures of Language and Writing Conventions	10	10
Short Response	12	12
Extended Response	12	12
TOTAL	34	34

Short response = constructed-response items;
Extended response = writing prompt

Grade 11

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.

For grade 11 writing, there is only one reporting category, "Extended Response," with a total possible raw score of 12 points. One hundred percent of the raw score points for writing was assigned to DOK Level 3.

Each item on the NECAP test in writing is assigned a DOK level according to the cognitive demand of the item. DOK is not synonymous with difficulty. The DOK level rates the complexity of the mental processing a student must use to answer the question. Each of the three levels is described in Figure 2-3:

Figure 2-3. 2012–13 NECAP: Depth of Knowledge—Writing

Level 1	This level requires the student to write or recite simple facts. This writing or recitation does not include complex synthesis or analysis but basic ideas.
Level 2	This level requires some mental processing. Students are beginning to connect ideas using a simple organizational structure. For example, students may be engaged in note-taking, outlining, or simple summaries.
Level 3	This level requires some higher-level mental processing. Students are engaged in developing compositions that include multiple paragraphs. These compositions may include complex sentence structure and may demonstrate some synthesis and analysis. Students show awareness of their audience and purpose through focus, organization, and the use of appropriate compositional elements. The use of appropriate compositional elements includes such things as addressing chronological order in a narrative or including supporting facts and details in an informational report.

Table 2-15 lists the percentage of actual score points assigned to each level of DOK in writing for grades 5 and 8.

Table 2-15. 2012–13 NECAP: Depth of Knowledge by Grade (in Percentage of Test)—Writing Grades 5 and 8

DOK	*Grade Tested*	
	5	*8*
Level 1	24%	24%
Level 2	41%	41%
Level 3	35%	35%
TOTAL	100%	100%

Table 2-16 lists the percentage of actual score points assigned to each level of DOK in writing for grade 11.

Table 2-16. 2012–13 NECAP: Depth of Knowledge in Terms of Percentage of Test—Writing Grade 11

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

* In grade 11, 100% of the writing test is assigned to DOK Level 3.

2.5 TEST DEVELOPMENT PROCESS

2.5.1 Item Development

Items used on the NECAP tests are developed and customized specifically for use on the NECAP tests and are consistent with NECAP GLE and GSE content standards. Measured Progress test developers work with Maine, New Hampshire, Rhode Island, and Vermont educators to verify the alignment of items to the appropriate NECAP content standards.

The development process combined the expertise of Measured Progress test developers and committees of educators to help ensure items meet the needs of the NECAP tests. All items used on the common portions of the NECAP tests were reviewed by a committee of content area experts and by a committee of bias experts. Tables 2-17 through 2-20 show the number of items developed within each content area for the 2012–2013 NECAP tests.

Table 2-17. 2012–13 NECAP: Annual English Language Arts Item Development—Grades 3–8

Passages	*MC*	*CR*
5 long passages (divided by literary and informational)	40	10
5 short passages (divided by literary and informational)	80	15
Standalones	20	
10 total passages	140	25

MC = multiple-choice; CR = constructed-response

Table 2-18. 2012–13 NECAP: Annual English Language Arts Item Development—Grade 11

Passages	*MC*	*CR*
5 long passages (divided by literary and informational)	40	10
5 short passages (divided by literary and informational)	80	15
Standalones	20	
10 total passages	140	25

MC = multiple-choice; CR = constructed-response

Table 2-19. 2012–13 NECAP: Annual Writing Item Development—Grade 11

Grade	*ER*
11	6

ER = extended-response writing prompt

Table 2-20. 2012–13 NECAP: Annual Mathematics Item Development —Grades 3–8 and 11

Grade	*MC*	*SA1*	*SA2*	*CR*
3	27	9	9	NA
4	27	9	9	NA
5	27	9	9	9
6	27	9	9	9
7	27	9	9	9
8	27	9	9	9
11	32	16	8	8

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

2.5.2 Item Reviews at Measured Progress

For the internal item review, the lead Measured Progress test developer within the content area performed the following activities:

- review of the formatted items, open-response scoring guides, and any reading selections and graphics
- evaluation of item integrity, content, and structure; appropriateness to designated content area; 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)
- assurance that, for each item, there was only one correct answer
- analysis of scorability and evaluation as to whether the scoring guide adequately addressed performance on the item

Fundamental questions the lead test 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 open-response 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, DOK, and identified contract)?
- Is the item appropriate for the designated grade level?

2.5.3 Item Reviews at State Level

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 four 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 2011–12 NECAP IRCs for each content area in grade levels 3 through 8 and 11 met in the spring of 2010. Committee members reviewed the entire set of embedded field-test items proposed for the 2011–12 operational test and made recommendations about selecting, revising, or eliminating specific items from the item pool. Members reviewed each item against the following criteria:

- GLE/GSE Alignment
  - Is the test item aligned to the appropriate GLE/GSE?
  - If the test item is not aligned to the appropriate GLE/GSE, then which GLE/GSE or grade level is more appropriate?
- Content 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 DOK?
  - If consensus cannot be reached, is there clarity around why the item might be on the borderline of two levels?
- Universal Design
  - Is there an appropriate use of simplified language? (Does it 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 Braille?
  - Is the item language clear?

[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.

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

2.5.4 Bias and Sensitivity Review

Bias review is an essential part of the development process. NECAP passages and items were reviewed by a committee of teachers, English language learner 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.) Passages and 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.

2.5.5 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 Maine, New Hampshire, Rhode Island, and Vermont Departments of Education content area specialists made the final selections with the assistance of Measured Progress test developers at a final face-to-face meeting.

2.5.6 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;
- 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 content area, 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.

2.5.7 Item Selection and 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 2012–13 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, short-answer, and constructed-response 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 set 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.

2.5.8 Operational Test Draft Review

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 and *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).

2.5.9 Alternative Presentations

Common items in each grade-level test were 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 enlarged to a 20-point font to create a large-print version of the test for visually impaired students.

CHAPTER 3 TEST ADMINISTRATION

3.1 RESPONSIBILITY FOR ADMINISTRATION

The *2012 NECAP Principal/Test Coordinator Manual* indicated that principals and/or their designated NECAP test coordinators were responsible for the proper administration of the NECAP tests. 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 coordinators were instructed to read the *Principal/Test Coordinator Manual* before testing and to be familiar with the instructions provided in the grade-level *Test Administrator Manual*. The *Principal/Test Coordinator Manual* 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 Manual* described the testing material sent to each school and 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 tests. 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 returned it to Measured Progress.

3.3 PARTICIPATION REQUIREMENTS AND DOCUMENTATION

The Departments of Education's intent is for all students in grades 3 through 8 and 11 to participate in the NECAP tests through either a standard administration an administration with accommodations, or via an alternate assessment. Furthermore, any student who is absent during any session of the NECAP tests 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 previous school year. Students who were alternately assessed in the 2011–12 school year were not required to participate in the NECAP in 2012–13. 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 Booklet 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, 2011, and is LEP (reading and writing only).

  A. 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 2012; however, these students were required to take the mathematics test in 2012.

- Student withdrew from school after October 1, 2012.

  B. If a student withdrew after October 1, 2012, 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, 2012.

  C. If a student enrolled after October 1, 2012, 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.

  D. 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, 2012, and did not complete test for reasons other than those listed above.

  E. 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."

Appendix B lists the participation rates of the four NECAP states combined in reading, mathematics, and writing.

3.4 Administrator Training

In addition to distributing the *Principal/Test Coordinator Manual* and *Test Administrator Manual*, the Maine, New Hampshire, Rhode Island, and Vermont Departments of Education, along with Measured Progress, conducted test administration workshops in each state to inform school personnel about the NECAP tests and to provide training on the policies and procedures regarding administration of the tests. In addition, an audio PowerPoint test administration workshop presentation was also prerecorded and provided to each state for inclusion on their Department of Education Web site.

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 by students—after testing was completed.

The first list of allowable accommodations was created by the three original NECAP states (New Hampshire, Rhode Island, and Vermont) at the beginning of the program in 2004. The list was later reviewed and revised in 2009 when the state of Maine joined the program. The four NECAP states worked together to change the coding system, revise existing accommodations, and add or delete certain accommodations. The resulting Table of Standard Test Accommodations is divided into accommodations for timing, setting, presentation, and response. Each accommodation is listed with details on how to deliver it to students. A *NECAP Accommodations Guide* was also produced to provide additional details on planning for and implementing accommodations. This guide was available on each state's Department of Education Web site. The states collectively made the decision that accommodations would continue to be made available to all students based on individual need, regardless of disability status. Decisions regarding accommodations were to be made by the student's educational team on an individual basis and were to be consistent with those used during the student's 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 (O)" and/or "Modifications (M)" sections.

Appendix C shows the accommodation frequencies by content area for the October, 2012 NECAP test administration. The accommodation codes (T1–T4, S1–S2, P1–P11, R1–R7, O1 and O2, and M1–M3) are defined in the Table of Standard Test Accommodations, which can 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 four states. The *Principal/Test Coordinator Manual* and *Test Administrator Manual* 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 or the principal or both 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 during test administration webinars as well as workshops that were conducted in all four states. The four 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 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

There were no test irregularities reported during the October, 2012 NECAP test administration.

3.8 Test Administration Window

The test administration window was October 1 through October 24, 2012.

3.9 NECAP Service Center

To provide additional support to schools before, during, and after testing, Measured Progress operates the NECAP Service Center. The support of a service center is essential to the successful administration of any statewide test program. The service center 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 that includes notes regarding the issue and resolution of each call.

The service center was staffed year-round and was available to receive calls from 8:00 a.m. to 4:00 p.m. Monday through Friday. Extra representatives were available as needed, beginning approximately two weeks before the start of the testing window and ending two weeks after the end of the testing window to assist with handling the additional call volume.

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

Multiple-choice 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 or is upside down or backward, 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 correct the situation.

4.1.2 Person-Scored Items

The images of student responses to constructed-response 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 on-site. (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, New Hampshire. Table 4-1 presents the locations where 2012–13 NECAP test item responses by grade and content area were scored.

Table 4-1. 2012–13 NECAP: Operational Scoring Locations by Content Area and Grade

Content Area	*Grade*	*Dover, NH*	*Menands, NY*	*Longmont, CO*
Mathematics	3			X
	4			X
	5			X
	6			X
	7			X
	8	X		
	11			X
Reading	3	X		
	4	X		
	5			X
	6			X
	7			X
	8			X
	11	X		
Writing	5		X	
	8		X	
	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 2012–13 NECAP responses:

- The NECAP scoring project manager, an employee of Measured Progress, was located in Dover, New Hampshire, 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, New Hampshire, and coordinated technical communication across all scoring sites.
- A chief reader in each content area (mathematics, reading, and writing) ensured consistency of scoring across all scoring sites for all grades tested in that content area. Chief readers also provided read-behind activities (defined in Section 4.1.2.7) for Quality Assurance Coordinators (QACs). Chief readers are employees of Measured Progress.
- Numerous 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, 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 senior readers to readers was approximately 1:11.
- Readers at scoring sites scored operational and field-test NECAP 2012–13 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 content area 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 that chief readers identified as illustrative midrange examples of the respective score points were selected. Chief readers presented these responses to the NECAP state content area 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 midrange 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 2012–13 NECAP tests, Measured Progress actively sought a diverse scoring pool representative of the population of the four NECAP states. The broad range of reader backgrounds included scientists, editors, business professionals, authors, teachers, graduate school students, and retired educators. Demographic information (e.g., gender, race, educational background) about readers 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., résumé and/or transcripts) of their qualifications.

Table 4-2 summarizes the qualifications of the 2012–13 NECAP scoring leadership and readers.

Table 4-2. 2012–13 NECAP: Qualifications of Scoring Leadership and Readers—Fall Administration

Scoring Responsibility	*Educational Credentials*				*Total*
	Doctorate	*Master's*	*Bachelor's*	*Other*	
Scoring Leadership	5.0%	32.7%	56.4%	5.9%[1]	100.0%
Readers	6.0%	34.6%	49.9%	9.5%[2]	100.0%

Scoring Leadership = chief readers, QACs, and senior readers
[1] 4 QAC/senior readers had associate's degrees and 2 had at least 48 college credits.
[2] 36 readers had associate's degrees and 16 had 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. 2012–13 NECAP: Possible Score Points for Polytomous Item Types

Polytomous Item Type	*Possible Score Point Range*
Writing prompt	0–6
Constructed-response	0–4
2-point short-answer (SA2)	0–2
1-point short-answer (SA1)	0–1
Nonscorable items	0

Nonscorable Items

Readers could designate a response as nonscorable 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 booklet (electronic copy or hard copy, as needed) to identify the correct location (in the answer booklet) 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-blind scoring. Single-scored items were scored by one reader. Double-blind scored items were scored independently by two readers, whose scores were tracked for "interrater agreement." (For further information on double-blind scoring and the interrater agreement, see Section 4.1.2.7 and Appendix P.)

4.1.2.5 Reader Training

Reader training began with an introduction of the on-site scoring staff and an overview of the NECAP'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 reviewed or scored the particular response set organized for that training: anchor sets, training sets, and qualifying sets. (These are defined on the following page.)

During training, readers could highlight or mark hard copies of the anchor and training sets (as well as the first qualifying sets after the qualification round), even if all or part of the set was also presented online.

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 area specialists representing the four 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 scoring staff. Each set contains one client-approved sample response per score point considered to be a midrange exemplar. The set included a second sample response if there was 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 a group discussion of one or two responses, directing readers' attentions to difficult scoring issues (e.g., the borderline between two score points). Throughout the training, trainers modeled 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 composed of constructed-response items, writing prompts, and all 2-point short-answer items for grades 3 and 4 mathematics. The 10 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 made available to readers after the qualification round 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 predetermined score on at least eight

of the ten responses) and at least 90% exact or adjacent agreement (i.e., to exactly match or be within one score point of the predetermined score on nine or ten of the ten responses), except 70% and 90%, respectively, for 6-point writing-prompt responses. In other words, readers were allowed one discrepant score (i.e., one score of 10 that was more than 1score point from the predetermined score) provided they had at least eight exact scores (seven for writing-prompt items).

To be eligible to score 1-point short-answer mathematics items (which were benchmarked "right" or "wrong") and 2-point short-answer 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 and Senior Reader Training

QACs and select senior readers were trained in a separate training session immediately prior to reader training. In addition to a discussion of the items and their responses, QAC and senior reader 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 senior readers to handle questions from the readers.

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 chief reader as appropriate (described on the following page), 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 stricter 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, it 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. The use of multiple monitoring techniques is critical toward monitoring reader accuracy during the process of live scoring.

Embedded CRRs

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 five and thirty 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 below).

Embedded CRRs were employed for all constructed-response items. They were not used for writing 6-point extended-response items, because these are 100% double-blind scored (defined on the following page). Embedded CRRs were also not used for mathematics 2-point short-answer 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 senior reader) scoring a response after a reader has already scored the response. The practice was applied to all open-response 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 senior reader. The iScore system allowed one, two, or three responses per reader to be placed into the read-behind queue at a time.

The senior reader entered his or her score into iScore before being allowed to see the reader's score. The senior reader 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 one point), the senior reader's score became the score of record. (A significant number of adjacent scores for a reader triggered an individual scoring consultation with the senior reader, after which the QAC determined whether or when the reader could resume scoring.)
- If the scores were discrepant (i.e., differed by more than one point), the senior reader's score became the score of record. (This triggered an individual consultation for the reader with the senior reader, after which the QAC determined whether or when the reader could resume scoring on that item.)

Table 4-4 illustrates how scores were resolved by read-behind.

Table 4-4. 2012–13 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/senior reader's score

Senior readers were tasked with conducting, on average, five read-behinds per reader throughout each half scoring day; however, senior readers 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 whether the response was to be double-blind scored. The practice was applied to all open-response item types. Table 4-5 shows by which method(s) the responses to both common and equating open-response item types for each operational test were scored.

Table 4-5. 2012–13 NECAP: Frequency of Double-Blind Scoring by Grade and Content

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

If there was a discrepancy (a difference greater than one score point) between double-blind scores, the response was placed into an arbitration queue. Arbitration responses were reviewed by scoring leadership (senior reader or QAC) without knowledge of the two readers' scores. Scoring leadership assigned the final score. Appendix P provides the NECAP 2012–13 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 on that item. Once the reader was allowed to resume scoring, scoring leadership carefully monitored the reader's accuracy by increasing the number of read-behinds.

Recalibration Sets

To determine whether readers were still calibrated to the scoring standard, they 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.

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.

Recalibration sets were employed for all constructed-response items. They were not used for writing 6-point extended-response items, which were 100% double-blind scored. They were also not used for 2-point short-answer 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.

CHAPTER 5 CLASSICAL ITEM ANALYSIS

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 *Standards for Educational and Psychological Testing* (AERA et al., 1999) and *Code of Fair Testing Practices in Education* (2004) include standards for identifying quality items. Items should assess only knowledge or skills that are identified as part of the domain being tested and should avoid assessing irrelevant factors. Items should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. In addition, items must not unfairly disadvantage students, in particular racial, ethnic, or gender groups.

Both qualitative and quantitative analyses are conducted to ensure that NECAP items meet these standards. Qualitative analyses are described in earlier chapters of this report; this chapter focuses on quantitative evaluations. Statistical evaluations are presented in four parts: (1) difficulty indices, (2) item-test correlations, (3) differential item functioning statistics, and (4) dimensionality analyses. The item analyses presented here are based on the statewide administration of NECAP in fall 2012. Note that the information presented in this chapter is based on the items common to all forms, since those are the items on which student scores are calculated. (Item analyses are also performed for field-test items, and the statistics are then used during the item review process and form assembly for future administrations.)

5.1 CLASSICAL DIFFICULTY AND DISCRIMINATION INDICES

All multiple-choice and constructed-response items are evaluated in terms of item difficulty according to standard classical test theory practices. Difficulty is defined as the average proportion of points achieved on an item and is measured by obtaining the average score on an item and dividing it by the maximum possible score for the item. Multiple-choice and one-point short-answer items are scored dichotomously (correct versus incorrect); so, for these items, the difficulty index is simply the proportion of students who correctly answered the item. Polytomously scored items include 2-point short-answer items, for which students can receive scores of 0, 1, or 2, and constructed-response items, which are worth 4 points total. By computing the difficulty index as the average proportion of points achieved, the indices for the different item types are placed on a similar scale, ranging from 0.0 to 1.0 regardless of the item type. Although this index is traditionally described as a measure of difficulty, it is properly interpreted as an easiness index, because larger values indicate easier items. An index of 0.0 indicates that all students received no credit for the item, and an index of 1.0 indicates that all students received full credit for the item.

Items that are answered correctly by almost all students provide little information about differences in student abilities, but do indicate knowledge or skills that have been mastered by most students. Similarly, items that are correctly answered by very few students provide little information about differences in student abilities, but may indicate knowledge or skills that have not yet been mastered by most students. In general, to

provide the best measurement, difficulty indices should range from near-chance performance of 0.25 (for four-option multiple-choice items or essentially 0 for constructed-response items) to 0.90, with the majority of items generally falling between approximately 0.4 and 0.7 for the lower mathematics grades and between about 0.6 and 0.8 for reading, writing, and the higher mathematics grades. However, 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.

A desirable characteristic of an item is for higher-ability students to perform better on the item than lower-ability students do. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of the 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 constructed-response items, the item discrimination index used was the Pearson product-moment correlation; for multiple-choice items, the corresponding statistic is commonly referred to as a point-biserial correlation. The theoretical range of these statistics is –1.0 to 1.0, with a typical observed range from 0.2 to 0.7.

Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the discrimination index can be thought of as a measure of construct consistency.

A summary of the item difficulty and item discrimination statistics for each content area and grade is presented in Table 5-1. Note that the statistics are presented for all items as well as by item type (multiple-choice, short-answer, constructed-response, and, for writing, writing prompt). Note also that only a single writing prompt is administered in grades 5 and 8; thus, it is not possible to calculate standard deviations of the difficulty and discrimination values. Furthermore, because the grade 11 writing test consists solely of a single prompt, no discrimination values or standard deviations could be calculated. The mean difficulty and discrimination values shown in the table are within generally acceptable and expected ranges.

Table 5-1. 2012–13 NECAP: Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade

Content Area	*Grade*	*Item Type*	*Number of Items*	*p-Value*		*Discrimination*	
				Mean	*Standard Deviation*	*Mean*	*Standard Deviation*
Mathematics	3	ALL	55	0.65	0.18	0.43	0.08
		MC	35	0.69	0.17	0.41	0.08
		SA	20	0.57	0.19	0.46	0.09
	4	ALL	55	0.65	0.16	0.42	0.08
		MC	35	0.68	0.17	0.39	0.06
		SA	20	0.59	0.13	0.47	0.09
	5	ALL	48	0.59	0.11	0.46	0.10
		CR	4	0.52	0.03	0.63	0.08
		MC	32	0.62	0.11	0.42	0.07
		SA	12	0.55	0.10	0.50	0.08

continued

Content Area	Grade	Item Type	Number of Items	p-Value		Discrimination	
				Mean	Standard Deviation	Mean	Standard Deviation
Mathematics	6	ALL	48	0.57	0.11	0.45	0.11
		CR	4	0.45	0.09	0.67	0.03
		MC	32	0.59	0.09	0.41	0.07
		SA	12	0.56	0.14	0.50	0.09
	7	ALL	48	0.55	0.14	0.46	0.10
		CR	4	0.33	0.05	0.64	0.06
		MC	32	0.60	0.11	0.42	0.07
		SA	12	0.46	0.14	0.51	0.09
	8	ALL	48	0.57	0.15	0.45	0.13
		CR	4	0.44	0.13	0.70	0.02
		MC	32	0.60	0.13	0.39	0.09
		SA	12	0.52	0.18	0.53	0.12
	11	ALL	46	0.45	0.18	0.50	0.13
		CR	4	0.31	0.12	0.71	0.01
		MC	24	0.54	0.18	0.42	0.10
		SA	18	0.35	0.12	0.56	0.06
Reading	3	ALL	34	0.67	0.15	0.41	0.09
		CR	6	0.57	0.24	0.50	0.08
		MC	28	0.69	0.12	0.39	0.08
	4	ALL	34	0.69	0.16	0.43	0.11
		CR	6	0.62	0.21	0.53	0.06
		MC	28	0.70	0.15	0.41	0.10
	5	ALL	34	0.68	0.18	0.43	0.08
		CR	6	0.38	0.08	0.55	0.07
		MC	28	0.74	0.12	0.40	0.06
	6	ALL	34	0.70	0.16	0.41	0.11
		CR	6	0.45	0.07	0.60	0.05
		MC	28	0.75	0.13	0.37	0.08
	7	ALL	34	0.70	0.13	0.40	0.13
		CR	6	0.51	0.06	0.63	0.04
		MC	28	0.74	0.10	0.35	0.07
	8	ALL	34	0.68	0.14	0.40	0.13
		CR	6	0.47	0.06	0.64	0.05
		MC	28	0.73	0.10	0.35	0.06
	10	ALL	34	0.72	0.12	0.44	0.13
		CR	6	0.53	0.03	0.67	0.02
		MC	28	0.76	0.09	0.39	0.08
Writing	5	ALL	14	0.72	0.16	0.40	0.14
		CR	3	0.56	0.03	0.61	0.03
		MC	10	0.79	0.10	0.32	0.04
		WP	1	0.40		0.60	
	8	ALL	14	0.66	0.12	0.40	0.17
		CR	3	0.61	0.04	0.65	0.01
		MC	10	0.69	0.13	0.30	0.04
		WP	1	0.56		0.66	
	11	ER (ALL)	1	0.51			

A comparison of indices across grade levels is complicated because these indices are population dependent. Direct comparisons would require that either the items or students were common across groups. Since that is not the case, it cannot be determined whether differences in performance across grade levels are because of differences in student abilities, differences in item difficulties, or both. With this caveat in mind, it appears generally that, for mathematics, students in higher grade levels found their items more difficult than students in lower grades found their items while, for reading, difficulty indices were more consistent across grades.

Comparing the difficulty indices of multiple-choice items and open-response (short-answer or constructed-response) items is inappropriate because multiple-choice items can be answered correctly by guessing. Thus, it is not surprising that the difficulty indices for multiple-choice items tend to be higher (indicating that students performed better on these items) than the difficulty indices for open-response items. Similarly, discrimination indices for the open-response items were larger than those for the dichotomous items because of the greater variability of the former (i.e., the partial credit these items allow) and the tendency for correlation coefficients to be higher given greater variances of the correlates.

In addition to the item difficulty and discrimination summaries presented above, item-level classical statistics and item-level score point distributions were also calculated. Item-level classical statistics are provided in Appendix E; item difficulty and discrimination values are presented for each item. The item difficulty and discrimination indices are within generally 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 low discrimination indices, but none was negative. While it is not inappropriate to include items with low discrimination values or with very high or very low item difficulty values to ensure that content is appropriately covered, there were very few such cases on NECAP. Item-level score point distributions are provided for open-response items in Appendix F; for each item, the percentage of students who received each score point is presented.

5.2 Differential Item Functioning

Code of Fair Testing Practices in Education (2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit and that actions should be taken to ensure that differences in performance are because of construct-relevant, rather than construct-irrelevant, factors. The *Standards for Educational and Psychological Testing* (AERA et al., 1999) includes similar guidelines. As part of the effort to identify such problems, an evaluation of the NECAP items was conducted in terms of Differential Item Functioning (DIF) statistics.

For NECAP, the standardization DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences. The standardization DIF procedure 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. In order to calculate DIF statistics, a minimum of 200 students must be in each comparison group.

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 reconsidered.

Computed DIF indices have a theoretical range from –1.0 to 1.0 for multiple-choice items, and the index is adjusted to the same scale for open-response items. Dorans and Holland (1993) suggested that index values between –0.05 and 0.05 should be considered negligible. The preponderance of NECAP 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 to 0.10 range (i.e., "high" DIF) are more unusual and should be examined very carefully.[2]

For the 2012–13 NECAP tests, seven subgroup comparisons were evaluated for DIF:

- male versus female
- no disability versus disability
- non-economically disadvantaged versus economically disadvantaged
- non-LEP versus LEP
- White versus Asian
- White versus Black
- White versus Hispanic

The tables in Appendix G present the numbers of items classified, overall and by group favored, as either "low" or "high" DIF.

5.3 DIMENSIONALITY ANALYSIS

Because tests are constructed with multiple content area subcategories, and their associated knowledge and skills, the potential exists for a large number of dimensions being invoked beyond the common primary dimension. Generally, the subcategories are highly correlated with each other; therefore, the

[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 of Education to determine whether to include the flagged item in a future operational test administration.

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 Item Response Theory (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 analyses performed on the 2012–13 NECAP common items for mathematics, reading, and writing are reported in Table 5-2. (Note: Only common items were analyzed since they are used for score reporting, and grade 11 writing was not analyzed because it consisted of a single assessment task.)

The dimensionality 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, and local *dependence* implies 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 displays 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 independent 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 0.2 indicate very weak multidimensionality (or near unidimensionality), values of 0.2 to 0.4 weak to moderate multidimensionality; values of 0.4 to 1.0 moderate to strong multidimensionality, and values greater than 1.0 very strong multidimensionality.

DIMTEST and DETECT were applied to the 2012–13 NECAP tests. The data for each grade and content area were split into a training sample and a cross-validation sample. Every grade/content area combination had at least 32,000 student examinees. Because DIMTEST was limited to using 24,000 students, the training and cross-validation samples for the DIMTEST analyses used 12,000 each, randomly sampled from the total sample. DETECT, on the other hand, had an upper limit of 50,000 students, so every training sample and cross-validation sample used with DETECT had at least 15,738 students. DIMTEST was then applied to every grade/content area. DETECT was applied to each dataset for which the DIMTEST null hypothesis was rejected in order to estimate the effect size of the multidimensionality.

The results of the DIMTEST analyses indicated that the null hypothesis was rejected at a significance level of 0.01 for every dataset. Because strict unidimensionality is an idealization that almost never holds exactly for a given dataset, these DIMTEST results were not surprising. Indeed, because of the very large sample sizes of NECAP, DIMTEST would be expected to be sensitive to even quite small violations of unidimensionality. Thus, it was important to use DETECT to estimate the effect size of the violations of local independence found by DIMTEST. Table 5-2 below displays the multidimensional effect size estimates from DETECT.

Table 5-2. 2012–13 NECAP: Multidimensionality Effect Sizes by Grade and Content Area

Content Area	*Grade*	*Multidimensionality Effect Size*	
		2012–13	*2011–12*
Mathematics	3	0.16	0.12
	4	0.12	0.12
	5	0.17	0.15
	6	0.18	0.15
	7	0.14	0.14
	8	0.13	0.16
	11	0.12	0.11
	Average	0.14	0.14
Reading	3	0.14	0.15
	4	0.17	0.20
	5	0.23	0.22
	6	0.20	0.21
	7	0.26	0.34
	8	0.33	0.30
	11	0.30	0.37
	Average	0.23	0.26
Writing	5	0.23	0.25
	8	0.23	0.20
	Average	0.23	0.23

All of the DETECT values for mathematics and writing indicated very weak to weak multidimensionality. For reading, grades 3 and 4 had very weak to weak multidimensionality, whereas grades 8 and 11 had DETECT values indicative of moderate multidimensionality. The average DETECT values for all three content areas indicated weak (reading, writing) or very weak (mathematics) multidimensionality. Also shown in Table 5-2 are the values reported in last year's dimensionality analyses. The individual values for the different grade levels as well as the averages for the three content areas are seen to be very similar to those from last year, with the exception of grade 7 and 11 reading, which had decreases of 0.08 and 0.07 (0.34 to 0.26 and 0.37 to 0.30) and, thus, went from moderate to weak multidimensionality. Looking back over the analyses of the past five years, the 2011–12 results represented peak values for multidimensional effect sizes, and the 2012–13 results are more similar to the average values across years and for other grades and subjects.

The way in which DETECT divided the tests into clusters was also investigated to determine whether there were any discernible patterns with respect to the multiple-choice and constructed-response item types. Inspection of the DETECT clusters indicated that multiple-choice/constructed-response separation occurred much more strongly with reading and writing than with mathematics, a pattern that has been consistent across all five years of dimensionality analyses for the NECAP fall tests. Specifically, for mathematics, grades 7, 8, and 11 showed some evidence of multiple-choice/constructed-response separation in that they each had at least one cluster dominated by multiple-choice points, and at least one cluster dominated by constructed-response points, as well as other clusters evenly mixed between multiple-choice and constructed-response points. Typically for these three mathematics tests, 60% or more of the points were accounted for by the multiple-choice and constructed-response dominated clusters. The remaining mathematics tests displayed no substantial separation of multiple-choice and constructed-response items in their DETECT solutions. In reading, on the other hand, grades 4–8, and 11 all displayed strong multiple-choice/constructed-response separation with one cluster containing all the multiple-choice items with no constructed-response and the remaining cluster(s) containing all the constructed-response items. Indeed, for reading grades 8 and 11 (the two with moderate multidimensionality), DETECT indicating a two-cluster solution in every case where one cluster was all multiple-choice and the other was almost all constructed-response. In reading, only grade 3 displayed no evidence of any multiple-choice/constructed-response separation; this finding was also reported for reading, grade 3 for the 2011–12 and 2012–13 NECAP tests. For writing, both grades displayed a strong multiple-choice/constructed-response separation in the same manner as occurred with reading, with one cluster having all the multiple-choice items and the remaining clusters having all the constructed-response items.

Thus, a tendency is suggested for multiple-choice and constructed-response items to sometimes measure statistically separable dimensions, especially in regard to the reading and writing tests. This has been consistent across all five years of analyses of the NECAP fall test administrations. However, it is important to emphasize that the degree of violation of unidimensional local independence has not been large in any of the three content areas over the four years of analysis. The degree to which these violations of local independence

can be attributed to item type differences tends to be greater for reading and writing than for mathematics. More investigation by content area experts would be required to better understand the violations of local independence that are due to sources other than item type.

In summary, for the 2012–13 analyses the violations of local independence, as evidenced by the DETECT effect sizes, were weak or very weak in all but three cases (reading in grades 8 and 11), both of which had effect sizes near the 0.30 border between weak and moderate. Thus, these effects do not seem to warrant any changes in test design or scoring, which are designed to be consistent across a range of grade levels. In addition, the magnitude of the violations of local independence have been consistently low over the years, and the patterns with respect to the multiple-choice and constructed-response items have also been consistent, with reading and writing tending to display more separation than mathematics.

CHAPTER 6 ITEM RESPONSE THEORY SCALING AND EQUATING

This chapter describes the procedures used to calibrate, equate, and scale the NECAP tests. During the course of these psychometric analyses, a number of quality-control procedures and checks on the processes were implemented. These procedures included evaluations of the calibration processes (e.g., checking the number of Newton cycles required for convergence for reasonableness, checking item parameters and their standard errors for reasonableness, examination of Test Characteristic Curves [TCCs] and Test Information Functions [TIFs] for reasonableness); evaluation of model fit; evaluation of equating items (e.g., delta analyses, rescore analyses, examination of α-plots and b-plots for reasonableness); and evaluation of the scaling results (e.g., parallel processing by the Psychometrics and Research department and Data Analysis department; comparing look-up tables to the previous year's). An equating report, which provided complete documentation of the quality-control procedures and results, was submitted to the member Departments of Education for their approval prior to production of student reports.

Table 6-1 lists items that required intervention either during item calibration or as a result of the evaluations of the equating items. For each flagged item, the table shows the reason it was flagged and what action was taken. The number of items identified for evaluation was very typical across the grades. Descriptions of the evaluations and results are included in Section 6.2, *Item Response Theory Results*, and Section 6.4, *Equating Results*.

Table 6-1. 2012–13 NECAP: Items That Required Intervention During IRT Calibration and Equating

Content Area	*Grade*	*Item Number*	*Reason*	*Action*
Mathematics	3	119752	c-parameter	fix c = 0
		124433	c-parameter	fix c = 0
		168364	c-parameter	fix c = 0
		168404	delta analysis	removed from equating
	4	120224	b/b analysis	removed from equating
		120274	b/b analysis	removed from equating
		124620	c-parameter	fix c = 0
		144645	delta analysis	removed from equating
		144648	c-parameter	fix c = 0
		255694	c-parameter	fix c = 0
	5	120654	c-parameter	fix c = 0
		150816	b/b analysis	removed from equating
		203997	delta analysis	removed from equating
	6	119232	c-parameter	fix c = 0
		119288	b/b analysis	removed from equating
		123513	b/b analysis	removed from equating
		123513	delta analysis	removed from equating
		125025	c-parameter	fix c = 0

continued

Content Area	Grade	Item Number	Reason	Action
Mathematics	7	120381	b/b analysis	removed from equating
		125286	c-parameter	fix c = 0
		199932	delta analysis	removed from equating
	8	120884	b/b analysis	removed from equating
		139845	c-parameter	fix c = 0
		204990	delta analysis	removed from equating
		205045	c-parameter	fix c = 0
	11	119544	delta analysis	removed from equating
		145469	c-parameter	fix c = 0
		169704	b/b analysis	removed from equating
Reading	3	117787	c-parameter	fix c = 0
		117790	c-parameter	fix c = 0
		117794	c-parameter	fix c = 0
		117807	b/b analysis	removed from equating
		117807	delta analysis	removed from equating
		148224	c-parameter	fix c = 0
		171525	c-parameter	fix c = 0
		171719	c-parameter	fix c = 0
		208039	c-parameter	fix c = 0
		208067	c-parameter	fix c = 0
	4	171904	c-parameter	fix c = 0
		208065	c-parameter	fix c = 0
		208802	delta analysis	removed from equating
	5	149079	b/b analysis	removed from equating
		149079	c-parameter	fix c = 0
		171822	c-parameter	fix c = 0
		171825	c-parameter	fix c = 0
		209368	c-parameter	fix c = 0
		209539	c-parameter	fix c = 0
		209546	c-parameter	fix c = 0
		209551	c-parameter	fix c = 0
		209593	c-parameter	fix c = 0
		209616	c-parameter	fix c = 0
		209617	c-parameter	fix c = 0
		209631	c-parameter	fix c = 0
		209700	c-parameter	fix c = 0
	6	172094	b/b analysis	removed from equating
		172111	c-parameter	fix c = 0
		172134	c-parameter	fix c = 0
		208944	c-parameter	fix c = 0
		208972	c-parameter	fix c = 0
		208974	c-parameter	fix c = 0
		209137	c-parameter	fix c = 0
		209142	c-parameter	fix c = 0
		209143	c-parameter	fix c = 0
		209267	c-parameter	fix c = 0
		209284	c-parameter	fix c = 0
	7	147203	c-parameter	fix c = 0
		147210	c-parameter	fix c = 0
		147215	c-parameter	fix c = 0

continued

Content Area	Grade	Item Number	Reason	Action
Reading	7	147633	c-parameter	fix c = 0
		173260	c-parameter	fix c = 0
		173269	c-parameter	fix c = 0
		173304	c-parameter	fix c = 0
		173319	c-parameter	fix c = 0
		173338	c-parameter	fix c = 0
		207521	c-parameter	fix c = 0
		207604	c-parameter	fix c = 0
		207625	c-parameter	fix c = 0
		207630	c-parameter	fix c = 0
		207633	c-parameter	fix c = 0
		207636	c-parameter	fix c = 0
		207640	c-parameter	fix c = 0
		207643	c-parameter	fix c = 0
		207647	c-parameter	fix c = 0
	8	118749	c-parameter	fix c = 0
		147516	c-parameter	fix c = 0
		172353	c-parameter	fix c = 0
		172356	c-parameter	fix c = 0
		172734	b/b analysis	removed from equating
		207586	c-parameter	fix c = 0
		207711	c-parameter	fix c = 0
		207715	c-parameter	fix c = 0
		207742	c-parameter	fix c = 0
		207746	c-parameter	fix c = 0
		207777	c-parameter	fix c = 0
		207840	c-parameter	fix c = 0
	11	147383	c-parameter	fix c = 0
		208312	b/b analysis	removed from equating
		208312	delta analysis	removed from equating
		208345	c-parameter	fix c = 0
		208426	c-parameter	fix c = 0
		208432	c-parameter	fix c = 0
		208482	c-parameter	fix c = 0
		208520	c-parameter	fix c = 0
		208524	c-parameter	fix c = 0
		208684	c-parameter	fix c = 0
		208689	c-parameter	fix c = 0
		208704	c-parameter	fix c = 0
		208729	c-parameter	fix c = 0
		208730	c-parameter	fix c = 0
		226071	c-parameter	fix c = 0

6.1 ITEM RESPONSE THEORY

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, all items are assumed to be 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 & van der Linden, 1997; Hambleton & 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 the 2012–13 NECAP tests, the three-parameter logistic (3PL) model was used for dichotomous (multiple-choice) items and the Graded-Response Model (GRM) was used for polytomous (open-response) items. The 3PL model for dichotomous items can be defined as:

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

where
i indexes the items,
j indexes students,
α represents item discrimination,
b represents item difficulty,
c is the pseudoguessing parameter,
ξ_i represents the set of item parameters (α, b, and c) for item i, and
D is a normalizing constant equal to 1.701.

In the GRM for polytomous items, an item is scored in a $k + 1$ graded category 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 a $k + 1$ category can be characterized by k Item Category Threshold Curves (ICTCs) of the two-parameter logistic form:

$$P^*_{ik}(1|\theta_j, \alpha_i, b_i, d_{ik}) = \frac{\exp[Da_i(\theta_j - b_i + d_{ik})]}{1+\exp[Da_i(\theta_j - b_i + d_{ik})]}$$

where
i indexes the items,
j indexes students,
k indexes threshold,
α 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 Item Characteristic Curve (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.

$$P_i\left(1\middle|\theta_j\right) = \sum_k^{m+1} w_{ik} P_{ik}\left(1\middle|\theta_j\right)$$

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 Results

The tables in Appendix H give the IRT item parameters of all common items on the 2012–13 NECAP tests by grade and content area. In addition, Appendix I shows graphs of the TCCs and 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

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

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), and

$E\left(X\middle|\theta_j\right)$ 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 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.

Table 6-1 above lists items that were flagged based on the quality-control checks implemented during the calibration process. (Note that some items were flagged as a result of the evaluations of the equating items; those results are described below.) In all cases, items flagged during this step were identified because of the pseudo-guessing parameter (*c* parameter) being poorly estimated. Difficulty in estimating the *c* parameter is not at all unusual and is well-documented in psychometric literature (see, for example, Nering & Ostini, 2010), especially when the item's discrimination is below 0.50. In all cases, fixing the *c* parameter resulted in reasonable and stable item parameter estimates and improved model fit.

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 fell within acceptable ranges.

Table 6-2. 2012–13 NECAP: Number of Newton Cycles Required for Convergence

Content Area	*Grade*	*Cycles*
Mathematics	3	37
	4	132
	5	58
	6	47
	7	141
	8	51
	11	82
Reading	3	97
	4	65
	5	57
	6	46
	7	47
	8	47
	11	45

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 2012–13 administration of the NECAP tests used a raw score-to-theta equating procedure in which test forms were equated to the theta scale established on the reference form (i.e., the form used in the most recent standard setting). 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.

The groups of students who took the equating items on the 2012–13 NECAP reading tests are not equivalent to the groups who took them in the reference years. IRT is particularly useful for equating scenarios that involve nonequivalent groups (Allen & 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 by utilizing a set of anchor items (also called equating items). 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 distributed across forms.

Item parameter estimates for the 2012–13 NECAP tests were placed on the 2011–12 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 2011–12 and 2012–13 NECAP tests should have the same item parameters. After the item parameters for each 2012–13 test were estimated using PARSCALE (Muraki & Bock, 2003), the Stocking and Lord method was employed to find the linear transformation (slope and intercept) that adjusted the equating items' parameter estimates such that the 2012–13 NECAP tests TCC for the equating items was as close as possible to that of the 2011–12 NECAP tests.

6.4 EQUATING RESULTS

Prior to calculating the Stocking and Lord transformation constants, a variety of evaluations of the equating items were conducted. Items that were flagged as a result of these evaluations are listed in Table 6-1 at the beginning of this chapter. These items were scrutinized and a decision was made as to whether to include the item as an equating item or to discard it. The procedures used to evaluate the equating items are described below.

Appendix J presents the results from the delta analysis. This procedure was used to evaluate adequacy of equating items; the discard status presented in the appendix indicates whether the item was flagged as potentially inappropriate for use in equating.

Also presented in Appendix J 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 one year to the next. All effect sizes were well below the criterion value for excluding an item as an equating item, 0.50 in absolute value.

Finally, α-plots and b-plots, which show IRT parameters for the 2012–13 NECAP tests plotted against the values of the 2011–12 NECAP tests, are presented in Appendix K. Any items that appeared as outliers in the plots were evaluated in terms of suitability for use as equating items.

Once all flagged items had been evaluated and appropriate action taken, the Stocking and Lord method of equating was used to place the item parameters onto the previous year's scale, as described above. The Stocking and Lord Transformation Constants are presented in Table 6-3.

Table 6-3. 2012–13 NECAP: Stocking and Lord Transformation Constants

Content Area	*Grade*	α-*slope*	b-*intercept*
Mathematics	3	1.008	0.097
	4	1.108	0.081
	5	1.028	0.147
	6	1.127	0.19
	7	1.06	0.244
	8	0.957	0.246
	11	1.028	0.218
Reading	3	1.015	0.035
	4	1.046	0.223
	5	0.966	0.165
	6	1.064	0.096
	7	1.144	0.13
	8	1.111	0.32
	11	1.086	0.39

The next administration of the NECAP test (in 2013–14) will be scaled to the 2012–13 NECAP administration using the same equating method described above.

6.5 ACHIEVEMENT STANDARDS

NECAP standards to establish achievement level cut scores in reading and mathematics for grades 3 through 8 were set in January, 2006, and in reading, mathematics, and writing for grade 11 in January, 2008. Details of the standard setting procedures can be found in the respective standard setting reports, as well as in the technical reports of those years.

Achievement standards for writing grades 5 and 8 were set in December, 2011; for complete details of the standard setting, please see the *2011–12 New England Common Assessment Program Standard Setting Report* (Measured Progress, 2011).

The cuts on the theta scale that were established via standard setting and used for reporting in fall 2012 are presented in Table 6-4 below. Also shown in the table are the cutpoints on the reporting score scale (described below). These cutpoints will remain fixed throughout the assessment program unless standards are reset for any reason.

Table 6-4. 2012–13 NECAP: Cutpoints on the Theta Metric and Reporting Scale by Content Area and Grade

Content Area	Grade	Theta			Scaled Score				
		Cut 1	Cut 2	Cut 3	Minimum	Cut 1	Cut 2	Cut 3	Maximum
Mathematics	3	-1.038	-0.269	0.970	300	332	340	353	380
	4	-1.150	-0.378	0.949	400	431	440	455	480
	5	-0.928	-0.285	1.031	500	533	540	554	580
	6	-0.874	-0.224	1.034	600	633	640	653	680
	7	-0.708	-0.079	1.099	700	734	740	752	780
	8	-0.644	-0.029	1.118	800	834	840	852	880
	11	-0.117	0.619	2.059	1100	1134	1140	1152	1180
Reading	3	-1.323	-0.497	1.031	300	331	340	357	380
	4	-1.173	-0.314	1.147	400	431	440	456	480
	5	-1.335	-0.428	1.040	500	530	540	556	580
	6	-1.478	-0.518	1.125	600	629	640	659	680
	7	-1.483	-0.522	1.206	700	729	740	760	780
	8	-1.525	-0.522	1.134	800	828	840	859	880
	11	-1.207	-0.310	1.004	1100	1130	1140	1154	1180
Writing	5	-1.283	-0.009	1.524	500	527	540	555	580
	8	-1.349	-0.106	1.268	800	827	840	854	880

The tables in Appendix L show achievement level distributions by content area and grade. Results are shown for each of the last three years for all grades of reading and mathematics and for writing grade 11. For writing grades 5 and 8, because standards were set in December 2011, results are shown only for the 2011–12 and 2012–13 NECAP administrations.

6.6 Reported Scaled Scores

Because the θ scale used in IRT calibrations is not readily understood by most stakeholders, reporting scales were developed for NECAP. The reporting scales are simple linear transformations of the underlying θ scale. The reporting scales are developed such that they range 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 is fixed 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 scores. 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 2012–13 NECAP tests were translated to scaled scores using a data analysis process called *scaling*. Scaling simply converts from one scale to another scale. 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 2012–13 NECAP tests can be expressed in raw or scaled scores.

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 grades and content areas. The raw cut score between Partially Proficient and Proficient could be, say, 35 in mathematics and 33 in reading, yet both of these raw scores would be transformed to scaled scores of *x*40. It is this uniformity across scaled 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.

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 and content area combination. For NECAP, the transformation function is determined by fixing the Partially Proficient/Proficient cut score and the bottom of the scale—that is, the *x*40 and the *x*00 values (e.g., 440 and 400 for grade 4). The *x*00 location on the θ scale is beyond (i.e., below) the scaling of all items. To determine this location, a chance score (approximately equal to a student's expected performance by guessing) is mapped to a value of –4.0 on the θ scale. A raw score of 0 is also assigned a scaled score of *x*00. The maximum possible raw score is assigned a scaled score of *x*80 (e.g., 480 in the case of grade 4). Because only two points within the θ scaled score space are fixed, the scaled score cutpoints between Substantially Below Proficient and Partially Proficient and between Proficient and Proficient with Distinction are free to vary across the grade and content area combinations.

Table 6-5 shows the slope and intercepts terms used to calculate the scaled scores for each content area and grade. Note that the values in Table 6-5 will not change unless the standards are reset.

Table 6-5. 2012–13 NECAP: Scaled Score Slope and Intercept by Content Area and Grade

Content Area	*Grade*	*Slope*	*Intercept*
Mathematics	3	10.720	342.878
	4	11.043	444.173
	5	10.766	543.063
	6	10.592	642.369
	7	10.201	740.803
	8	10.072	840.288
	11	8.660	1134.640
Reading	3	11.419	345.675
	4	10.852	443.410
	5	11.197	544.788
	6	11.487	645.950
	7	11.502	746.007
	8	11.502	846.009
	11	10.840	1143.359
Writing	5	10.022	540.087
	8	10.272	841.088

Appendix M contains raw score to scaled score look-up tables for the 2012–13 NECAP tests. These are the actual tables used to determine student scaled scores, error bands, and achievement levels.

Appendix N contains scaled score distribution graphs for each grade and content area. These distributions were calculated using the sparse data matrix files that were used in the IRT calibrations.

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 misread an item, or mistakenly fill in the wrong bubble when he or she knew the answer. Collectively, extraneous factors that affect 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, 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 two 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 with 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 into 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, α (alpha), 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 2012–13 NECAP tests:

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

where
i indexes the item,
n is the total number of items,
$\sigma^2_{(Y_i)}$ represents individual item variance, and
σ^2_x represents the total test variance.

7.1 Reliability and Standard Errors of Measurement

Table 7-1 presents descriptive statistics, Cronbach's α coefficient, and the raw score Standard Errors of Measurements (SEMs) for each grade and content area. (Statistics are based on common items only.) Note that reliability could not be calculated for grade 11 writing because the test consists of a single writing prompt.

Table 7-1. 2012–13 NECAP: Raw Score Descriptive Statistics, Cronbach's Alpha, and Standard Errors of Measurement (SEM) by Content Area and Grade

Content Area	*Grade*	*Number of Students*	*Raw Score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard Deviation*		
Mathematics	3	43,978	65	40.50	12.71	0.93	3.39
	4	43,558	65	41.15	12.72	0.93	3.47
	5	44,020	66	37.55	14.67	0.93	3.97
	6	44,582	66	35.80	14.90	0.93	4.02
	7	44,518	66	32.52	14.68	0.93	3.85
	8	44,586	66	35.46	14.69	0.93	3.95
	11	31,249	64	25.99	14.69	0.94	3.58
Reading	3	43,834	52	32.80	8.74	0.88	3.04
	4	43,433	52	34.62	8.99	0.89	3.05
	5	43,892	52	29.91	8.63	0.89	2.89
	6	44,482	52	31.77	8.25	0.88	2.82
	7	44,427	52	33.08	9.13	0.88	3.19
	8	44,520	52	31.60	9.01	0.88	3.09
	11	31,234	52	33.98	9.38	0.90	2.96
Writing	5	43,832	34	19.42	4.85	0.76	2.40
	8	44,366	34	20.99	5.35	0.76	2.61
	11	31,183	12	6.18	2.02		

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.

7.2 SUBGROUP RELIABILITY

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2012–13 NECAP test. Appendix O presents reliabilities for various subgroups of interest. Subgroup Cronbach's α's were calculated using the formula defined above based only on the members of the subgroup in question in the computations; values are only calculated for subgroups with 10 or more students.

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 O 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 & 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 REPORTING SUBCATEGORY RELIABILITY

Of even more interest are reliabilities for the reporting subcategories within NECAP content areas, described in Chapter 2. Cronbach's α coefficients for subcategories were calculated via the same formula defined previously using just the items of a given subcategory in the computations. Results are presented in Appendix O. Once again as expected, because results are based on a subset of items rather than the full test, computed subcategory reliabilities were lower (sometimes substantially so) than were overall test reliabilities, and interpretations should take this into account. The subcategory reliabilities were lower than those based on the total test and approximately to the degree one would expect based on classical test theory. Qualitative differences between grades and content areas once again preclude valid inferences about the quality of the full test based on statistical comparisons among subcategories.

7.4 INTERRATER CONSISTENCY

Chapter 4 of this report describes in detail the processes that were implemented to monitor the quality of the hand-scoring of student responses for constructed-response items. One of these processes was double-blind scoring: approximately 2% of student responses were randomly selected and scored independently by two different scorers, except for the writing prompts in grades 5, 8, and 11, which were 100% double scored.

Results of the double-blind scoring were used during the scoring process to identify scorers who required retraining or other intervention and are presented here as evidence of the reliability of the NECAP tests. A summary of the interrater consistency results are presented in Table 7-2 below. Results in the table are collapsed across the hand-scored items by grade and content area. The table shows the number of score categories, number of included scores, percent exact agreement, percent adjacent agreement, correlation between the first two sets of scores, and percentage of responses that required a third score. This same information is provided at the item level in Appendix P.

Table 7-2. 2012–13 NECAP: Summary of Interrater Consistency Statistics Collapsed Across Items by Content Area and Grade

Content Area	*Grade*	*Number of*		*Percent*		*Correlation*	*Percent of Third Scores*
		Score Categories	*Included Scores*	*Exact*	*Adjacent*		
Mathematics	3	2	8,686	98.88	1.12	0.98	0.00
		3	8,741	95.39	4.50	0.96	0.11
	4	2	8,695	98.84	1.16	0.98	0.00
		3	8,729	92.93	6.69	0.93	0.38
	5	2	5,300	98.00	2.00	0.96	0.00
		3	5,227	92.12	7.50	0.94	0.38
		5	3,551	88.57	9.88	0.95	1.55
	6	2	5,365	98.04	1.96	0.96	0.00
		3	5,322	92.48	7.23	0.94	0.28
		5	3,733	81.03	16.56	0.94	2.38
	7	2	5,272	98.08	1.92	0.96	0.00
		3	5,225	94.07	5.61	0.95	0.33
		5	3,583	87.44	11.36	0.95	1.17
	8	2	5,336	96.63	3.37	0.93	0.00
		3	5,301	95.13	4.66	0.96	0.21
		5	3,698	83.15	15.49	0.95	1.33
	11	2	7,196	98.71	1.29	0.97	0.00
		3	3,459	93.29	6.01	0.94	0.69
		5	2,398	88.53	9.76	0.94	1.71
Reading	3	5	5,415	82.27	16.69	0.94	0.98
	4	5	5,327	79.59	19.07	0.92	1.33
	5	5	5,263	66.73	31.79	0.81	1.39
	6	5	5,456	67.50	31.36	0.79	1.08
	7	5	5,725	63.88	33.82	0.82	2.31
	8	5	5,462	60.14	37.17	0.77	2.44
	11	5	3,754	59.54	37.61	0.75	2.64
Writing	5	5	2,740	67.59	30.84	0.76	1.57
		7	43,119	59.25	36.97	0.68	3.72
	8	5	2,669	64.71	34.10	0.75	1.16
		7	43,340	57.06	40.02	0.74	2.73
	11	7	30,248	58.19	38.92	0.77	2.68

7.5 Reliability of Achievement Level Categorization

While related to reliability, the accuracy and consistency of classifying students into achievement categories are even more important statistics in a standards-based reporting framework (Livingston & Lewis, 1995). After the achievement levels were specified and students were classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications. For NECAP, students are classified into one of four achievement levels: Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction. This section of the report explains the methodologies used to assess the reliability of classification decisions, and results are given.

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 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 (1995) technique was used for the 2012–13 NECAP tests because it is easily adaptable to all types of testing formats, including mixed format tests.

The accuracy and consistency estimates reported in Appendix Q make use of "true scores" in the classical test theory 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 (1995) method, estimated true scores are used to categorize students into their "true" classifications.

For the 2012–13 NECAP tests, after various technical adjustments (described in Livingston & Lewis, 1995), a four-by-four contingency table of accuracy was created for each grade and content area, where cell $[i, j]$ represented the estimated proportion of students whose true score fell into classification i (where $i = 1$ to 4) and observed score into classification j (where $j = 1$ to 4). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate 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 four-by-four contingency table was created for each grade and content area and populated by the proportion of students who would be categorized into each combination of classifications 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 classification i (where $i = 1$ to 4) and whose observed score on the second form would fall into classification j (where $j = 1$ to 4). The sum of the diagonal entries

(i.e., the proportion of students categorized by the two forms into exactly the same classification) signified overall consistency.

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:

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

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.

7.5.1 Accuracy and Consistency Results

The accuracy and consistency analyses described above are provided in Table Q-1 of Appendix Q. The table includes overall accuracy and consistency indices, including kappa. Accuracy and consistency values conditional upon achievement level are also given. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.84 for Substantially Below Proficient for grade 3 mathematics. This figure indicates that among the students whose true scores placed them in this classification, 84% would be expected to be in this classification when categorized according to their observed scores. Similarly, a consistency value of 0.77 indicates that 77% of students with observed scores in the Substantially Below Proficient level would be expected to score in this classification again if a second, parallel test form were used.

For some testing situations, the greatest concern may be decisions around level thresholds. For example, in testing done for No Child Left Behind Act (NCLB) accountability purposes, the primary concern is distinguishing between students who are proficient and those who are not yet proficient. In this case, the accuracy of the Partially Proficient-Proficient threshold is of greatest interest. For the 2012–13 NECAP tests, Table Q-2 in Appendix Q provides accuracy and consistency estimates at each cutpoint as well as false positive and false negative decision rates. (A false positive is the proportion of students whose observed scores were above the cutpoint and whose true scores were below the cutpoint. A false negative is the proportion of students whose observed scores were below the cutpoint and whose true scores were above the cutpoint.)

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 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 symmetrical, indicating that the two parallel forms have the same statistical properties. This second reason is consistent with the notion of forms that are parallel; that is, it is more intuitive and interpretable for two parallel forms to have the same statistical distribution.

Note that, as with other methods of evaluating reliability, accuracy, and consistency statistics calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Appendix Q should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare accuracy and consistency statistics between grades and content areas.

CHAPTER 8 SCORE REPORTING

8.1 TEACHING YEARS VERSUS TESTING YEAR REPORTING

The data used for the NECAP reports are the results of the fall 2012 NECAP test administration. It is important to note that the NECAP tests are based on the Grade Level Expectations (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.

8.2 PRIMARY REPORTING DELIVERABLES

The following reporting deliverables were produced for the 2012–13 NECAP tests:

- Student Report
- School and District Results Report
- School and District Summary Report
- School and District Student-Level Data File
- Analysis & Reporting System

With the exception of the Student Report, these reports and data files were available for schools and districts to view or download via the NECAP Analysis & Reporting System, a password-secure Web site hosted by Measured Progress. Each of these reporting deliverables is described in the following sections. Sample reports are provided in Appendix R.

Support is provided by the state Departments of Education and Measured Progress to stakeholders who use the various reporting deliverables by hosting report interpretation workshops and by providing the *Guide to Using the 2012 NECAP Reports*. These resources help foster proper use and interpretation of NECAP results.

The *Guide to Using the 2012 NECAP Reports* includes a table that shows the number of scaled score points that would indicate a statistically significant difference between two equally sized groups of students.

The calculations are performed by computing the standard error of the difference in means ($\sigma_{\overline{x}_1-\overline{x}_2}$) for different values of n, based on the observed scaled score standard deviations for each grade and content area. The formula for the variance error of the difference in means is:

$$\sigma^2_{\overline{x}_1-x_2} = \sigma_2^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$$

By assuming $n_1 = n_2 = n$ and $\sigma_1^2 = \sigma_2^2 = \sigma_w^2$, this equation simplifies to:

$$\sigma^2_{\overline{x}_1-x_2} = \sigma_w^2\left(\frac{1}{n}+\frac{1}{n}\right) = \frac{2\sigma^2}{n}$$

and

$$\sigma_{\overline{x}_1-x_2} = \sqrt{2\sigma^2/n}$$

Therefore, if the difference in scaled scores of two equally sized groups is greater than or equal to $\sigma_{\overline{x}_1-x_2}$, you can be 67% certain that there is a true difference in performance between the two groups. Differences between two unequally sized groups can be interpreted, conservatively, by using the value associated with the size of the smaller group.

The *Guide to Using the 2012 NECAP Reports* also includes a second table that shows corresponding values based on percentages of students, to help interpret differences in percentages of students in performance level categories. The calculations for this table are based on the variance error of a proportion:

$$s_p^2 = \frac{S^2}{n} = \frac{p(1-p)}{n}$$

and

$$s_p = \sqrt{p(1-p)/n}$$

Together, these two tables in the *Guide to Using the 2012 NECAP Reports* help teachers, schools, and districts interpret differences in scores between two groups of students and, in this way, support appropriate interpretation of NECAP scores.

8.3 STUDENT REPORTS

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 report and the achievement level descriptions. The reverse side of the *NECAP Student Report* provides a complete picture of

an individual student's performance on the NECAP test, 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. Grade 5 and 8 writing report Multiple Choice, Short Responses, and Extended Response as categories. Grade 11 writing only reports Extended Response as a category.

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

8.4 Item Analysis Reports

The *NECAP Item Analysis Report* provides a roster of all students in a school and provides their performance 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, 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 Depth of Knowledge (DOK) code for the item, the item type, the correct response key for multiple-choice items, and the total possible points.

For each student, multiple-choice items are marked either with a plus sign (+), indicating that the student chose the correct multiple-choice response, or a letter (from A to D), indicating the incorrect response chosen by the student. For short-answer and constructed-response 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. Subcategory Points Earned columns show points earned by the student in each content area subcategory 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 percent correct for each multiple-choice item and average scores for the short-answer and constructed-response items 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 strand, 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. FERPA requires that access to individual student results be restricted to the student, the student's parents/guardians, and authorized school personnel.

8.5 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, the results for the content areas, and the disaggregated content area results.

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, 2012. The total number of students enrolled is defined as the number of students tested plus the number of students not tested.

Data are provided for the following groups of students who are considered tested in NECAP:

- **Students Tested**: This category provides the total number of students tested.
- **Students Tested with an Approved Accommodation**: Students in this category tested with an accommodation and did not have their test invalidated.
- **Current LEP Students**: Students in this category are currently receiving Limited English Proficiency (LEP) services.
- **Current LEP Student Tested with an Approved Accommodation**: Students in this category are currently receiving LEP services, tested with an accommodation, and did not have their test invalidated.
- **IEP Students**: Students in this category have an Individual Education Plan (IEP).
- **IEP Student tested with an Approved Accommodation**: Students in this category have an IEP, tested with an accommodation, and did not have their test invalidated.

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.

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 2011–12 school year.
- **First-Year LEP**: Students in this category are defined as being new to the United States after October 1, 2010, 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 tests.
- **Withdrew after October 1**: Students withdrawing from a school after October 1, 2012, 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, 2012, 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 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.

For reading and mathematics, 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). For writing in grades 5 and 8, information is provided by item type (multiple-choice, short-response, and extended-response). The purpose of these sections is to help schools determine the extent to which their curricula are effective in helping students to achieve the particular standards and benchmarks contained in the GLEs and GSEs. The content area results pages provide data for the 2010–11, 2011–12, and 2012–13 NECAP individual test administrations as well as cumulative data for the three years in reading and mathematics. For writing grades 5 and 8, data are only provided for the 2011–12 and the 2012–13 NECAP test administrations as well as cumulative data for the two years. Data do not exist for the 2010–11 NECAP test administration for writing in grades 5 and 8 because the test was a pilot and results were not produced.

Information about each content area (reading and mathematics for all grades and writing for grades 5 and 8) for school, district, and state levels include:

- 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 and mathematics in all grades and about each item type for writing in grades 5 and 8 includes the following:

- The total possible points for that reporting category or item type. 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 percentage of total possible points for each sub-content area/reporting category for the school, district, and state. In this graphic display, there are symbols representing school, district, and state performances. 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).

The first content area results page for the grade 11 writing test provides data for the 2010–11, 2011–12, and 2012–13 NECAP individual test administrations as well as cumulative data for the three years. Information provided for the 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 raw score.

The bottom half of this page includes a table that lists the type of writing (genre) for the common prompt (i.e., the prompt for which the results on the top half of the page are about) for each of the last three test administrations. The type of writing and a description of that type are included for each year.

The second content area results page lists the types of writing that are assessed in the grade 11 writing test. The list is comprised of the types of writing used to make both a common prompt (one that is administered to all students) and matrix prompts (ones that vary across the eight different forms of the test). The first column on this page provides the name and a description of each type of writing. The second column

provides separate rows for the current year (2012–13) and the previous year that each type of writing was assessed. The symbol "(C)" indicates the type of writing that was common in that year's test. The number of students tested and the mean raw score are provided for the school, district, and state levels. A graphic display is also provided for each year and type of writing that shows the average score attained on the 0-to-12 scale for the school, district, and state levels. The scale of 0 to 12 on the graphic display represents the possible score range for the writing prompt. The two scores assigned to the student's response from the 6-point rubric are added; the result is a score from 0 to 12. A score of 7 depicted on the scale represents the score needed to be proficient.

Finally, the third page of the grade 11 writing content area results contains a table that presents information on the distribution of scores across the 0-to-12 range. The first column of the table lists the possible scores from 12 down to 0. The next two columns ("Score 1" and "Score 2") represent the two independent scores assigned to a student's response to the common writing prompt. The student's total score on the common writing prompt is the sum of these two scores. The next four columns list the total number of students, N, and the percentage, %, of students for each score on the 0-to-12 scale for the school and district. The last column provides the percentage, %, of students for each score on the 0-to-12 scale for the state. The 6-point scoring rubric that is used to score student responses to the common writing prompt is also included on this page of the report.

The disaggregated content area results pages (all grades and content areas) present the relationship between performance and student reporting categories (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 (or mean raw score for grade 11 writing) by each reporting category.

The list of student reporting categories is as follows:

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

The data for achievement levels and mean scaled score (or mean raw score for grade 11 writing) 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 no data were reported for the 504 Plan in any of the content areas for New Hampshire and Vermont. Additionally, no data were reported for Title I in any of the content areas for Vermont.

8.6 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 GLEs and GSEs.

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 reports and district reports were produced for each grade level tested.

8.7 School and District Student-Level Data Files

In addition to the reports described above, districts and, for the first time this year, schools received access to and were able to download student-level data files from the Analysis & Reporting System for each grade of students tested within their district or school. Student-level data files were produced for both "teaching year" and "testing year."

The student-level data files list students alphabetically within each school and contain all of the demographic information that was provided by the state for each student. Student records contain the scaled score, achievement level, and subscores earned by the student for each content area tested. In addition, the student records contain each student's actual performance on each of the released items for each content area tested as well as the student's responses to the student questionnaire.

The data collected from the optional reports field, if it was coded by schools on page 2 of the Student Answer Booklets, are also available for each student in the student-level data file. The optional reports field was provided to allow schools the option of grouping individual students into additional categories (e.g., by

class or by previous year's teacher). This allows schools to make comparisons between subgroups that are not already listed on the disaggregated results pages of the school and district results reports.

The file layout of the student-level data files that lists all of the field names, variable information, and valid values for each field was also available to districts and schools on the Analysis & Reporting System.

8.8 Analysis & Reporting System

NECAP results for the 2012–13 test administration were accessible online via the Analysis & Reporting System. In addition to accessing and downloading reports and student-level data files in the same manner as in previous years, this new system includes interactive capabilities that allow school and district users to sort and filter item and subgroup data to create custom reports.

8.8.1 Interactive Reports

There are four interactive reports that were available from the Analysis & Reporting System: "Item Analysis Report," "Achievement Level Summary," "Released Items Summary Data," and "Longitudinal Data." Each of these interactive reports is described in the following sections. To access these four interactive reports, the user needed to click the "Interactive" tab on the home page of the system and select the report desired from the drop-down menu. Next, the user had to apply basic filtering options such as the name of the district or school and the grade level/content area test to open the specific report. At this point, the user had the option of printing the report for the entire grade level or applying advanced filtering options to select a subgroup of students for which to analyze their results. Advanced filtering options included gender, ethnicity, LEP, IEP, and SES. Users also needed to select either the "Teaching" or "Testing" cohort of students using the "Filter by Group" drop-down menu. All interactive reports, with the exception of the Longitudinal Data Report, allowed the user to provide a custom title for the report.

8.8.1.1 Item Analysis Report

The Item Analysis Report provides individual student performance data on the released items and total test results for a selected grade and content area. Please note that when advanced filtering criteria are applied by the user, the School and District Percent Correct and Average Score rows at the bottom of the report are blanked out and only the Group row and the State row for the group selected will contain data. This report can be saved, printed, or exported as a PDF.

8.8.1.2 Achievement Level Summary

The Achievement Level Summary provides a visual display of the percentages of students in each achievement level for a selected grade and content area. The four achievement levels (Proficient with Distinction, Proficient, Partially Proficient, and Substantially Below Proficient) are represented by various

colors in a pie chart. A separate table is also included below the chart that shows the number and percentage of students in each achievement level. This report can be saved, printed, or exported as a PDF or JPG file.

8.8.1.3 Released Items Summary Data

The Released Items Summary Data report is a school-level report that provides a summary of student responses to the released items for a selected grade and content area. The report is divided into two sections by item type (multiple-choice and open-response). For multiple-choice items, the content strand and GE code linked to the item are included as well as the total number and percentage of students who answered the item correctly and the number of students who chose each incorrect option or provided an invalid response. An invalid response on a multiple-choice item is defined as "the item was left blank" or "the student selected more than one option for the item." For open-response items, the content strand and GE code linked to the item are included as well as the point value and average score for the item. Users are also able to view the actual released items within the report. If a user clicks on a particular magnifying glass icon next to a released item number, a pop-up box will open, displaying the released item.

8.8.1.4 Longitudinal Data

The Longitudinal Data report is a confidential student-level report that provides individual student performance data for multiple test administrations. Fall 2012 NECAP scores and achievement levels are provided for each tested student in reading, mathematics, and writing. In addition, fall NECAP 2008, 2009, 2010, and 2011 reading, mathematics, and writing scores and achievement levels as well as spring NECAP science scores and achievement levels are also included for students in New Hampshire, Rhode Island, and Vermont. Maine students in grades 3 through 8 will show fall 2009, 2010, 2011 and 2012 NECAP scores and achievement levels in reading and mathematics, since this is only the fourth test administration for Maine since joining NECAP. Student performance on future test administrations will be included on this report over time. This report can be saved, printed, or exported as a PDF file.

8.8.2 User Accounts

In the Analysis & Reporting System, principals have the ability to create unique user accounts by assigning specific usernames and passwords to educators, such as teachers, curriculum coordinators, or special education coordinators, in their school. Once the accounts have been created, individual students may be assigned to each user account. After users have received their usernames and passwords, they are able to log in to their accounts and access the interactive reports, which will be populated only with the subgroup of students assigned to them.

Information about the interactive reports and setting up user accounts is available in the *Analysis & Reporting System User Manual* that is available for download on the Analysis & Reporting System.

8.9 DECISION RULES

To ensure that reported results for the 2012–13 NECAP tests 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 2012 administration of the NECAP is found in Appendix S.

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.

8.9.1 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 Data and Reporting Services and Psychometrics and Research departments, 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 received scaled scores and/or was 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 of samples 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 with schools that 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.

CHAPTER 9 VALIDITY

Because interpretations of test scores, and not a test itself, are evaluated for validity, the purpose of the *2012–13 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 *Standards for Educational and Psychological Testing* (AERA, et al., 1999) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. The evidence around test content, response processes, internal structure, relationship to other variables, and consequences of testing speaks to different aspects of validity but are not distinct *types* of validity. Instead, each aspect of validity contributes to a body of evidence about the comprehensive validity of score interpretations.

Evidence on test content validity is meant to determine how well the assessment tasks represent the curriculum and standards for each grade level and content area. Content validation is informed by the item development process, including how the test blueprints and test items align to the curriculum and standards. Viewed through this lens provided by the standards, evidence based on test content was extensively described in Chapters 2 and 3. Item alignment with NECAP 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 validity evidence based on test content. As discussed earlier, all NECAP questions are aligned by educators from the member states to specific NECAP content standards, and undergo several rounds of review for content fidelity and appropriateness. Items are presented to students in multiple formats (constructed-response, short-answer, and multiple-choice). Finally, tests are administered according to state-mandated standardized procedures, with allowable accommodations, and all test coordinators and administrators are required to familiarize themselves with and adhere to all of the procedures outlined in the *NECAP Principal/Test Coordinator Manual* and the *Test Administrator Manual*.

The scoring information in Chapter 4 describes the steps taken to train and monitor hand-scorers, as well as quality control procedures related to scanning and machine scoring. To speak to student response processes, however, additional studies would be helpful and might include an investigation of students' cognitive methods using think-aloud protocols.

Evidence based on internal structure is presented in great detail in the discussions of item analyses, scaling and equating, and reliability in Chapters 5 through 7. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), differential item functioning analyses, dimensionality analyses, reliability, standard errors of measurement,

and item response theory parameters and procedures. Each test is equated to the same grade and content area test from the prior year in order to preserve the meaning of scores over time. In general, item difficulty and discrimination indices were in acceptable and expected ranges. Very few items were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most items were assessing consistent constructs, and students who performed well on individual items tended to perform well overall.

Evidence based on the consequences of testing is addressed in the scaled score information in Chapter 6 and the reporting information in Chapter 8, as well as in the *Guide to Using the 2012 NECAP Reports*, which is a separate document. Each of these chapters speaks to the efforts undertaken to promote accurate and clear information provided to the public regarding test scores. Scaled scores offer the advantage of simplifying the reporting of results across content areas, grade levels, and subsequent years. Achievement levels provide users with reference points for mastery at each grade and content area, which is another useful and simple way to interpret scores. Several different standard reports are provided to stakeholders. Additional evidence of the consequences of testing could be supplemented with broader investigation of the effect of testing on student learning.

To further support the validation of the assessment program, additional studies might be considered to provide evidence regarding the relationship of NECAP results to other variables, including the extent to which scores from NECAP converge with other measures of similar constructs, and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

9.1 QUESTIONNAIRE DATA

External validity of the NECAP assessment is conveyed by the relationship of test scores and situational variables such as time-spent patterns and attitude toward content matter. These situational variables were all based on student questionnaire data collected during the administration of the NECAP test. Note that no inferential statistics are included in the results presented below; however, because the numbers of students are quite large, differences in average scores may be statistically significant.

9.1.1 Difficulty of Assessment

Examinees in all grades and content areas were asked how difficult the test was relative to their regular schoolwork. In the sections below, results are presented for selected grade levels for each content area.

9.1.1.1 Difficulty: Reading

Figures 9-1 and 9-2 below show that students in grades 8 and 11 who thought the test was easier than their regular reading schoolwork did better overall than those who thought it was more difficult.

Question: How difficult was the reading test?

Figure 9-1. 2012–13 NECAP: Reading Grade 8 Questionnaire Responses—Difficulty

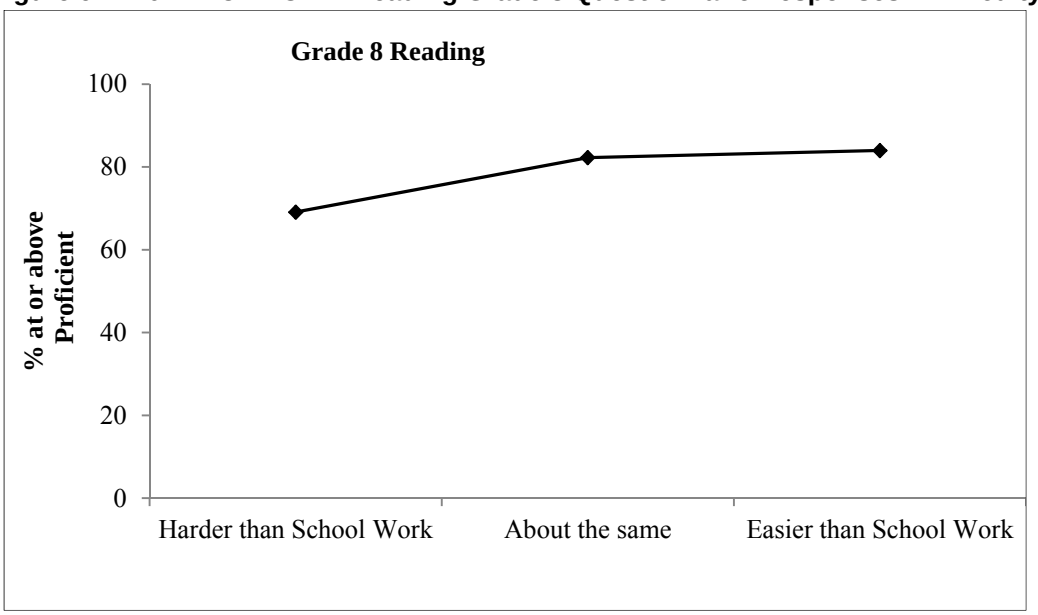
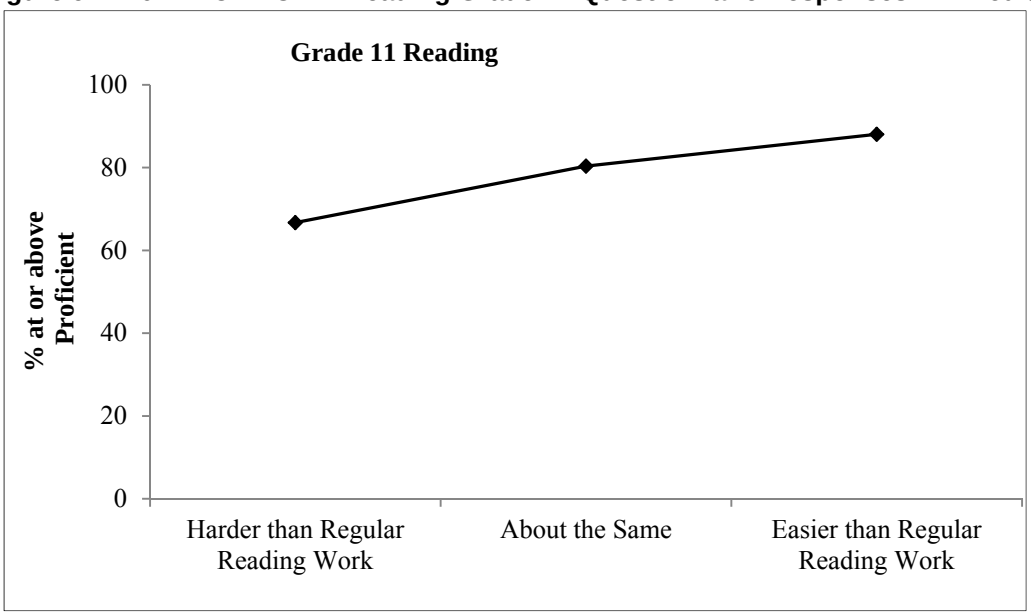


Figure 9-2. 2012–13 NECAP: Reading Grade 11 Questionnaire Responses—Difficulty



9.1.1.2 Difficulty: Mathematics

Figures 9-3 and 9-4 below show a very similar pattern to that for reading: students in grades 8 and 11 who thought the test was easier than their regular mathematics schoolwork did better overall than those who thought it was more difficult.

Question: How difficult was the mathematics test?

Figure 9-3. 2012–13 NECAP: Mathematics Grade 8 Questionnaire Responses—Difficulty

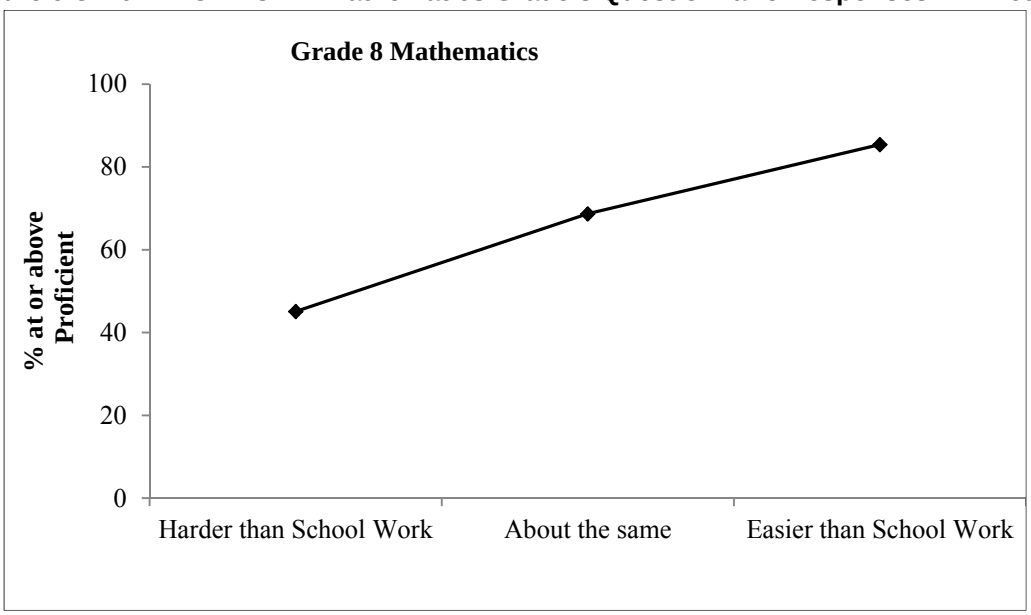
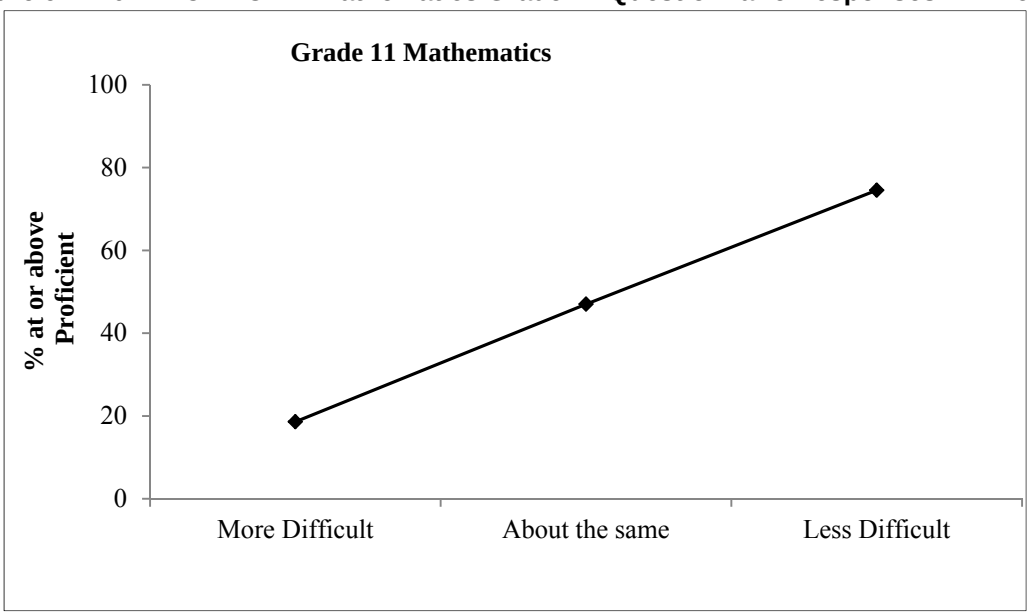


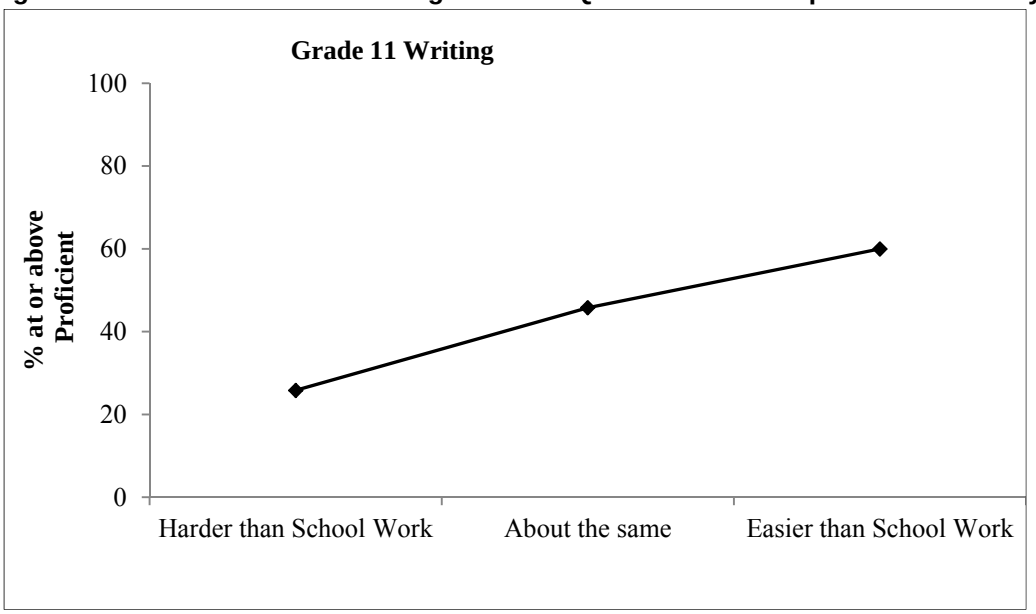
Figure 9-4. 2012–13 NECAP: Mathematics Grade 11 Questionnaire Responses—Difficulty



9.1.1.3 Difficulty: Writing

For writing, as shown in Figure 9-5 below, there was a pronounced relationship between perception of the difficulty of the test and student performance at grade 11.

Figure 9-5. 2012–13 NECAP: Writing Grade 11 Questionnaire Responses—Difficulty



9.1.2 Content

Across grades, examinees were asked about the frequency with which they engage in academic activities (specific to content area) that are expected to be related to test performance. In the sections below, results are presented for selected grade levels for each content area.

9.1.2.1 Content: Reading

Examinees in reading were asked how often they are asked to write at least one paragraph for reading/English language arts (grades 3 through 8) or reading (grade 11) class. Figures 9-6 through 9-9 show that students who indicated they write at least one paragraph a few times a week perform better than any of the other groups.

Figure 9-6. 2012–13 NECAP: Reading Grade 3 Questionnaire Responses—Content

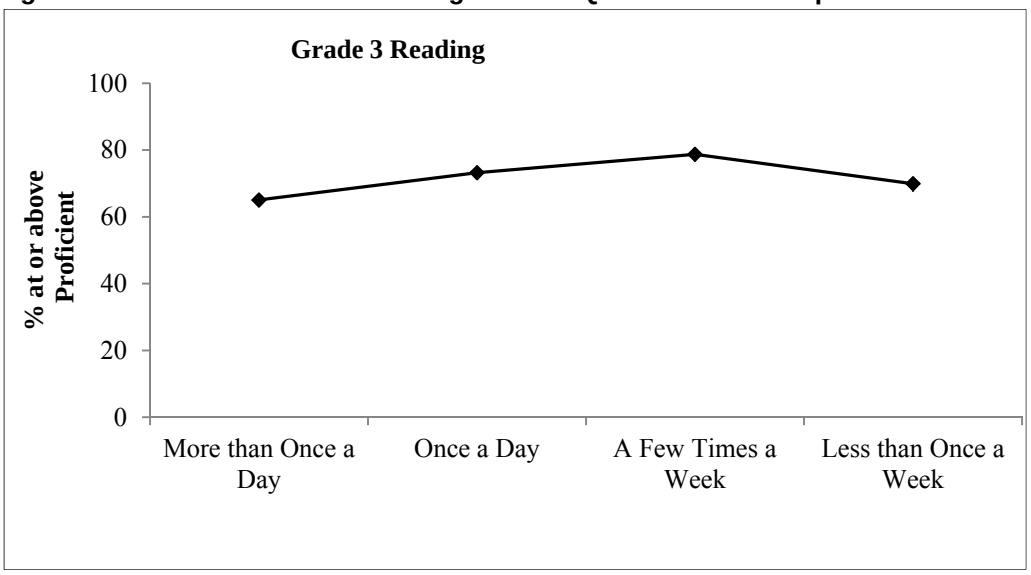


Figure 9-7. 2012–13 NECAP: Reading Grade 4 Questionnaire Responses—Content

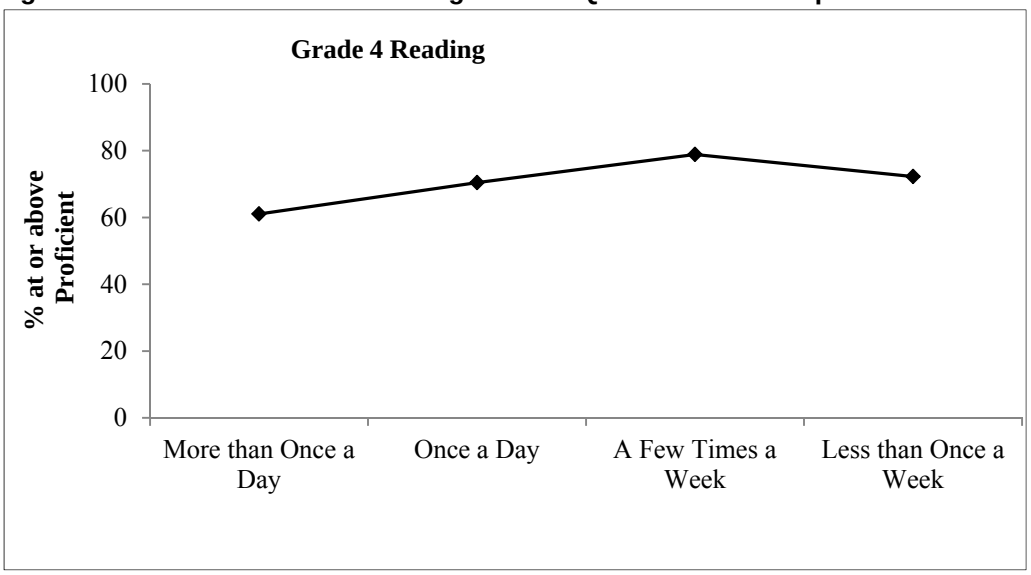


Figure 9-8. 2012–13 NECAP: Reading Grade 7 Questionnaire Responses—Content

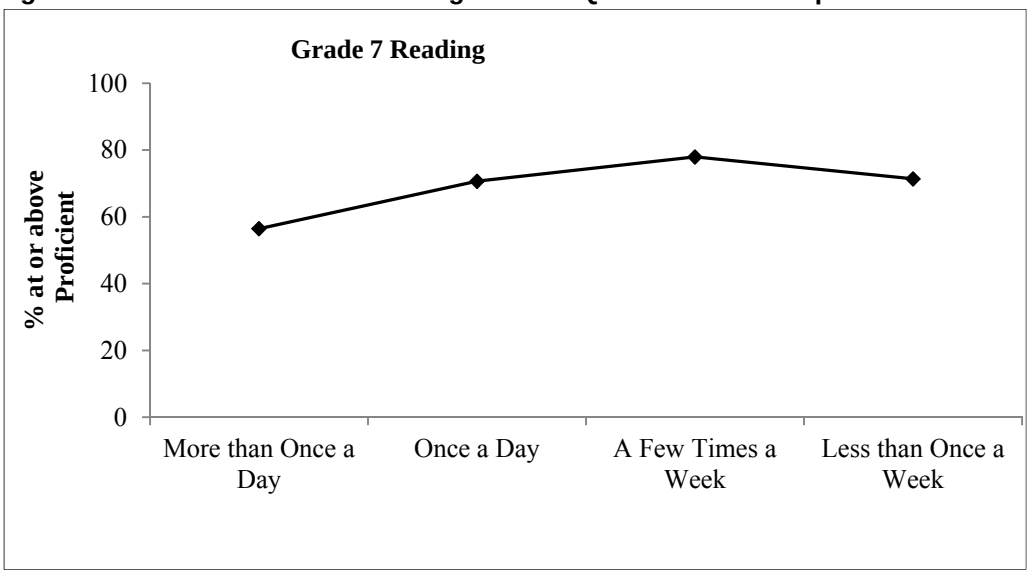
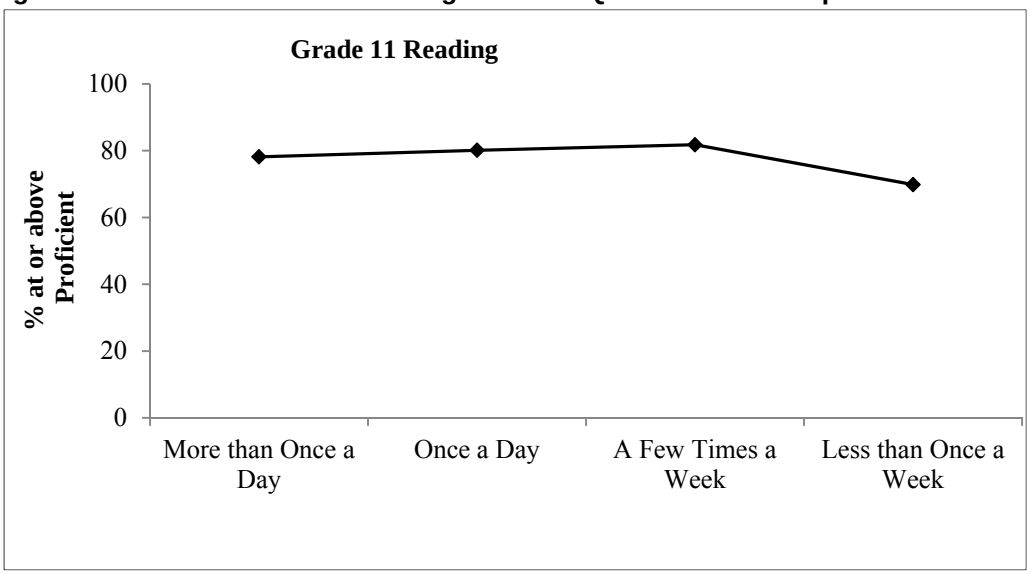


Figure 9-9. 2012–13 NECAP: Reading Grade 11 Questionnaire Responses—Content



9.1.2.2 Content: Mathematics

For mathematics, examinees in grades 7 and 8 were asked whether they were currently enrolled in an Algebra I or higher mathematics class. In grade 11, examinees were asked which mathematics course they last completed (e.g., Geometry). Figures 9-10 through 9-12 seem to suggest that students with more exposure to mathematics coursework tend to perform better than students who have been exposed to fewer mathematics courses.

Figure 9-10. 2012–13 NECAP: Mathematics Grade 7 Questionnaire Responses—Content

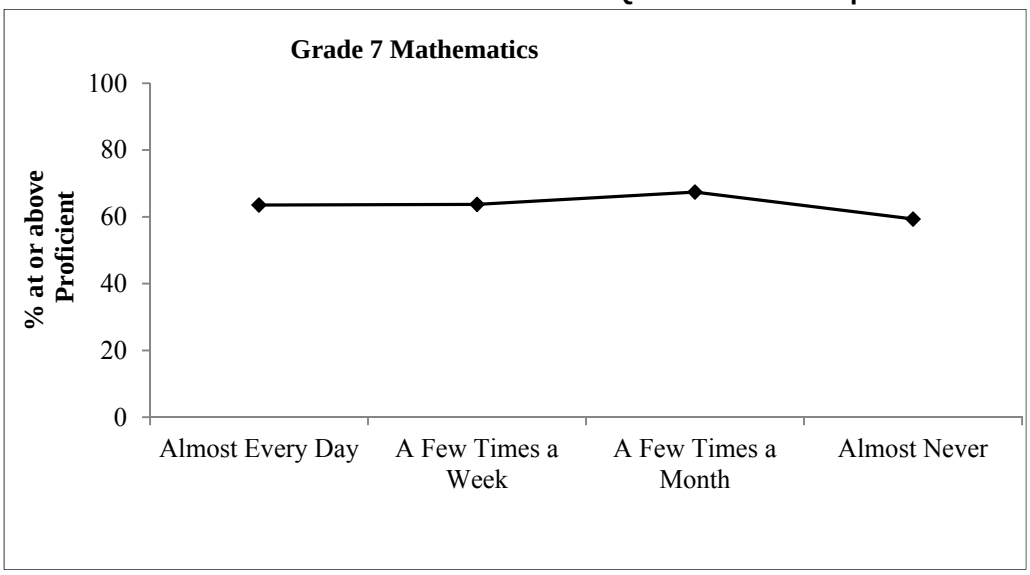


Figure 9-11. 2012–13 NECAP: Mathematics Grade 8 Questionnaire Responses—Content

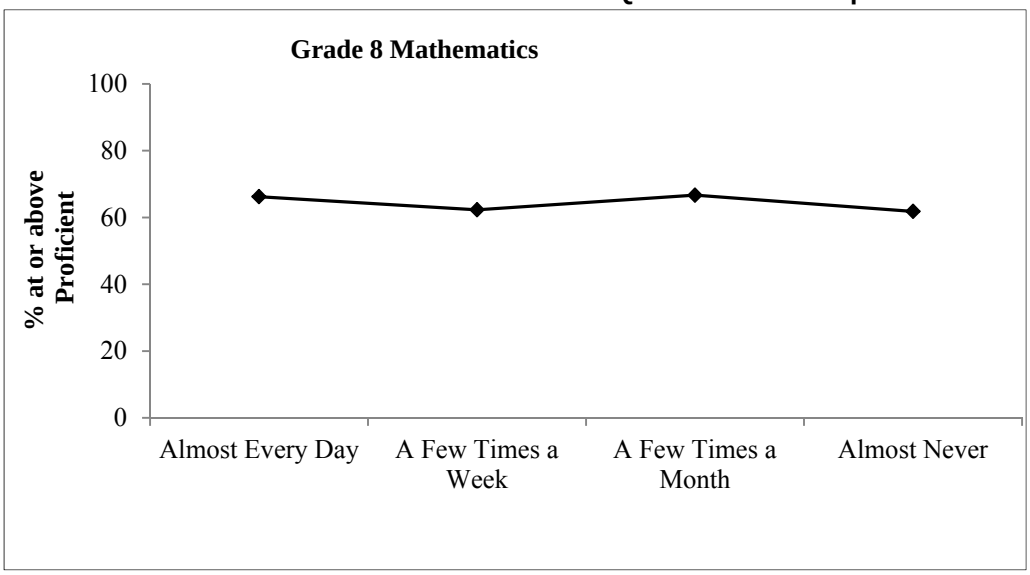
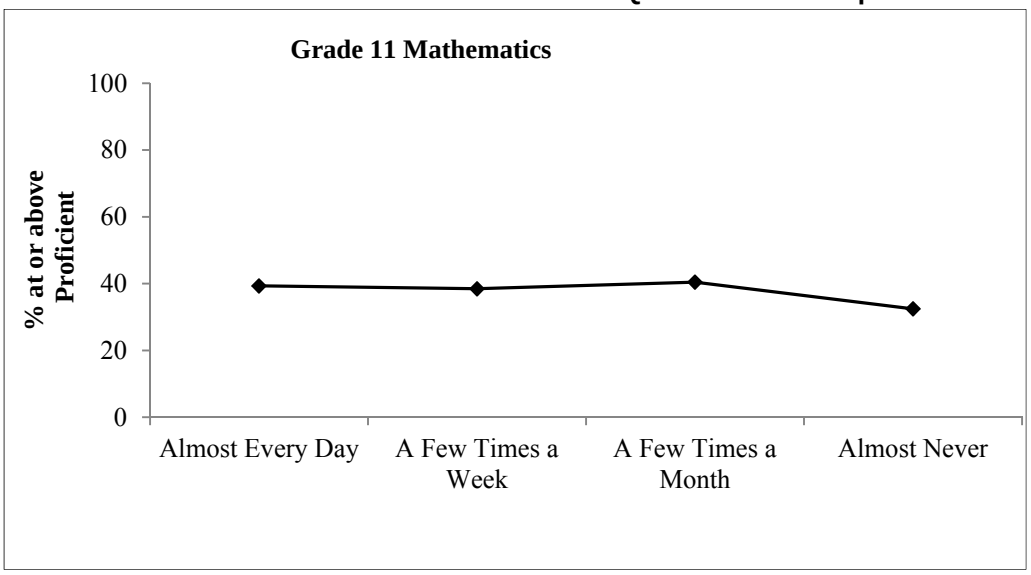


Figure 9-12. 2012–13 NECAP: Mathematics Grade 11 Questionnaire Responses—Content



9.1.2.3 Content: Writing

Examinees in writing were asked how often they write more than one draft. Figures 9-13 through 9-15 show that students who indicated they write multiple drafts more frequently did better than students who write multiple drafts less frequently, although the differences at grade 5 were slight.

Figure 9-13. 2012–13 NECAP: Writing Grade 5 Questionnaire Responses—Content

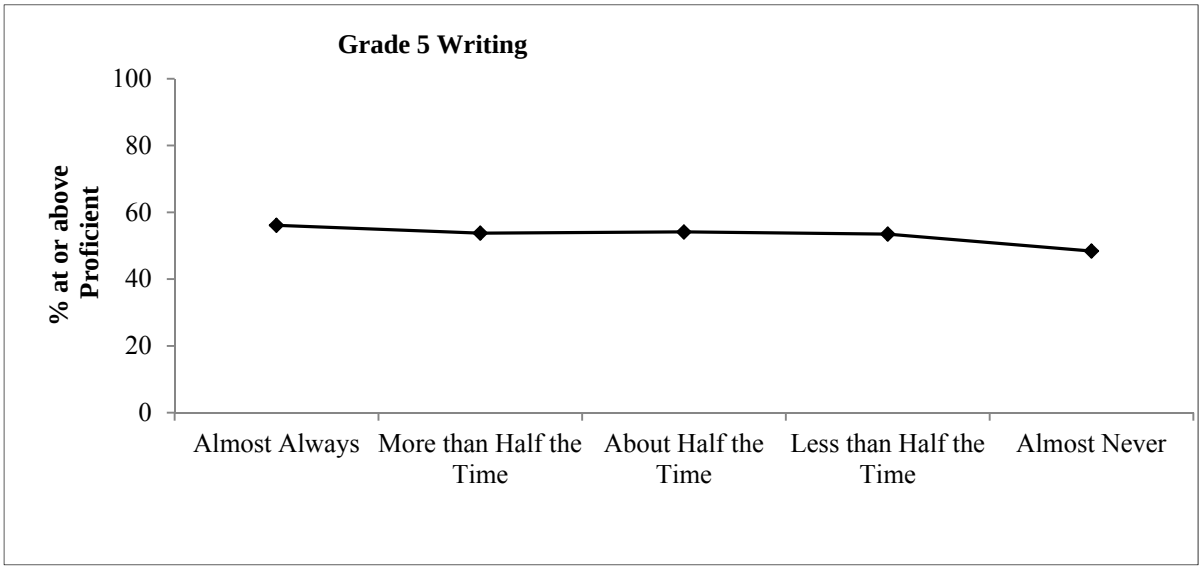


Figure 9-14. 2012–13 NECAP: Writing Grade 8 Questionnaire Responses—Content

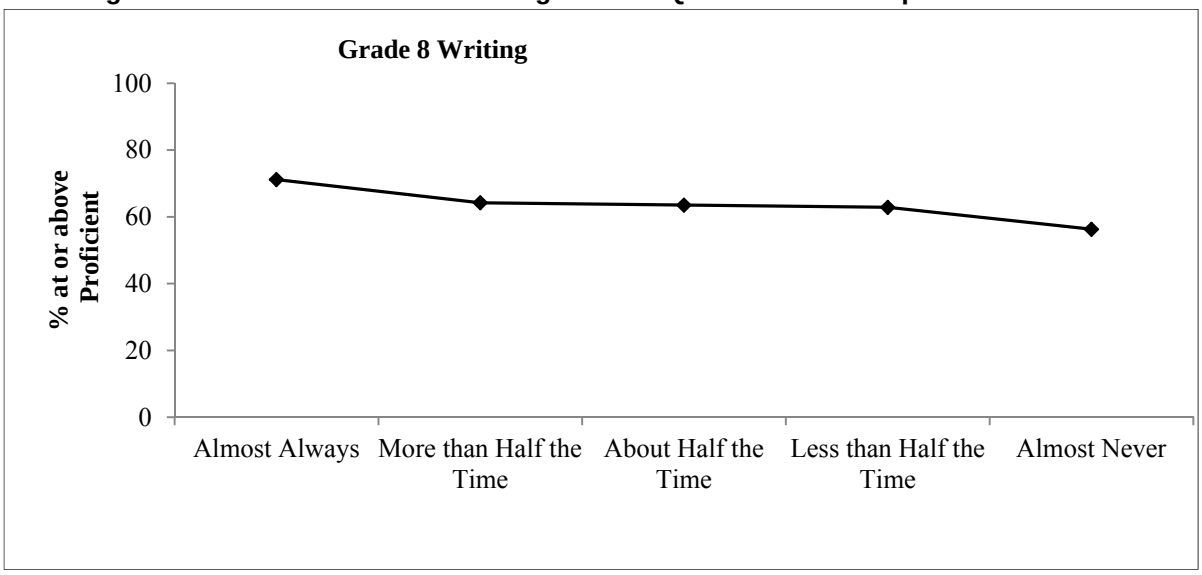
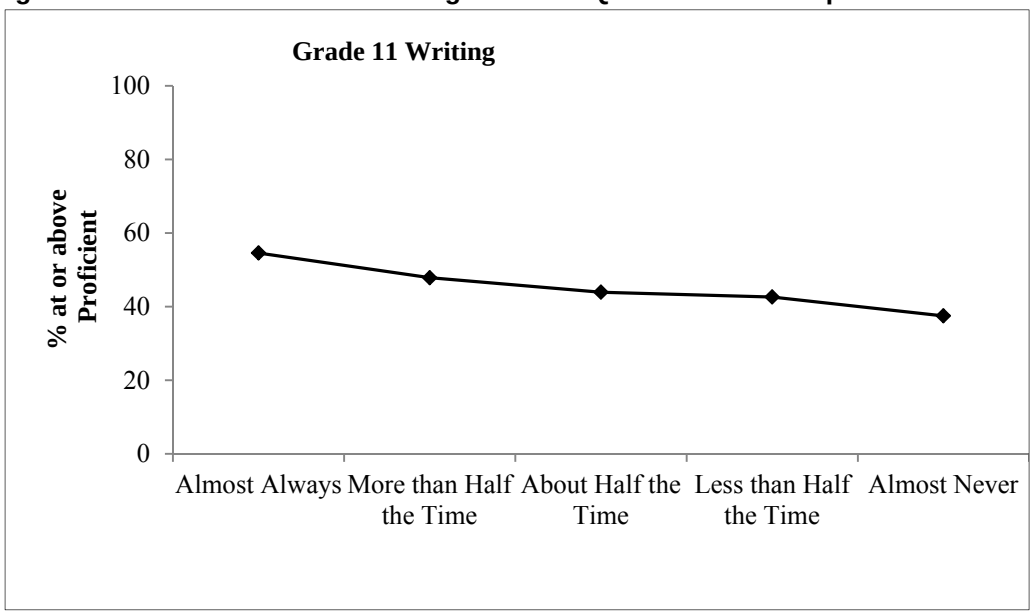


Figure 9-15. 2012–13 NECAP: Writing Grade 11 Questionnaire Responses—Content



9.1.3 Homework

Examinees in all grades in reading and mathematics were asked how often they have homework. In the sections below, results are presented for selected grade levels for each content area.

9.1.3.1 Homework: Reading

Figures 9-16 through 9-18 below show that students in grades 4, 7, and 11 who indicated they had reading homework more frequently performed better than students who had less homework. The relationship is more pronounced in the higher grades.

Figure 9-16. 2012–13 NECAP: Reading Grade 4 Questionnaire Responses—Homework

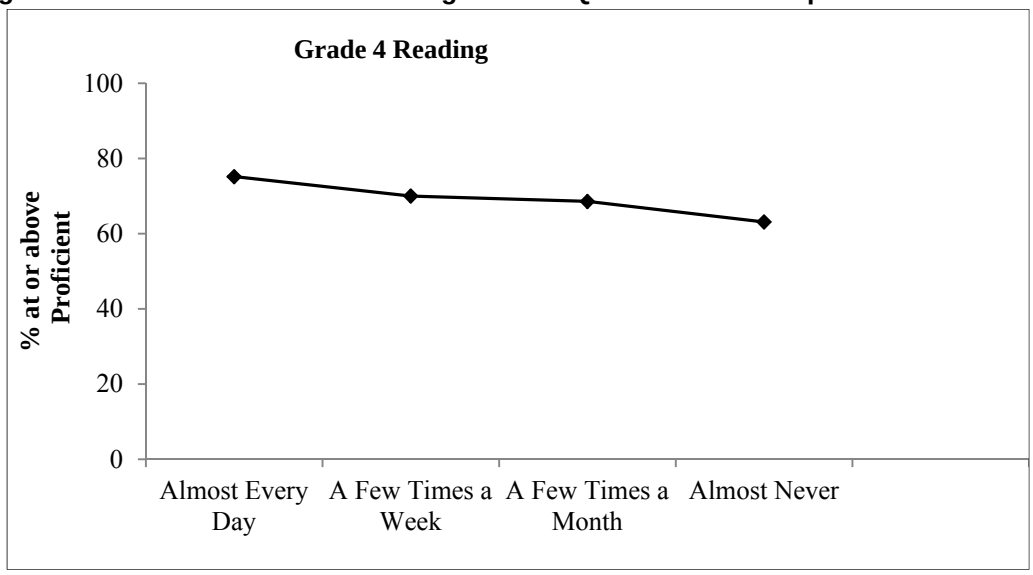


Figure 9-17. 2012–13 NECAP: Reading Grade 7 Questionnaire Responses—Homework

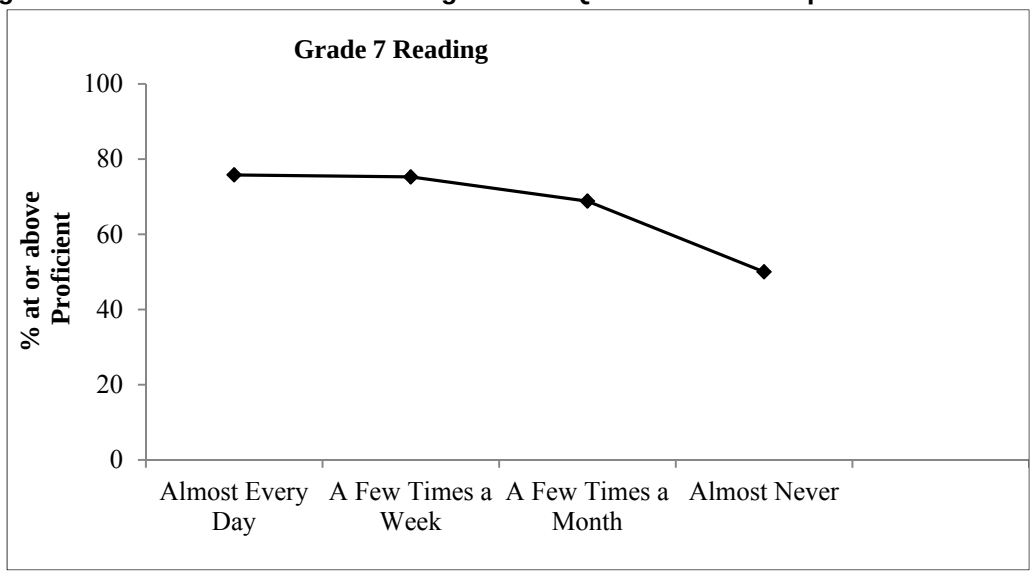
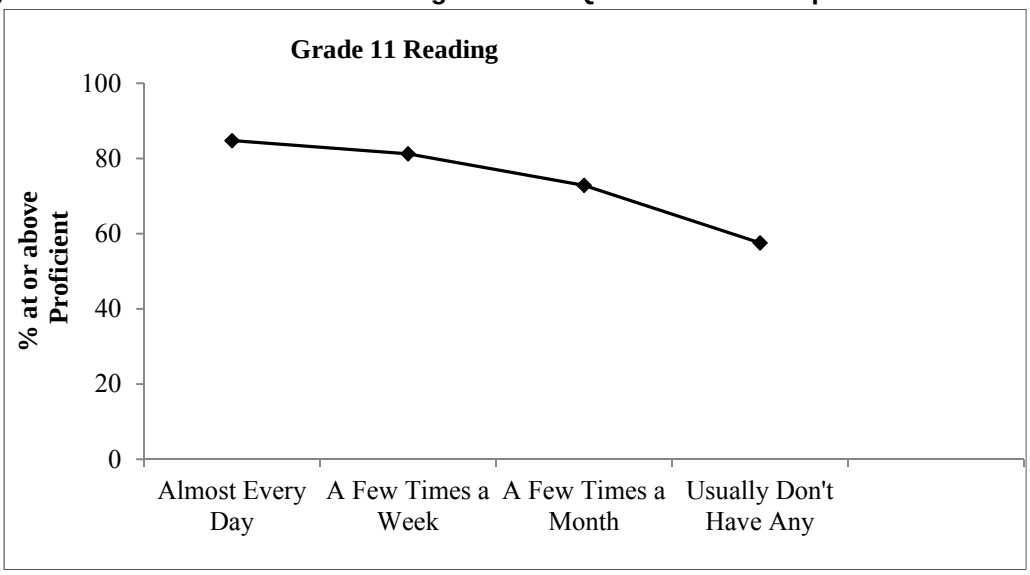


Figure 9-18. 2012–13 NECAP: Reading Grade 11 Questionnaire Responses—Homework



9.1.3.2 Homework: Mathematics

Figures 9-19 through 9-22 below show results that are very similar to those for reading; students in grades 4, 5, 8, and 11 who indicated they had mathematics homework more frequently performed better than students who had less homework. Again, the pattern is more pronounced in the higher grades.

Figure 9-19. 2012–13 NECAP: Mathematics Grade 4 Questionnaire Responses—Homework

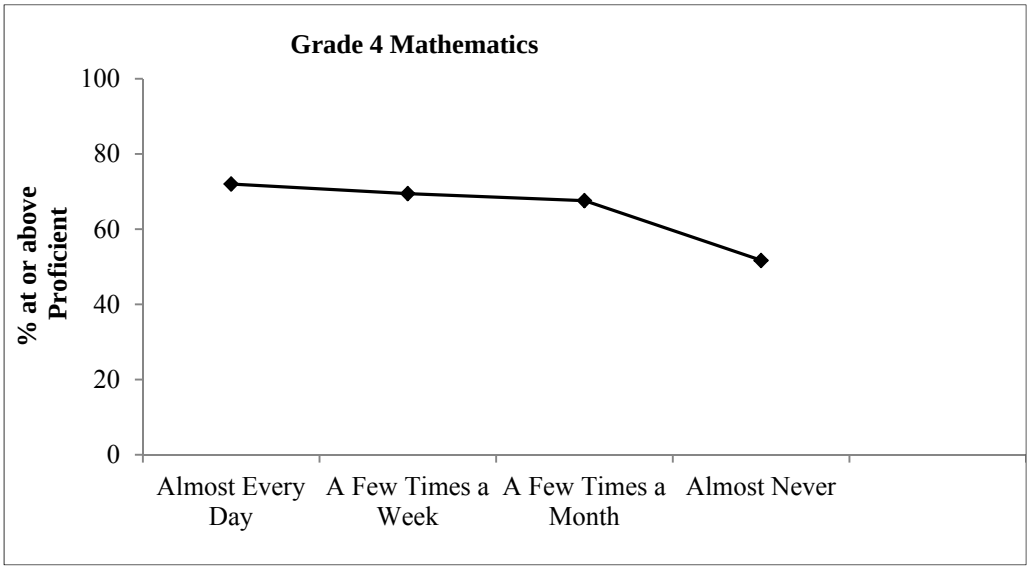


Figure 9-20. 2012–13 NECAP: Mathematics Grade 5 Questionnaire Responses—Homework

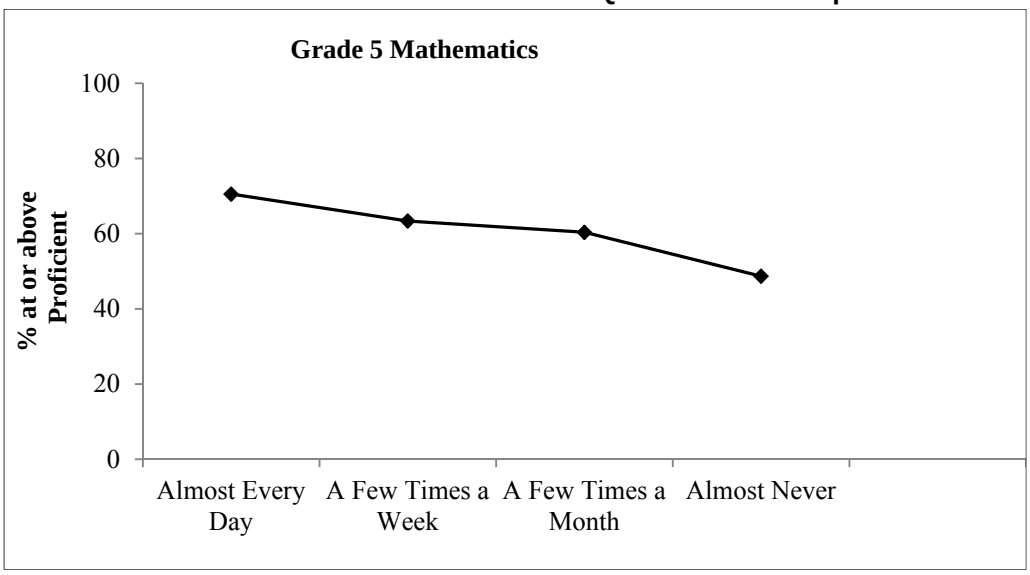


Figure 9-21. 2012–13 NECAP: Mathematics Grade 8 Questionnaire Responses—Homework

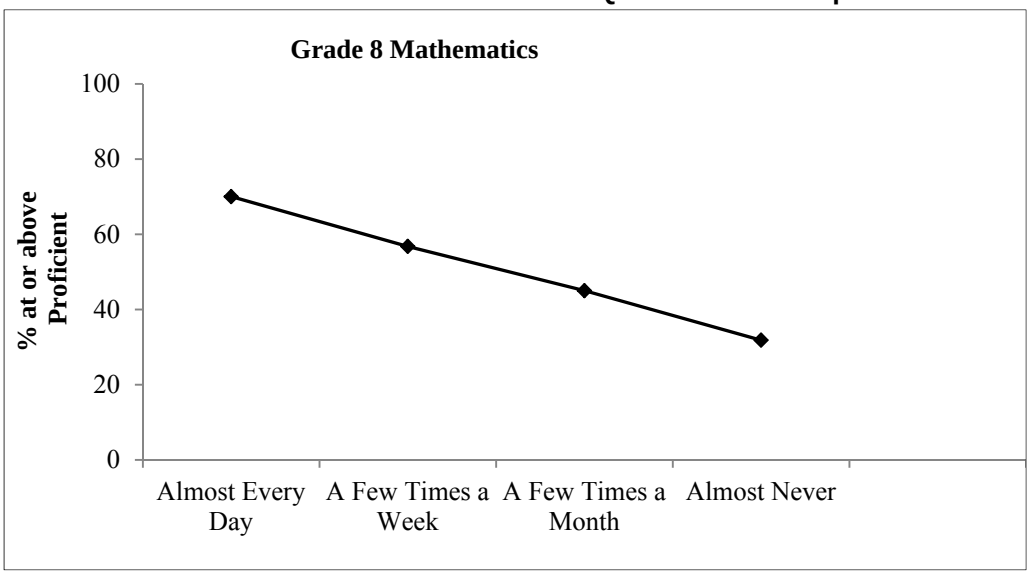
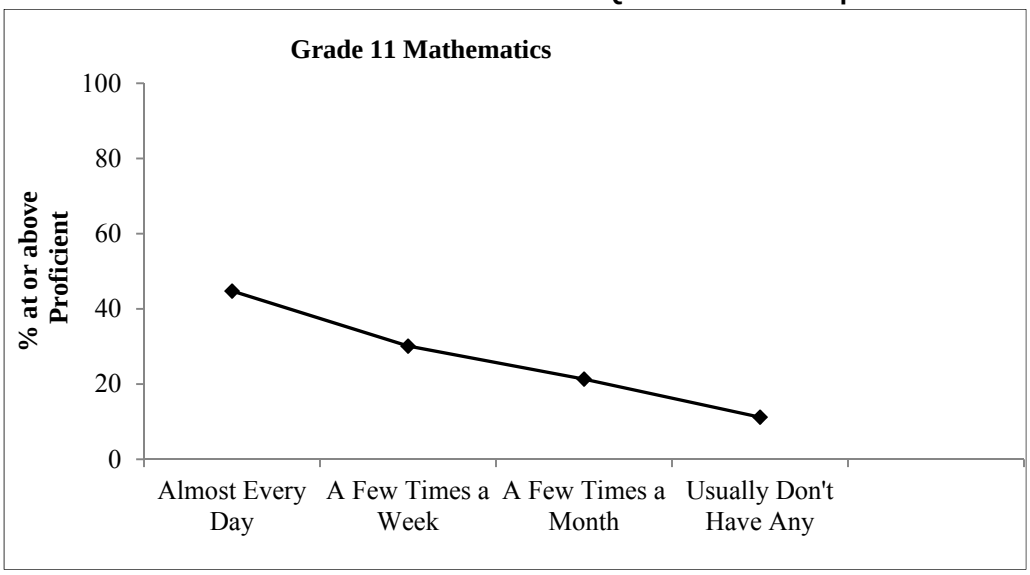


Figure 9-22. 2012–13 NECAP: Mathematics Grade 11 Questionnaire Responses—Homework



9.1.4 Performance in Courses

Students in grade 11 for both reading and mathematics were asked what their most recent course grade was. Figures 9-23 and 9-24 indicate that, for both reading and mathematics, there was a strong positive relationship between the most recent course grade and NECAP scores in that content area.

Figure 9-23. 2012–13 NECAP: Reading Grade 11 Questionnaire Responses—Performance

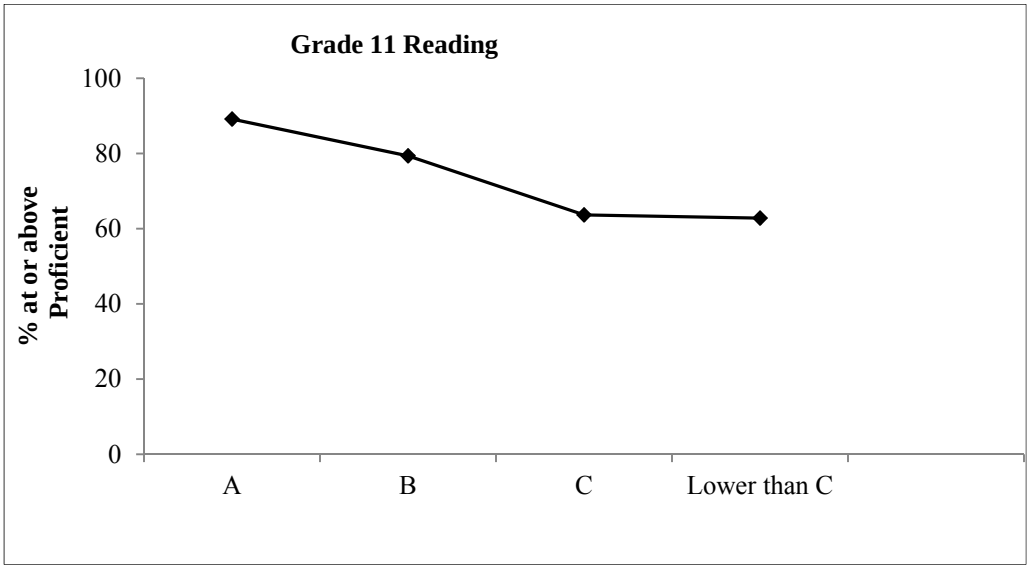
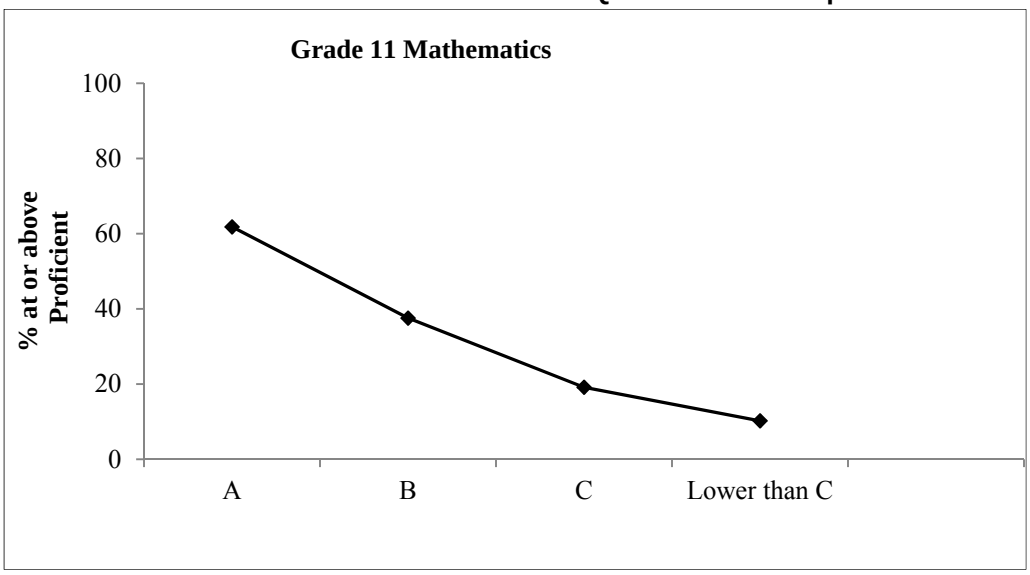


Figure 9-24. 2012–13 NECAP: Mathematics Grade 11 Questionnaire Responses—Performance



REFERENCES

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Belmont, Calif.: Wadsworth, Inc.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.

Baker, F. B., & Kim, S. H. (2004). *Item response theory: Parameter estimation techniques* (2nd ed.). New York: Marcel Dekker, Inc.

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

Chicago Manual of Style (15th ed., 2003). Chicago: University of Chicago Press.

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

Dorans, N. J., & 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., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York: John Wiley and Sons, Inc.

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

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

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

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

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

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

Nering, M., & Ostini, R. (2010). *Handbook of polytomous item response theory models*. New York: Routledge.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 221–262).

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7,* 201–210.

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

Stout, W. F., Froelich, A. G., & 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., & 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 MEMBERS
MAY 8 & 9, 2012

Name	*Association/Affiliation*
Dale Carlson	NAEP Coach, NAEO -Westat
Bill Erpenbach	WJE Consulting, LTD.
Jeff Nellhaus	PARCC Assessment, Achieve, Inc.
Jim Pellegrino	University of Illinois
Joe Ryan	Consultant
Steve Slater	Oregon Department of Education
Martha Thurlow	NCEO, University of Minnesota
Craig Wells	University of Massachusetts Amherst
Laurie Wise	Principal Scientist HumRRO

MAINE ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Beth	Ahlholm	Medomak Middle School	Grade 8 Language Arts & Social Studies Teacher
Marshalyn	Baker	Messalonskee Middle School	Math Teacher
Melissa	Biehn	Hermon Middle School	Literacy Coach
Kathy	Brooks	Hollis School	Grade 4 Teacher
Steven	Corman	Lincoln Middle School	CH 104 Math Consultant
Wendy	Dunbar	Mt. Jefferson Jr. High School	ELA Teacher/ Title 1A Coordinator/ Literacy Specialist
Louise	Erskine	James H. Bean School	Grades 4 & 5 Teacher
Melissa	Goeke	Hope Elementary School	Teacher
Heidi	Goodwin	Margaret Chase Smith School	Literacy Coach
Kate	Greeley	MSAD #75	Elem. Technology Integrator/ Math Lead Teacher
Kathy	Kauffman	Oak Hill Middle School	Grades 7 & 8 ELA Teacher
Evelyn	Krahn	Paris Elementary School	Grade 3 Teacher
Tracy	Lapointe	Marshwood Middle School	Grade 8 ELA Teacher
Karen	Mayo	James H. Bean School	Grade 3 Teacher
Eleanor	Merrick	Emerson School	Principal
Kayanne	Nadeau	Wallagrass Elementary School	Literacy Coach
Eileen	Pew	Meroby Elementary School	Literacy Coach
Narda	Plant	Fort Fairfield Elementary School	Special Education Teacher
Gail	Saldanha	Riverton Elementary School	Title 1 Math
Beth	Soules	Wentworth Intermediate School	Grades 3-5 Math Support Teacher
Terry	Taiani	Jordan-Small Middle School	Grades 5 & 6 Math & Science Teacher

NEW HAMPSHIRE ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Beverly	Alexander	Milford Middle School	Grade 6 Math Teacher
Coleen	Bridle	Iber Holmes Gove Middle School	Math Teacher
Shane	Clifford	Bernise A. Ray School	Grade 5 Teacher
Donna	Dubey	Winnisquam Regional High School	Math Teacher
Lisa	Dwyer	Merrimack Valley Middle School	Grade 8 Reading & Language Arts Teacher
Jack	Finley	Franklin High School	Teacher
Katherine	Franz	Milford Middle School	Grade 8 ELA Teacher
Jara	Freeman	Milford Middle School	ELA Teacher
Michael	Gile	Dover High School	Math Teacher
Linda	Guidi	Nashua School District	K-8 Math Coach
Katrina	Hall	Hollis Brookline Middle School	Math Teacher
Heidi	Hanson-Gillis	Henry Wilson Memorial School	Grade 8 Language Arts Teacher
Heidi	Hutchison	Dover Middle School	Grade 8 Language Arts Teacher
Erica	Johnson	Dover High School	Math Teacher
Sharon	Knapp	Kearsarge Regional at New London	Grade 2 Teacher
Natalie	LaFlamme	Fairgrounds Middle School	Grade 6 Math Teacher
Patricia	Maestranzi	Woodbury School	Literacy Specialist
William	Maniotis Jr.	Merrimack High School	English Teacher/ Literary Specialist
Leslie	McRobie	Idlehurst Elementary School	Grade 5 Teacher
Susan	Megas-Russell	Cutler School	Literacy/ Reading Specialist
Nick	Miranda	Berlin High School	Grades 9-12 Math Teacher
Jerrell	Moodie	Conval High School	Grades 9-12 English & Reading Teacher
Jake	Moore	Berlin High School	Teacher
Linda	Nelson	Barnstead Elementary	Grades 4-6 Reading Enrichment
Sarah	Potter	Deerfield Community School	Grade 5 Teacher
Deborah	Rockwell	Dover Middle School	Grade 8 Math Teacher

NEW HAMPSHIRE ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012 *(Continued)*

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Judy	Ross	Lebanon School District	K-8 Math & Science Coordinator
Chris	Saunders	Nashua High School North	English Development Head
Karen	Shackford	Freedom Elementary	Reading Teacher
Bonnie	Skogsholm	Main Dunstable School	Grade 2 Teacher
Robert	Stone	Bernice A. Ray School	Grade 5 Teacher
Shelly	Vessels	Londonderry Middle School	Teacher
Anne	Wallace	Hampstead Middle School	Math Specialist
Ann	West	Bow Elementary School	Grade 3 Teacher
Kathy	Woodbury	Gilsum Elementary	Reading Specialist/ Special Education Care Manager
Kathi	Yusko	Newport Middle and High School	Title 1/ Grades 7 & 8 Teacher/ High School Math Dept. Head

RHODE ISLAND ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Erin	Bagwell	Davisville Middle School	Math Teacher
Cynthia	Ballard	Alan Feinstein Middle School of Coventry	English Teacher
Mary	Bettez	Alan Feinstein Middle School of Coventry	ELA Curriculum Coordinator
Joseph	Cavaco	William E Tolman Senior High School	Math Teacher
Christina	Cipolla	Wm. M. Davies Jr. Career-Technical High School	Grades 9-12 Reading Specialist
Leanne	Dube	Mt. Pleasant High School	ELA Teacher Leader
Donna	Gattinella	Central Falls School District	District Math Coach
Mia	Godbout	Dr. Earl F. Calcutt Middle School	Literacy Coach
Maria	Gould	Frank D. Spaziano Elementary School	Reading Coach
Nicolle	Greene	E.G. Robertson School	Grade 3 Teacher
Adam	Heywood	Pleasant View District	Reading Coach
Kenneth	Kard	Segue Institute for Learning	Grade 7 Math Teacher
Tracy	Lafreniere	North Smithfield Elementary School	Reading Specialist
Julie	Lamarre	Segue Institute for Learning	Grade 7 ELA Teacher
Luciana	Lancellotta	Westerly Middle School	Special Educator/ Reading Specialist
Debra	Lancia	Cranston Public Schools	District Math Coach
Joanne	Lynch	Greystone Elementary School	Grade 4 Teacher
Charlee	McElroy	Sherman School	Grade 3 Teacher
Carolyn	Melillo	Central Falls School District	Literacy Coach
Lisa	Melmed	Mt. Hope High School	English Teacher
Jenna	Mendonsa	Segue Institute for Learning	Numeracy Coach
Ryan	Mullen	Pilgrim High School	Math Teacher
Tracey	Nangle	North Smithfield Middle School	ELA & Social Studies Teacher
Nicola	Pekrul	Alan Feinstein Middle School of Coventry	Curriculum Coordinator/ Math Teacher

RHODE ISLAND ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012 *(Cont.)*

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Susan	Round	Central Falls School District	K-2 Literacy Coach
Monique	Rousselle Maynard	West Warwick High School	Math Teacher
Rebecca	Russell	Blackstone Valley Prep	Grade 2 Teacher
Kevin	Seekell	Alan Shawn Feinstein Middle School Of Coventry	Math Curriculum Coordinator
Eva	Silver	Central Falls School District	Math Coach
Eileen	Szymkowicz	Pilgrim High School	Math Teacher

VERMONT ITEM REVIEW COMMITTEE
FEBRUARY 6–8, 2012

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Lorraine	Bargmann Metz	Rutland Intermediate School	Literacy & Enrichment Coordinator
Renee	Berthiaume	North Country Union High School	High School Language Arts Teacher
Jennifer	Bisson	Bedford Elementary School	Grade 6 Teacher
Suzanne	Buck	Vergennes Union High School	Special Education/ Assessment Coordinator
Ellen	Cameron	Cavendish Town Elementary School	Grade 3 Teacher
Tamara	Cooley	Barre Town School	Grade 4 Teacher
Barbara	Dall	Barre City Elementary And Middle School	Grades 5-8 ELL Teacher
Sandy	Frizzell	North Country Union High School	English Teacher
Linda	Gilbert	Dothan Brook School	Grade 4 Teacher/ Math Leader
Kelley	Green	Central Elementary School	Grade 3 Teacher
Jennifer	Harper	Cavendish Town Elementary School	Grade 4 Teacher
Kathleen	Jacob	Mill River Union High School	Grade 7 Math Teacher
Dana	Johnson	Rutland Intermediate School	K-6 District Math Coach
Lindy	Johnson	Washington Central Supervisory Union	Literacy Coordinator
Stefanie	Kingzett	Barre Town School	Grade 6 Math Teacher
Geoff	Lawrence	Proctor Jr/Sr High School	English Teacher
Karen	Lawson	Concord School	Grade 5 Math Teacher
Diana	Martel	Barre Town Middle & Elementary	Grade 4 Teacher
Susan	McEwing	St. Albans Town Educational Center	Math Content/ Tech. Integration Specialist
Cathy	Newton	Dothan Brook School	Language Arts Teacher/ Title 1
John	Pandolfo	U-32	Math Teacher/ Dept. Head/ Curriculum Leader
Bernie	Peatman	Rutland South Supervisory Union	Curriculum Coordinator
Nancy	Pejouhy	Woodside Union Middle School	Grade 8 Math Teacher
Wendy	Pratt	Whitcomb Jr Sr High School	Math Teacher
Travis	Redman Jr.	Rutland Town School	Grade 8 Teacher

BIAS AND SENSITIVITY REVIEW COMMITTEE
FEBRUARY 6–8, 2012

MAINE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Shelly	Chasse-Johndro	University of Maine	ESL Teacher Trainer
Sheree	Granger	Sweetser Children's Services	Special Education Teacher
Deborah	Howard	Consultant	Teacher for the Deaf/ ESL
Bruno	Yomoah	Deering High School	ESL Teacher

NEW HAMPSHIRE

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Chris	Blackstone	Franklin Middle and High School	Curriculum Coordinator
Marisa	Bozek	Epping Middle School	Teacher
Kathaleen	Cobb	Cutler School	Grade 4 Teacher
Christine	Leach	Exeter High School	Guidance Counselor
Linda	Scribner	Dover Middle School	Special Education Teacher/ LD Specialist

RHODE ISLAND

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Darline	Berrios	Paul Cuffee Charter School	Grade 5 Teacher
Kimberly	Cabana	Alan Shawn Feinstein Elementary at Broad Street	K-5 Reading Coach
Johanna	Connell	Sherlock Center on Disabilities	Teacher for the Visually Impaired
Danielle	Smith	Mount Pleasant High School	ELL Teacher Leader

VERMONT

First Name	*Last Name*	*School/Association/Affiliation*	*Position*
Katherine	Blair	J F Kennedy Elementary School	Special Educator
Colleen	Fiore	Long Trail School	Special Services Director
Trisha	Klien	VT Association for the Blind and Visually Impaired	Certified Orientation & Mobility Instructor
Kirsten	Kollgaard	J F Kennedy Elementary School	ELL Teacher
Susanna	Mason	Rutland Intermediate School	K-12 ELL Teacher

APPENDIX B—PARTICIPATION RATES

Table B-1. 2012–13 NECAP: Summary of Participation by Demographic Category—Mathematics

Description	*Tested*	
	Number	*Percent*
All Students	296,491	100.00
Male	152,512	51.44
Female	143,677	48.46
Gender Not Reported	302	0.10
Hispanic or Latino	22,366	7.54
American Indian or Alaskan Native	1,580	0.53
Asian	7,061	2.38
Black or African American	11,058	3.73
Native Hawaiian or Pacific Islander	302	0.10
White (non-Hispanic)	247,475	83.47
Two or More Races (non-Hispanic)	5,830	1.97
No Primary Race/Ethnicity Reported	819	0.28
Currently Receiving LEP Services	9,433	3.18
Former LEP Student – Monitoring Year 1	1,403	0.47
Former LEP Student – Monitoring Year 2	1,368	0.46
LEP: All Other Students	284,287	95.88
Students with an IEP	43,152	14.55
IEP: All Other Students	253,339	85.45
Economically Disadvantaged Students	117,834	39.74
SES: All Other Students	178,657	60.26
Migrant Students	124	0.04
Migrant: All Other Students	296,367	99.96
Students Receiving Title 1 Services	52,579	17.73
Title 1: All Other Students	243,912	82.27
Plan 504	5,296	1.79
Plan 504: All Other Students	291,195	98.21

Table B-2. 2012–13 NECAP: Summary of Participation by Demographic Category—Reading

Description	*Tested*	
	Number	*Percent*
All Students	295,822	100.00
Male	152,191	51.45
Female	143,330	48.45
Gender Not Reported	301	0.10
Hispanic or Latino	21,948	7.42
American Indian or Alaskan Native	1,580	0.53
Asian	6,872	2.32
Black or African American	10,924	3.69
Native Hawaiian or Pacific Islander	301	0.10
White (non-Hispanic)	247,555	83.68
Two or More Races (non-Hispanic)	5,827	1.97
No Primary Race/Ethnicity Reported	815	0.28

continued

Description	Tested	
	Number	Percent
Currently Receiving LEP Services	8,638	2.92
Former LEP Student – Monitoring Year 1	1,403	0.47
Former LEP Student – Monitoring Year 2	1,368	0.46
LEP: All Other Students	284,413	96.14
Students with an IEP	43,210	14.61
IEP: All Other Students	252,612	85.39
Economically Disadvantaged Students	117,470	39.71
SES: All Other Students	178,352	60.29
Migrant Students	115	0.04
Migrant: All Other Students	295,707	99.96
Students Receiving Title 1 Services	52,985	17.91
Title 1: All Other Students	242,837	82.09
Plan 504	5,302	1.79
Plan 504: All Other Students	290,520	98.21

Table B-3. 2012–13 NECAP: Summary of Participation by Demographic Category—Writing

Description	Tested	
	Number	Percent
All Students	119,381	100.00
Male	61,252	51.31
Female	58,001	48.58
Gender Not Reported	128	0.11
Hispanic or Latino	8,738	7.32
American Indian or Alaskan Native	592	0.50
Asian	2,717	2.28
Black or African American	4,387	3.67
Native Hawaiian or Pacific Islander	124	0.10
White (non-Hispanic)	100,196	83.93
Two or More Races (non-Hispanic)	2,204	1.85
No Primary Race/Ethnicity Reported	423	0.35
Currently Receiving LEP Services	2,851	2.39
Former LEP Student – Monitoring Year 1	517	0.43
Former LEP Student – Monitoring Year 2	669	0.56
LEP: All Other Students	115,344	96.62
Students with an IEP	17,681	14.81
IEP: All Other Students	101,700	85.19
Economically Disadvantaged Students	44,364	37.16
SES: All Other Students	75,017	62.84
Migrant Students	36	0.03
Migrant: All Other Students	119,345	99.97
Students Receiving Title 1 Services	16,717	14.00
Title 1: All Other Students	102,664	86.00
Plan 504	2,446	2.05
Plan 504: All Other Students	116,935	97.95

APPENDIX C—ACCOMMODATION FREQUENCIES BY CONTENT AREA

Table C-1. 2012–13 NECAP: Numbers of Students Tested with and without Accommodations by Content Area and Grade

Content Area	*Grade*	*Number of Students Tested:*	
		with Accommodations	*without Accommodations*
Mathematics	3	8,709	35,269
	4	9,004	34,554
	5	9,490	34,530
	6	8,216	36,366
	7	7,691	36,827
	8	7,131	37,455
	11	5,234	26,015
Reading	3	8,371	35,463
	4	8,952	34,481
	5	9,482	34,410
	6	8,435	36,047
	7	7,736	36,691
	8	7,283	37,237
	11	4,376	26,858
Writing	5	8,656	35,176
	8	6,499	37,867
	11	3,649	27,534

Table C-2. 2012–13 NECAP: Numbers of Students Tested with Accommodations by Accommodation Type and Subject—Mathematics

Accommodation Code	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*	*Grade 11*
mataccomT1	3,671	4,077	4,892	4,217	3,903	3,618	3,393
mataccomT2	182	169	209	221	135	123	54
mataccomT3	3,501	3,671	3,611	2,929	2,493	2,357	750
mataccomT4	176	162	193	152	125	130	51
mataccomS1	5,786	6,132	5,991	4,492	3,632	3,366	2,172
mataccomS2	31	37	28	56	45	62	19
mataccomP1	1,251	1,363	1,220	1,027	829	651	192
mataccomP2	6,123	6,296	6,228	5,601	5,259	4,903	1,938
mataccomP3	4,749	4,919	4,454	3,239	2,435	1,958	411
mataccomP4	489	511	485	507	601	478	206
mataccomP5	2,675	2,808	2,865	2,275	1,930	1,818	278
mataccomP6	17	39	17	18	26	17	10
mataccomP7	2,823	3,099	2,794	2,075	1,580	1,374	412
mataccomP8	42	42	50	44	47	40	15
mataccomP9	5	4	3	3	3	2	4
mataccomP10	32	12	16	45	54	76	75
mataccomP11	159	191	182	162	83	54	10
mataccomR1	1,621	1,727	1,612	1,149	736	555	59
mataccomR2	37	27	26	30	36	23	4
mataccomR3	318	317	317	164	101	70	18
mataccomR4	27	29	77	116	112	134	47
mataccomR5	132	134	134	87	54	40	23
mataccomR6	9	6	13	10	3	15	3

continued

Accommodation Code	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
mataccomR7	10	7	7	29	28	41	51
mataccomO1	6	7	6	2	6	9	2
mataccomM1	11	15	4	11	9	39	58
mataccomM3	1	2	0	2	0	2	3

Table C-3. 2012–13 NECAP: Numbers of Students Tested with Accommodations by Accommodation Type and Subject—Reading

Accommodation Code	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 11
reaaccomT1	3,559	4,301	5,043	4,468	3,952	3,813	2,502
reaaccomT2	180	170	225	224	140	122	66
reaaccomT3	3,542	3,778	3,661	2,968	2,540	2,367	785
reaaccomT4	183	173	191	164	129	131	50
reaaccomS1	5,641	6,071	5,980	4,513	3,653	3,383	2,119
reaaccomS2	30	30	25	57	46	55	16
reaaccomP1	1,261	1,358	1,225	1,037	854	667	197
reaaccomP2	5,951	6,206	6,193	5,626	5,269	4,897	1,950
reaaccomP4	645	739	728	626	674	546	247
reaaccomP5	2,591	2,780	2,845	2,255	1,915	1,799	282
reaaccomP6	18	37	19	20	27	16	7
reaaccomP7	2,762	3,027	2,765	2,038	1,580	1,378	416
reaaccomP8	43	44	51	44	47	41	13
reaaccomP9	5	4	3	3	3	2	3
reaaccomP11	209	255	200	141	82	54	10
reaaccomR1	1,929	2,092	1,980	1,481	1,032	707	73
reaaccomR2	44	37	30	37	45	26	4
reaaccomR3	353	332	330	176	104	74	34
reaaccomR4	40	47	125	211	262	321	116
reaaccomR5	125	138	128	89	59	40	21
reaaccomR6	11	7	13	10	3	17	3
reaaccomO1	7	8	16	4	5	7	3
reaaccomM2	12	10	7	7	45	9	11
reaaccomM3	0	2	1	3	0	2	4

Table C-4. 2012–13 NECAP: Numbers of Students Tested with Accommodations by Accommodation Type and Subject—Writing

Accommodation Code	Grade 5	Grade 8	Grade 11
wriaccomT1	4,108	3,091	1,586
wriaccomT2	188	98	30
wriaccomT3	3,324	2,143	785
wriaccomT4	180	120	50
wriaccomS1	5,631	3,105	2,111
wriaccomS2	23	52	17
wriaccomP1	1,084	592	201
wriaccomP2	5,900	4,556	2,028
wriaccomP3	4,023	1,789	498
wriaccomP4	453	451	241
wriaccomP5	2,701	1,713	284

continued

Accommodation Code	*Grade 5*	*Grade 8*	*Grade 11*
wriaccomP6	18	17	7
wriaccomP7	2,671	1,292	414
wriaccomP8	49	36	13
wriaccomP9	3	2	3
wriaccomP10	9	46	68
wriaccomP11	167	52	9
wriaccomR3	301	66	33
wriaccomR4	194	410	153
wriaccomR5	112	30	16
wriaccomR6	12	16	2
wriaccomR7	5	18	44
wriaccomO1	9	10	2
wriaccomO2	58	54	29
wriaccomM3	3	1	6

APPENDIX D—TABLE OF STANDARD ACCOMMODATIONS

NECAP Table of Standard Accommodations

Revised August 2009

Any accommodation(s) used for the assessment of an individual student will be the result of a team decision made at the local level. All decisions regarding the use of accommodations must be made on an individual student basis – not for a large group, entire class, or grade level. Accommodations are available to all students on the basis of individual need regardless of disability status and should be consistent with the student's normal routine during instruction and assessment. This table is not intended to be used as a stand-alone document and should always be used in conjunction with the *NECAP Accommodations Guide*.

T. Timing		
Code	Tests were administered	Details on Delivery of Accommodations
T1	with time to complete a session extended beyond the scheduled administration time within the same day.	NECAP tests are not designed to be timed or speeded tests. The scheduled administration time already includes additional time and the vast majority of students complete the test session within that time period. Extended time within a single sitting may be needed by students who are unable to meet time constraints. A test session may be extended until the student can no longer sustain the activity.
T2	so that only a portion of the test session was administered on a particular day.	In rare and severe cases, the extended time accommodation (T1) may not be adequate for a student not able to complete a test session within a single day. A test session may be administered to a student as two or more "mini-sessions" if procedures are followed to maintain test security and ensure that the student only has access to the items administered on that day (see the *NECAP Accommodations Guide* for details).
T3	with short, supervised breaks.	Multiple or frequent breaks may be required by a student whose attention span, distractibility, or physical condition, requires shorter working periods.
T4	at the time of day or day of week that takes into account the student's medical needs or learning style.	Individual scheduling may be used for a student whose school performance is noticeably affected by the time of day or day of the school week on which it is done. This accommodation may not be used specifically to change the order of administration of test sessions. This accommodation must not result in the administration of a test session to an individual student prior to the regularly scheduled administration time for that session for all students.

S. Setting		
Code	Tests were administered	Details on Delivery of Accommodations
S1	in a separate location within the school by trained school personnel.	A student or students may be tested individually or in small groups in an alternative site within the school to reduce distractions for themselves or others, or to increase physical access to special equipment.
S2	in an out-of-school setting by trained school personnel.	Out-of-school testing may be used for a student who is hospitalized or tutored because they are unable to attend school. The test must be administered by trained school personnel familiar with test administration procedures and guidelines. Relatives/guardians of the student may not be used as the test administrator.

P. Presentation		
Code	Tests were administered	Details on Delivery of Accommodations
P1	individually.	Individual or small group testing may be used to minimize distractions for a student or students whose test is administered out of the classroom or so that others will not be distracted by other accommodations being used (e.g., dictation)
P2	in a small group.	
P3	with test and directions read aloud in English or signed to the student. (**NOT allowed for the Reading test.**)	A reader may be used for a student whose inability to read would hinder performance on the Mathematics, Science, or Writing test. Words must be read as written. Guidelines for reading mathematical symbols must be followed. No translations (with the exception of signed language) or explanations are allowed. Trained personnel may use sign language to administer the test.
P4	with only test directions read aloud or signed to the student.	A reader may be used for a student whose inability to read or locate directions would hinder performance on the test. Note that most directions on the NECAP test occur at the beginning of the test session and are already read aloud by the test administrator. Guidelines for what are and are not "test directions" must be followed. With the exception of sign language and the case of students enrolled in a program where the test administrator routinely presents information in a foreign language, directions may not be translated.
P5	with administrator verification of student understanding following the reading of test directions.	After test directions have been read, the test administrator may ask the student to explain what he/she has been asked to do. If directions have been misunderstood by the student, the test directions may be paraphrased or demonstrated. Test items MUST NOT be paraphrased or explained.
P6	using alternative or assistive technology that is part of the student's communication system.	The test may be presented through his/her regular communication system to a student who uses alternative or assistive technology on a daily basis. Technology may not be used to "read" the Reading test to the student.
P7	by trained school personnel known to the student other than the student's classroom teacher.	A student may be more comfortable with a test administrator who works with the student on a regular basis, but is not the student's regular teacher for the general curriculum or other staff assigned as test administrator. All test administrators must be trained school personnel familiar with test administration and accommodations procedures and guidelines.
P8	using a large-print version of assessment.	Both large-print and Braille versions of the assessment require special preparation and processing and must be pre-ordered. Directions for ordering these materials are included in communications sent to school principals prior to the test.
P9	using Braille version of assessment.	
P10	using a word-to-word translation dictionary for ELL students. (**NOT allowed for the Reading test.**)	A student with limited English proficiency may have a word-to-word dictionary available for individual use as needed. A word-to-word dictionary is one that does not include any definitions. Information on acceptable dictionaries is provided on the departments' websites.
P11	using visual or auditory supports.	The test may be presented using visual aids such as visual magnification devices, reduction of visual print by blocking or other techniques, or acetate shields; or auditory devices such as special acoustics, amplification, noise buffers, whisper phones, or calming music.

R. Response		
Code	Tests were administered	Details on Delivery of Accommodations
R1	with a student dictating responses to school personnel. (**NOT allowed for the Writing test.** See O2 – using a scribe for the Writing test.)	A student may dictate answers to constructed-response or short-answer questions to locally trained personnel or record oral answers in an individual setting so that other students will not benefit by hearing answers or be otherwise disturbed. Policies regarding recorded answers must be followed prior to returning test materials.
R2	with a student dictating responses using alternative or assistive technology/devices that are part of the student's communication system. (**NOT allowed for the Writing test.** See O2 – using a scribe for the Writing test.)	Technology is used to permit a student to respond to the test. When using a computer, word processing device, or other assistive technology, spell and grammar checks must be turned off. Policies regarding recorded answers must be followed prior to returning test materials.
R3	with a student using approved tools or devices to minimize distractions.	Noise buffers, place markers, carrels, etc. may be used to minimize distractions for the student. This accommodation does NOT include assistive devices such as templates, graphic organizers, or other devices intended specifically to help students organize thinking or develop a strategy for a specific question.
R4	with a student writing responses using separate paper, a word processer, computer, brailler, or similar device.	A student may use technological or other tools (e.g., large-spaced paper) to write responses to constructed-response, short-answer, and extended response items. A key distinction between this accommodation and R2 is that the student using this accommodation is responding in writing rather than dictating. When using a computer, word processing device, or other assistive technology, spell and grammar checks must be turned off, as well as access to the Web. This accommodation is intended for unique individual needs, not an entire class. Policies regarding recorded answers must be followed prior to returning test materials.
R5	with a student indicating responses to multiple-choice items to school personnel.	A student unable to write or otherwise unable to fill-in answers to multiple-choice questions may indicate a response to trained school personnel. The school personnel records the student's response in the student answer booklet.
R6	with a student responding with the use of visual aids.	Visual aids include any optical or non-optical devices used to enhance visual capability. Examples include magnifiers, special lighting, markers, filters, large-spaced paper, color overlays, etc. An abacus may also be used for student with severe visual impairment or blindness on the Mathematics and Science tests. Note that the use of this accommodation still requires student responses to be recorded in a student answer booklet.
R7	with a student with limited English proficiency responding with use of a word-to-word dictionary. (**NOT allowed for the Reading test.**)	A student with limited English proficiency may have a word-to-word dictionary available for individual use as needed when responding. A word-to-word dictionary is one that does not include any definitions. Information on acceptable dictionaries is provided on each Department's website.

O. Other These accommodations require DOE approval.		
Code	Tests were administered	Details on Delivery of Accommodations
O1	using other accommodation(s) not on this list, requested by the accommodations team.	An IEP team or other appropriate accommodation team may request that a student be provided an accommodation not included on this standard list of accommodations. Like all other accommodations, these should be consistent with the student's normal routine during instruction and/or assessment. Requests should be made to the DOE when accommodation plans are being made for a student prior to testing. DOE approval must be received for the requested accommodation to be coded as an O1 accommodation. Non-approved accommodations used during test administration will be coded as an M3 modification.
O2	with a scribe used on the Writing test.	The use of a scribe for students dictating a response to the Writing test may only be used under limited circumstances and must be approved by the DOE. When approved as an accommodation, the scribe must follow established guidelines and procedures.

M. Modifications All modifications result in impacted items being scored as incorrect.		
Code	Tests were administered	Details on Delivery of Accommodations
M1	using a 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	Inappropriate use of a calculator or other tools will result in impacted items being scored as incorrect.
M2	with the test administrator reading the Reading test.	The read aloud accommodation (P3) is not allowed for the Reading test. If it is used, all reading items in the sessions that are read aloud will be scored as incorrect.
M3	using an accommodation on this list not approved for a particular test or an accommodation not included on this list without prior approval of the DOE.	Inappropriate use of an accommodation included on this list or use of another accommodation without prior approval of the DOE will result in impacted items being scored as incorrect.

Note: English Language Learners may qualify for any of the accommodations listed as appropriate and determined by a team. Refer to the *NECAP Accommodations Guide* for additional information.

APPENDIX E—ITEM-LEVEL CLASSICAL STATISTICS

Table E-1. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 3

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
119678	MC	0.78	0.51	1
119701	SA	0.79	0.30	1
119704	MC	0.50	0.41	1
119712	MC	0.79	0.39	1
119716	MC	0.87	0.44	1
119720	MC	0.72	0.52	1
119721	MC	0.79	0.46	1
119731	SA	0.83	0.48	1
119747	SA	0.69	0.50	2
119776	MC	0.51	0.23	1
119780	SA	0.44	0.52	2
119784	SA	0.81	0.29	1
119808	MC	0.83	0.44	1
119811	MC	0.88	0.43	0
119825	MC	0.71	0.44	1
119840	MC	0.74	0.44	1
119852	SA	0.35	0.47	1
119862	SA	0.49	0.57	1
119891	MC	0.46	0.23	1
119920	SA	0.64	0.59	1
122953	MC	0.38	0.28	2
123045	MC	0.94	0.37	0
124305	SA	0.76	0.40	2
124334	MC	0.81	0.48	1
124370	MC	0.88	0.41	0
124433	MC	0.86	0.40	1
124436	MC	0.79	0.48	2
139552	MC	0.83	0.45	1
139569	MC	0.89	0.43	1
139658	MC	0.55	0.40	2

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
139659	MC	0.62	0.33	1
139665	MC	0.54	0.36	2
139695	SA	0.59	0.45	1
144607	MC	0.63	0.40	1
145250	SA	0.60	0.48	1
145252	SA	0.74	0.47	1
145505	SA	0.25	0.47	2
145674	MC	0.87	0.21	1
145733	MC	0.71	0.48	1
168355	MC	0.79	0.43	2
168364	MC	0.69	0.45	1
168409	SA	0.68	0.39	1
168416	SA	0.55	0.52	1
168582	MC	0.82	0.30	1
168752	MC	0.50	0.46	1
168754	MC	0.46	0.48	4
203056	SA	0.57	0.62	2
203063	SA	0.33	0.44	1
203099	MC	0.74	0.46	1
203129	SA	0.63	0.36	1
203182	MC	0.34	0.41	2
203198	SA	0.39	0.48	1
203218	MC	0.44	0.48	1
203221	MC	0.62	0.50	2
215701	SA	0.21	0.44	1

Table E-2. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 4

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
120033	MC	0.89	0.39	0
120045	MC	0.54	0.32	1
120071	SA	0.59	0.53	1
120083	SA	0.69	0.48	1
120108	MC	0.72	0.28	1
120131	SA	0.77	0.46	0
120159	SA	0.70	0.47	1
120173	MC	0.61	0.40	2
120178	MC	0.94	0.34	1
120187	MC	0.71	0.43	2
120193	MC	0.39	0.30	1
120199	SA	0.46	0.33	1
120203	SA	0.61	0.52	1
120214	MC	0.86	0.45	1
120221	SA	0.77	0.38	1
120222	SA	0.51	0.55	1
120251	MC	0.89	0.43	0
120255	MC	0.71	0.50	1
120272	MC	0.69	0.40	1
120284	SA	0.65	0.50	1
123105	MC	0.89	0.37	0
124560	MC	0.73	0.45	1
124620	MC	0.92	0.35	0
124723	SA	0.66	0.55	1
124737	SA	0.63	0.53	1
139437	MC	0.83	0.39	1
139450	SA	0.39	0.54	1
139466	MC	0.68	0.39	1
139487	SA	0.48	0.50	0
139491	MC	0.71	0.47	1
139497	MC	0.51	0.42	1
139511	MC	0.94	0.36	0
139527	SA	0.68	0.53	1
144634	MC	0.56	0.42	1
144646	MC	0.73	0.39	1
144648	MC	0.66	0.35	0
145085	SA	0.77	0.46	1
145538	MC	0.45	0.40	2
145539	MC	0.88	0.44	1
145802	MC	0.73	0.30	0
145821	MC	0.71	0.48	1
145848	MC	0.66	0.37	0
168431	MC	0.48	0.31	1
168457	SA	0.56	0.30	0
169097	MC	0.54	0.36	1
169107	MC	0.44	0.47	1
169111	SA	0.40	0.49	1
201329	MC	0.76	0.50	0
201331	MC	0.60	0.35	1
201339	MC	0.65	0.49	1
201356	MC	0.41	0.35	1
201358	MC	0.34	0.38	1
201364	SA	0.57	0.62	1
201383	SA	0.42	0.41	0
201425	SA	0.43	0.27	1

Table E-3. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 5

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
120626	MC	0.55	0.43	0
120630	MC	0.80	0.44	1
120639	SA	0.64	0.55	1
120654	MC	0.72	0.39	0
120657	MC	0.61	0.54	1
120659	SA	0.69	0.44	0
120678	MC	0.57	0.41	0
120682	MC	0.68	0.39	0
120683	MC	0.48	0.37	0
120694	SA	0.41	0.36	2
120702	MC	0.55	0.40	0
120720	MC	0.65	0.47	0
120746	SA	0.36	0.57	2
120749	MC	0.54	0.43	0
120762	MC	0.60	0.34	0
120783	SA	0.55	0.40	0
120791	SA	0.53	0.60	1
120801	MC	0.55	0.55	0
120818	MC	0.58	0.41	0
120831	MC	0.70	0.46	2
120850	SA	0.60	0.53	1
120857	CR	0.56	0.64	1
124040	MC	0.57	0.40	1
124804	SA	0.67	0.47	0
124806	MC	0.61	0.46	1
124937	MC	0.61	0.37	0

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
139344	MC	0.64	0.43	1
144664	CR	0.51	0.74	1
145102	SA	0.59	0.49	1
145112	MC	0.35	0.22	1
145294	MC	0.71	0.40	0
145571	SA	0.48	0.55	2
145584	MC	0.56	0.46	0
145597	MC	0.49	0.41	1
145870	MC	0.70	0.36	0
145871	MC	0.67	0.53	1
145890	MC	0.87	0.36	0
145942	CR	0.50	0.55	1
167675	MC	0.74	0.43	1
167743	SA	0.47	0.62	1
167794	CR	0.50	0.60	1
167797	MC	0.66	0.30	0
167859	MC	0.53	0.47	0
167878	SA	0.62	0.43	1
204245	MC	0.82	0.47	0
204285	MC	0.53	0.42	0
204683	MC	0.70	0.49	0
217161	MC	0.48	0.34	0

Table E-4. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 6

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
119175	MC	0.63	0.40	0
119203	SA	0.53	0.52	1
119205	SA	0.37	0.55	2
119216	MC	0.67	0.47	0
119219	MC	0.67	0.50	0
119221	MC	0.55	0.39	0
119227	SA	0.60	0.46	1
119229	SA	0.54	0.62	1
119231	MC	0.61	0.31	1
119232	MC	0.74	0.25	0
119248	SA	0.72	0.46	1
119263	SA	0.31	0.31	1
119281	SA	0.59	0.56	2
119283	MC	0.55	0.47	0
119299	MC	0.55	0.41	0
119311	MC	0.54	0.36	0
119316	MC	0.63	0.52	0
119319	SA	0.69	0.50	0
119326	MC	0.48	0.40	0
119341	SA	0.45	0.46	2
119377	MC	0.54	0.42	2
119385	MC	0.68	0.47	0
119393	MC	0.49	0.29	0
124983	MC	0.53	0.54	1
125025	MC	0.65	0.39	0
125128	SA	0.51	0.61	2

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
125135	MC	0.59	0.38	0
139269	MC	0.53	0.29	0
139274	MC	0.61	0.46	0
139336	MC	0.66	0.43	0
139370	SA	0.75	0.38	1
145124	MC	0.59	0.49	0
145130	MC	0.73	0.44	0
145346	MC	0.70	0.45	0
145606	MC	0.78	0.32	1
145978	MC	0.47	0.33	0
167684	MC	0.56	0.42	0
167709	MC	0.60	0.48	1
167831	CR	0.31	0.68	1
167910	MC	0.44	0.30	0
167914	MC	0.39	0.41	1
167982	CR	0.50	0.68	1
202497	MC	0.58	0.37	0
202514	MC	0.60	0.44	0
202530	MC	0.57	0.49	1
202623	SA	0.68	0.56	1
202642	CR	0.47	0.63	1
202876	CR	0.50	0.68	1

Table E-5. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 7

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
119242	MC	0.52	0.52	0
119378	MC	0.80	0.39	0
120334	MC	0.66	0.42	0
120351	MC	0.64	0.47	1
120400	MC	0.52	0.42	1
120413	SA	0.31	0.39	3
120419	MC	0.65	0.41	0
120421	MC	0.72	0.46	0
120440	MC	0.56	0.46	1
120446	MC	0.49	0.50	0
120458	CR	0.29	0.64	1
120462	MC	0.55	0.47	0
120468	SA	0.40	0.47	2
120496	MC	0.46	0.22	0
120511	MC	0.74	0.37	1
120514	MC	0.55	0.40	0
125286	MC	0.70	0.29	0
125310	MC	0.76	0.39	0
125344	CR	0.29	0.56	2
125349	MC	0.45	0.33	1
139925	MC	0.54	0.47	0
140280	SA	0.68	0.39	1
140292	CR	0.37	0.70	1
144729	MC	0.60	0.58	0
145144	SA	0.39	0.59	2
145151	MC	0.43	0.36	0
145157	SA	0.62	0.50	1
145166	MC	0.73	0.48	0
145364	MC	0.55	0.45	1
145370	MC	0.69	0.33	0
145375	MC	0.66	0.42	1
146126	MC	0.51	0.46	0
150893	SA	0.23	0.50	4
169224	SA	0.65	0.53	1
169226	CR	0.37	0.67	2
169486	SA	0.52	0.62	1
169533	SA	0.44	0.50	1
204567	SA	0.36	0.55	2
204667	SA	0.52	0.68	1
204888	MC	0.45	0.48	0
204917	MC	0.66	0.42	0
204925	SA	0.43	0.41	1
204932	MC	0.65	0.47	0
204939	MC	0.75	0.39	1
204998	MC	0.48	0.45	0
205033	MC	0.63	0.42	1
205049	MC	0.72	0.34	0
205054	MC	0.48	0.54	0

Table E-6. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 8

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
120876	SA	0.57	0.63	2
120879	MC	0.56	0.38	0
120894	MC	0.57	0.42	0
120910	MC	0.44	0.33	1
120932	SA	0.77	0.37	1
120936	SA	0.59	0.64	3
120950	MC	0.37	0.32	1
120961	MC	0.58	0.41	0
120969	MC	0.77	0.47	0
120971	MC	0.78	0.47	0
120998	MC	0.43	0.49	0
121002	CR	0.37	0.69	1
121016	CR	0.36	0.71	2
121020	MC	0.53	0.34	1
121024	MC	0.82	0.37	0
121025	MC	0.75	0.47	0
121031	SA	0.58	0.67	2
121033	MC	0.55	0.50	0
121040	SA	0.69	0.30	1
121056	MC	0.60	0.31	0
121061	SA	0.23	0.44	2
122522	MC	0.63	0.45	1
125498	SA	0.53	0.58	1
125528	CR	0.63	0.71	1
125542	MC	0.69	0.55	1
125592	MC	0.73	0.30	0

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
139756	MC	0.60	0.52	0
139782	MC	0.40	0.23	0
139845	MC	0.71	0.25	0
139848	MC	0.52	0.39	1
139865	SA	0.68	0.55	1
139885	SA	0.57	0.53	1
139894	CR	0.41	0.67	1
145197	MC	0.70	0.50	1
145386	MC	0.63	0.27	0
145393	SA	0.16	0.48	3
145641	MC	0.78	0.41	0
145643	MC	0.40	0.35	0
146483	SA	0.50	0.63	2
169368	MC	0.44	0.43	0
169404	MC	0.36	0.26	1
169406	MC	0.61	0.48	0
169411	MC	0.66	0.37	0
204961	SA	0.42	0.58	2
205021	MC	0.62	0.44	0
205045	MC	0.52	0.25	1
205064	MC	0.55	0.32	0
205073	MC	0.80	0.44	1

Table E-7. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Mathematics Grade 11

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
119419	MC	0.82	0.31	1
119437	MC	0.59	0.45	1
119460	CR	0.22	0.72	9
119464	SA	0.40	0.62	4
119489	CR	0.24	0.72	7
119553	SA	0.41	0.56	5
119562	SA	0.29	0.54	6
119573	MC	0.79	0.36	0
119576	MC	0.57	0.31	1
119603	SA	0.23	0.50	5
119620	SA	0.19	0.45	5
119623	CR	0.28	0.70	7
119625	MC	0.59	0.52	1
125758	MC	0.34	0.37	1
125769	MC	0.43	0.33	1
125790	SA	0.20	0.56	8
125822	SA	0.26	0.58	11
125854	SA	0.24	0.57	8
140013	MC	0.35	0.46	1
140033	SA	0.23	0.59	12
140042	MC	0.54	0.56	1
140086	MC	0.87	0.35	0
140089	SA	0.31	0.63	7
140181	MC	0.60	0.53	1
141322	CR	0.49	0.71	5

Item		*Difficulty*	*Discrimination*	*Percent Omitted*
Number	*Type*			
144783	MC	0.33	0.25	1
144797	MC	0.24	0.29	1
144820	MC	0.77	0.45	1
145230	MC	0.21	0.39	1
145233	MC	0.56	0.51	0
145234	SA	0.36	0.53	5
145461	SA	0.58	0.52	2
145469	MC	0.69	0.25	1
169559	MC	0.55	0.41	1
169568	MC	0.54	0.60	1
169577	MC	0.43	0.59	1
169588	SA	0.48	0.42	3
169595	SA	0.52	0.66	9
169711	SA	0.33	0.61	7
169792	SA	0.37	0.63	9
205745	MC	0.40	0.44	1
205785	MC	0.70	0.46	0
205802	MC	0.49	0.42	0
205985	SA	0.34	0.59	5
206305	MC	0.55	0.43	0
210417	SA	0.57	0.59	2

Table E-8. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 3

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
117787	MC	0.70	0.34	1
117789	MC	0.79	0.48	1
117790	MC	0.69	0.29	2
117793	MC	0.75	0.35	1
117794	MC	0.69	0.27	1
117796	MC	0.45	0.38	2
117797	MC	0.55	0.38	2
117798	MC	0.79	0.43	1
117801	CR	0.71	0.48	1
117802	CR	0.82	0.39	0
148224	MC	0.85	0.40	1
171325	MC	0.77	0.42	0
171327	MC	0.82	0.49	0
171329	MC	0.80	0.42	1
171334	MC	0.70	0.31	1
171335	CR	0.36	0.50	1
171508	MC	0.46	0.24	2
171511	MC	0.84	0.48	0
171515	MC	0.84	0.48	1
171525	MC	0.71	0.21	1
171554	CR	0.34	0.50	1
171719	MC	0.69	0.44	0
171725	MC	0.56	0.34	1
171765	MC	0.76	0.45	1
208039	MC	0.68	0.40	1
208046	MC	0.58	0.44	1
208067	MC	0.78	0.37	2
208072	MC	0.54	0.50	2
208076	MC	0.54	0.38	1
208086	MC	0.56	0.34	2
208088	MC	0.64	0.45	2
208093	MC	0.71	0.49	3
208096	CR	0.82	0.48	1
208103	CR	0.35	0.63	2

Table E-9. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 4

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
118014	MC	0.86	0.57	0
118015	MC	0.50	0.35	1
118017	MC	0.59	0.46	1
118019	MC	0.74	0.55	1
118020	MC	0.67	0.48	1
118021	MC	0.65	0.40	1
118022	MC	0.76	0.41	1
118023	MC	0.80	0.50	2
118026	CR	0.70	0.54	1
118028	CR	0.90	0.47	0
171904	MC	0.86	0.37	0
171908	MC	0.58	0.38	0
171911	MC	0.80	0.43	0
171970	MC	0.80	0.42	1

continued

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
171975	CR	0.55	0.63	1
172219	MC	0.88	0.43	2
208048	MC	0.76	0.49	0
208052	MC	0.89	0.43	0
208061	MC	0.66	0.15	1
208065	MC	0.86	0.33	1
208080	CR	0.39	0.52	1
208641	MC	0.66	0.53	0
208667	MC	0.42	0.33	1
208672	MC	0.36	0.20	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
208683	MC	0.66	0.28	2
208686	MC	0.72	0.34	1
208701	MC	0.71	0.51	1
208713	MC	0.41	0.23	1
208714	MC	0.65	0.49	2
208727	CR	0.80	0.53	0
208740	CR	0.40	0.48	1
208849	MC	0.87	0.41	0
208860	MC	0.78	0.44	0
208865	MC	0.74	0.50	1

Table E-10. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 5

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
118120	MC	0.64	0.38	0
149127	MC	0.74	0.44	0
171822	MC	0.77	0.41	0
171825	MC	0.92	0.38	0
171837	MC	0.81	0.43	0
171842	MC	0.84	0.36	1
171848	CR	0.42	0.51	1
209325	MC	0.73	0.49	0
209352	MC	0.67	0.35	0
209359	MC	0.82	0.47	0
209363	MC	0.79	0.47	1
209368	MC	0.48	0.33	1
209370	MC	0.89	0.43	1
209372	MC	0.77	0.47	1
209377	MC	0.53	0.37	1
209380	CR	0.32	0.59	1
209383	CR	0.33	0.66	1
209539	MC	0.83	0.38	0

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
209546	MC	0.78	0.36	0
209551	MC	0.86	0.41	0
209552	MC	0.40	0.26	1
209558	CR	0.53	0.48	1
209591	MC	0.76	0.27	0
209593	MC	0.75	0.37	0
209595	MC	0.59	0.44	0
209599	MC	0.79	0.42	1
209616	MC	0.83	0.43	1
209617	MC	0.77	0.39	1
209618	MC	0.63	0.34	1
209631	MC	0.82	0.44	1
209634	CR	0.36	0.51	1
209640	CR	0.33	0.57	1
209700	MC	0.79	0.38	0
209705	MC	0.70	0.47	0

Table E-11. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 6

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
118387	MC	0.72	0.33	0
172111	MC	0.85	0.37	0
172129	MC	0.70	0.39	0
172131	MC	0.75	0.42	0
172134	MC	0.88	0.41	0
172140	CR	0.54	0.51	0
172242	MC	0.51	0.26	0
172264	MC	0.81	0.42	0
208944	MC	0.77	0.23	0
208965	MC	0.87	0.46	0
208967	MC	0.72	0.38	0
208972	MC	0.84	0.36	0
208974	MC	0.87	0.44	0
208981	MC	0.82	0.50	1
208986	MC	0.71	0.41	1
208987	MC	0.60	0.33	1
208989	MC	0.82	0.39	1
208991	CR	0.39	0.64	1

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
208995	CR	0.38	0.57	1
209118	MC	0.57	0.30	0
209119	MC	0.93	0.43	0
209127	MC	0.70	0.47	0
209137	MC	0.63	0.27	1
209142	MC	0.72	0.34	1
209143	MC	0.88	0.42	1
209146	MC	0.87	0.45	1
209148	MC	0.81	0.41	1
209151	CR	0.47	0.62	1
209153	CR	0.41	0.62	1
209267	MC	0.90	0.32	0
209284	MC	0.72	0.26	0
209288	MC	0.51	0.22	0
209293	MC	0.49	0.32	1
209299	CR	0.52	0.62	1

Table E-12. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 7

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
147203	MC	0.89	0.32	0
147210	MC	0.74	0.26	0
147213	MC	0.79	0.36	1
147215	MC	0.74	0.27	0
147224	CR	0.47	0.61	1

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
147623	MC	0.81	0.43	0
147625	MC	0.74	0.38	0
147626	MC	0.70	0.27	0
147633	MC	0.60	0.30	1
147637	CR	0.55	0.57	1

continued

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
147686	MC	0.56	0.40	0
173246	MC	0.76	0.31	0
173247	MC	0.70	0.25	0
173249	MC	0.70	0.43	0
173254	MC	0.84	0.46	1
173255	MC	0.57	0.31	1
173260	MC	0.68	0.32	1
173261	MC	0.79	0.51	1
173269	MC	0.65	0.32	1
173286	CR	0.49	0.67	1
173293	CR	0.49	0.62	1
173338	MC	0.71	0.23	0

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
207521	MC	0.63	0.32	0
207560	MC	0.74	0.41	0
207625	MC	0.92	0.28	0
207630	MC	0.82	0.33	0
207633	MC	0.86	0.37	0
207636	MC	0.87	0.41	0
207640	MC	0.66	0.30	0
207643	MC	0.70	0.36	1
207644	MC	0.77	0.40	0
207647	MC	0.91	0.36	0
207649	CR	0.60	0.66	1
207654	CR	0.45	0.65	1

Table E-13. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 8

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
118749	MC	0.67	0.20	0
172353	MC	0.92	0.34	0
172356	MC	0.79	0.36	0
172359	MC	0.83	0.32	0
172361	MC	0.82	0.38	0
172373	CR	0.53	0.62	0
172713	MC	0.81	0.29	0
207706	MC	0.68	0.38	0
207708	MC	0.65	0.36	0
207710	MC	0.81	0.44	0
207711	MC	0.78	0.36	0
207715	MC	0.81	0.41	0
207717	MC	0.65	0.31	0
207742	MC	0.81	0.29	0
207746	MC	0.81	0.25	1

Item Number	Item Type	Difficulty	Discrimination	Percent Omitted
207754	CR	0.42	0.62	1
207757	CR	0.55	0.68	1
207763	MC	0.61	0.28	0
207769	MC	0.81	0.36	0
207772	MC	0.75	0.35	0
207777	MC	0.70	0.37	1
207782	CR	0.48	0.55	1
207829	MC	0.72	0.47	0
207831	MC	0.58	0.38	1
207832	MC	0.48	0.27	1
207833	MC	0.68	0.38	1
207835	MC	0.75	0.40	1
207840	MC	0.79	0.40	1
207844	MC	0.56	0.32	1
207850	MC	0.53	0.35	1

continued

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
207852	CR	0.42	0.68	1
207854	CR	0.44	0.66	1

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
208253	MC	0.73	0.36	0
208279	MC	0.77	0.28	0

Table E-14. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Reading Grade 11

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
147381	MC	0.63	0.42	1
147382	MC	0.76	0.46	1
147383	MC	0.76	0.45	1
147385	MC	0.79	0.52	1
147388	CR	0.50	0.68	3
147908	MC	0.72	0.29	0
172855	MC	0.67	0.36	1
208337	MC	0.87	0.32	0
208345	MC	0.68	0.30	0
208478	MC	0.77	0.41	1
208480	MC	0.76	0.38	1
208482	MC	0.86	0.31	0
208486	MC	0.79	0.45	0
208502	MC	0.74	0.46	0
208519	MC	0.76	0.51	0
208520	MC	0.84	0.42	0
208524	MC	0.86	0.39	0
208528	CR	0.57	0.67	1

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
208545	CR	0.53	0.66	2
208680	MC	0.64	0.33	1
208682	MC	0.82	0.52	1
208684	MC	0.68	0.28	1
208689	MC	0.87	0.34	1
208700	CR	0.49	0.65	2
208704	MC	0.90	0.26	1
208710	MC	0.80	0.46	1
208721	MC	0.75	0.40	1
208726	MC	0.76	0.38	1
208729	MC	0.81	0.43	1
208730	MC	0.82	0.42	1
208742	MC	0.79	0.43	1
208744	CR	0.54	0.65	2
208747	CR	0.53	0.69	3
226071	MC	0.46	0.25	1

Table E-15. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Writing Grade 5

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
148013	MC	0.61	0.24	1
148060	MC	0.77	0.32	0
148115	MC	0.65	0.30	0
148134	MC	0.83	0.27	0
148181	MC	0.87	0.37	0
148235	MC	0.88	0.34	0
148288	MC	0.78	0.31	0
148292	MC	0.88	0.35	0
148314	MC	0.86	0.36	0
148323	MC	0.81	0.30	0
150155	WP	0.40	0.60	0
150161	CR	0.53	0.64	1
150201	CR	0.57	0.59	1
150231	CR	0.58	0.61	1

Table E-16. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Writing Grade 8

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
148443	MC	0.77	0.28	0
148486	MC	0.74	0.32	0
148548	MC	0.47	0.24	1
148591	MC	0.82	0.31	0
148615	MC	0.51	0.33	0
148657	MC	0.75	0.37	0
148704	MC	0.58	0.24	0
148766	MC	0.68	0.30	0
148780	MC	0.69	0.30	0
148850	MC	0.87	0.28	0
150778	CR	0.57	0.65	2
150794	CR	0.64	0.66	1
150813	WP	0.56	0.66	0
150830	CR	0.63	0.64	1

Table E-17. 2012–13 NECAP: Item-Level Classical Test Theory Statistics
Writing Grade 11

Item		Difficulty	Discrimination	Percent Omitted
Number	Type			
211263	WP	0.51		

APPENDIX F—ITEM-LEVEL SCORE POINT DISTRIBUTIONS

Table F-1. 2012–13 NECAP: Item-Level Score Distributions for Constructed Response Items—Mathematics

Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
			0	1	2	3	4
3	119701	1	20.35	78.76			
	119731	1	16.13	82.79			
	119747	1	29.52	68.74			
	119780	2	30.20	47.59	20.29		
	119784	1	17.96	80.90			
	119852	1	63.28	35.32			
	119862	2	29.29	40.65	29.14		
	119920	2	25.22	18.69	54.78		
	124305	2	12.12	20.03	65.87		
	139695	2	6.83	66.47	25.69		
	145250	1	38.92	60.15			
	145252	1	25.00	74.14			
	145505	2	60.37	24.74	13.01		
	168409	1	31.44	67.61			
	168416	2	29.75	28.50	40.38		
	203056	2	26.29	29.96	42.01		
	203063	1	65.80	33.22			
	203129	1	36.13	62.94			
	203198	2	48.11	22.98	27.95		
	215701	2	64.34	26.28	8.13		
4	120071	1	40.25	58.69			
	120083	2	14.84	31.17	53.47		
	120131	1	22.97	76.62			
	120159	1	29.76	69.59			
	120199	1	52.87	45.69			
	120203	2	18.32	38.48	42.05		
	120221	1	22.32	76.87			
	120222	2	12.48	70.25	16.08		
	120284	2	19.80	29.58	49.85		
	124723	2	19.34	28.43	51.38		
	124737	1	36.03	63.36			
	139450	2	52.05	15.16	31.36		
	139487	2	22.29	58.10	19.14		
	139527	2	12.07	38.24	48.98		
	145085	1	22.48	76.72			
	168457	1	43.65	55.96			
	169111	1	59.38	40.08			
	201364	2	34.45	16.06	48.98		
	201383	1	57.50	42.24			
	201425	2	41.71	28.95	28.60		
5	120639	2	28.76	11.26	58.60		
	120659	1	30.80	68.76			
	120694	2	46.14	22.35	29.41		
	120746	2	45.91	31.07	20.90		
	120783	1	44.61	54.92			
	120791	2	38.20	16.15	44.78		

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
			0	1	2	3	4
5	120850	1	38.58	60.43			
	120857	4	7.32	14.42	47.65	6.89	23.02
	124804	1	32.98	66.62			
	144664	4	7.96	29.30	31.13	11.88	19.10
	145102	1	40.32	58.95			
	145571	2	41.54	16.29	40.05		
	145942	4	18.42	15.64	30.07	15.28	19.47
	167743	2	40.47	23.08	35.63		
	167794	4	22.63	4.44	40.46	9.75	21.60
	167878	1	37.07	61.92			
6	119203	2	23.10	44.74	30.87		
	119205	2	53.95	13.34	30.32		
	119227	1	39.11	60.37			
	119229	2	31.42	26.99	40.48		
	119248	1	27.32	72.04			
	119263	1	67.90	31.47			
	119281	2	20.60	36.32	41.21		
	119319	1	30.89	68.62			
	119341	2	36.43	33.45	28.54		
	125128	2	24.32	44.63	28.57		
	139370	1	24.19	75.06			
	167831	4	42.10	20.58	17.90	4.74	13.56
	167982	4	18.12	16.68	23.83	25.33	15.13
	202623	1	31.52	67.92			
	202642	4	30.66	8.03	27.73	7.08	25.60
	202876	4	16.59	16.78	23.50	30.14	11.76
7	120413	2	40.18	51.61	5.07		
	120458	4	48.28	18.15	8.44	12.77	11.03
	120468	1	57.92	40.10			
	125344	4	30.85	34.20	23.01	4.98	5.40
	140280	1	31.50	67.75			
	140292	4	18.24	38.27	27.40	3.56	11.28
	145144	2	43.66	31.63	22.96		
	145157	1	36.84	61.87			
	150893	2	69.26	8.12	18.95		
	169224	1	33.75	65.46			
	169226	4	28.84	27.98	17.80	11.24	12.59
	169486	2	31.41	29.35	37.82		
	169533	1	55.25	43.65			
	204567	2	55.99	11.49	30.57		
	204667	2	31.18	31.96	35.61		
	204925	1	56.53	42.83			
8	120876	2	20.29	42.14	35.61		
	120932	1	21.85	77.39			
	120936	2	31.05	13.95	52.07		
	121002	4	33.28	23.58	15.80	10.88	15.34
	121016	4	31.75	25.11	16.83	10.48	13.72
	121031	2	21.61	37.04	39.19		
	121040	1	29.98	69.13			

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
			0	1	2	3	4
8	121061	1	74.71	23.22			
	125498	1	45.76	52.78			
	125528	4	18.87	7.76	11.12	20.54	40.38
	139865	1	30.54	68.41			
	139885	2	25.02	32.22	41.32		
	139894	4	18.73	24.87	31.60	18.97	4.74
	145393	2	75.79	9.27	11.61		
	146483	2	24.55	47.70	26.07		
	204961	1	55.29	42.48			
11	119460	4	37.34	29.29	17.49	4.53	2.16
	119464	1	56.08	39.65			
	119489	4	54.47	4.74	18.03	6.44	8.88
	119553	1	54.63	40.80			
	119562	1	64.26	29.46			
	119603	1	72.30	23.11			
	119620	2	73.08	5.04	16.67		
	119623	4	31.71	23.19	28.30	6.82	2.48
	125790	2	63.46	16.50	11.66		
	125822	1	63.93	25.55			
	125854	1	68.08	24.25			
	140033	2	59.02	12.86	16.35		
	140089	1	61.89	30.86			
	141322	4	16.51	2.48	44.60	21.19	10.26
	145234	1	59.04	36.44			
	145461	1	39.32	58.36			
	169588	1	49.28	47.52			
	169595	2	34.25	9.98	47.09		
	169711	2	38.30	43.92	10.62		
	169792	2	51.44	5.57	34.31		
	205985	1	61.06	34.18			
	210417	1	41.72	56.56			

Table F-2. 2012–13 NECAP: Item-Level Score Distributions for Constructed Response Items—Reading

Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
			0	1	2	3	4
3	117801	4	3.37	5.03	15.76	54.17	20.76
	117802	4	0.49	0.81	10.82	43.50	43.92
	171335	4	16.42	35.70	33.87	11.18	1.59
	171554	4	13.18	45.72	31.78	7.15	1.21
	208096	4	2.34	9.25	10.74	11.28	65.74
	208103	4	23.19	27.82	33.43	11.57	2.35
4	118026	4	6.57	10.14	14.49	29.42	38.43
	118028	4	0.39	1.69	6.49	19.11	71.95
	171975	4	11.95	11.19	30.47	32.30	13.21
	208080	4	10.26	34.78	44.70	8.47	1.11
	208727	4	0.82	6.36	16.24	24.09	52.02
	208740	4	8.14	37.29	38.96	12.24	2.33

continued

Grade	Item Number	Total Possible Points	Percent of Students at Score Point				
			0	1	2	3	4
5	171848	4	12.56	32.69	31.22	18.07	4.74
	209380	4	18.23	44.25	27.98	7.30	1.26
	209383	4	15.91	43.16	30.33	7.78	1.39
	209558	4	5.58	20.40	37.92	24.38	11.04
	209634	4	5.70	49.52	35.09	7.56	0.75
	209640	4	17.15	42.44	29.55	8.22	1.62
6	172140	4	3.01	18.37	43.93	27.38	6.90
	208991	4	12.00	34.77	36.88	12.94	2.14
	208995	4	11.40	38.28	37.41	10.11	2.01
	209151	4	4.84	24.60	49.38	18.13	2.45
	209153	4	12.01	26.66	44.79	13.40	1.94
	209299	4	3.46	15.14	55.71	20.71	4.46
7	147224	4	11.05	23.10	38.45	20.04	6.83
	147637	4	2.58	19.83	41.99	25.26	9.78
	173286	4	7.58	22.65	40.80	20.17	7.56
	173293	4	22.83	14.73	19.50	24.61	17.10
	207649	4	3.09	12.11	38.92	31.76	13.58
	207654	4	11.67	22.97	40.19	17.77	6.28
8	172373	4	5.20	21.08	38.83	26.53	7.94
	207754	4	14.50	28.03	34.59	17.28	4.41
	207757	4	5.00	17.92	36.00	30.08	10.37
	207782	4	7.87	25.55	38.65	21.76	5.56
	207852	4	10.11	32.81	37.60	14.76	3.71
	207854	4	11.31	28.25	35.87	17.72	5.41
11	147388	4	5.23	18.09	46.08	21.83	6.06
	208528	4	4.06	14.48	36.62	33.15	10.50
	208545	4	10.77	13.60	34.97	28.01	11.14
	208700	4	6.81	20.26	44.46	21.34	5.28
	208744	4	3.66	16.28	44.62	26.20	7.59
	208747	4	5.64	12.45	45.08	26.41	7.75

Table F-3. Item-Level Score Distributions for Constructed–Response Items by Grade—Writing

Grade	Item Number	Total Possible Points	Percent of Students at Score Point												
			0	1	2	3	4	5	6	7	8	9	10	11	12
5	150155	12	1.32		8.93	8.72	27.29	17.72	21.69	7.88	4.33	1.38	0.47	0.05	0.01
	150161	4	1.86	20.83	45.96	24.55	6.22								
	150201	4	2.82	14.56	42.84	30.95	8.32								
	150231	4	2.70	11.25	41.86	36.79	6.51								
8	150778	4	1.23	15.09	41.73	30.05	10.14								
	150794	4	0.65	9.50	37.80	36.04	15.46								
	150813	12	0.60		2.74	2.29	6.83	7.46	23.54	17.43	21.51	11.02	5.62	0.44	0.09
	150830	4	1.02	10.03	35.36	39.82	13.23								
11	211263	12	0.92		5.00	3.26	12.08	10.15	23.48	16.79	18.44	6.63	2.63	0.44	0.15

APPENDIX G—DIFFERENTIAL ITEM FUNCTIONING RESULTS

Table G-1. 2012–13 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored—Mathematics

Grade	*Group*		*Item Type*	*Number of Items*	*Number "Low"*			*Number "High"*		
	Reference	*Focal*			*Total*	*Favoring*		*Total*	*Favoring*	
						Reference	*Focal*		*Reference*	*Focal*
3	Male	Female	MC	35	0	0	0	0	0	0
			SA	20	1	1	0	0	0	0
	No Disability	Disability	MC	35	4	4	0	0	0	0
			SA	20	3	2	1	0	0	0
	Non-EconDis	EconDis	MC	35	0	0	0	0	0	0
			SA	20	0	0	0	0	0	0
	Non-LEP	LEP	MC	35	4	3	1	0	0	0
			SA	20	0	0	0	0	0	0
	White	Asian	MC	35	3	3	0	0	0	0
			SA	20	1	0	1	0	0	0
		Black	MC	35	0	0	0	0	0	0
			SA	20	1	1	0	0	0	0
		Hispanic	MC	35	3	3	0	0	0	0
			SA	20	1	1	0	0	0	0
4	Male	Female	MC	35	4	4	0	0	0	0
			SA	20	2	0	2	0	0	0
	No Disability	Disability	MC	35	3	3	0	0	0	0
			SA	20	1	1	0	0	0	0
	Non-EconDis	EconDis	MC	35	1	1	0	0	0	0
			SA	20	1	1	0	0	0	0
	Non-LEP	LEP	MC	35	3	1	2	0	0	0
			SA	20	2	1	1	1	1	0
	White	Asian	MC	35	6	3	3	0	0	0
			SA	20	3	2	1	0	0	0
		Black	MC	35	4	2	2	1	1	0
			SA	20	4	3	1	1	1	0
		Hispanic	MC	35	5	3	2	0	0	0
			SA	20	3	2	1	1	1	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring		Total	Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
5	Male	Female	CR	4	0	0	0	0	0	0
			MC	32	2	1	1	0	0	0
			SA	12	0	0	0	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
			MC	32	5	4	1	0	0	0
			SA	12	0	0	0	0	0	0
	Non-EconDis	EconDis	CR	4	0	0	0	0	0	0
			MC	32	1	1	0	0	0	0
			SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	1	0	1	0	0	0
			MC	32	2	2	0	1	1	0
			SA	12	0	0	0	0	0	0
	White	Asian	CR	4	1	0	1	0	0	0
			MC	32	2	2	0	0	0	0
			SA	12	1	1	0	0	0	0
		Black	CR	4	0	0	0	0	0	0
			MC	32	2	2	0	0	0	0
			SA	12	0	0	0	0	0	0
		Hispanic	CR	4	0	0	0	0	0	0
			MC	32	3	2	1	0	0	0
			SA	12	0	0	0	0	0	0
6	Male	Female	CR	4	0	0	0	0	0	0
			MC	32	7	5	2	0	0	0
			SA	12	0	0	0	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
			MC	32	1	1	0	0	0	0
			SA	12	1	1	0	0	0	0
	Non-EconDis	EconDis	CR	4	0	0	0	0	0	0
			MC	32	0	0	0	0	0	0
			SA	12	0	0	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
6	Non-LEP	LEP	CR	4	1	1	0	0	0	0
			MC	32	3	2	1	0	0	0
			SA	12	2	2	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
			MC	32	4	3	1	0	0	0
			SA	12	1	1	0	0	0	0
		Black	CR	4	0	0	0	0	0	0
			MC	32	5	4	1	0	0	0
			SA	12	4	3	1	0	0	0
		Hispanic	CR	4	1	1	0	0	0	0
			MC	32	6	6	0	0	0	0
			SA	12	2	2	0	0	0	0
7	Male	Female	CR	4	0	0	0	0	0	0
			MC	32	7	4	3	0	0	0
			SA	12	3	1	2	0	0	0
	No Disability	Disability	CR	4	0	0	0	0	0	0
			MC	32	2	2	0	0	0	0
			SA	12	2	2	0	0	0	0
	Non-EconDis	EconDis	CR	4	0	0	0	0	0	0
			MC	32	0	0	0	0	0	0
			SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
			MC	32	7	6	1	1	1	0
			SA	12	1	1	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
			MC	32	3	3	0	0	0	0
			SA	12	2	1	1	0	0	0
		Black	CR	4	0	0	0	0	0	0
			MC	32	4	3	1	2	2	0
			SA	12	1	1	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
7	White	Hispanic	CR	4	0	0	0	0	0	0
			MC	32	4	3	1	1	1	0
			SA	12	2	1	1	0	0	0
8	Male	Female	CR	4	1	0	1	0	0	0
			MC	32	3	3	0	0	0	0
			SA	12	2	0	2	0	0	0
	No Disability	Disability	CR	4	1	1	0	0	0	0
			MC	32	1	1	0	0	0	0
			SA	12	1	1	0	0	0	0
	Non-EconDis	EconDis	CR	4	0	0	0	0	0	0
			MC	32	0	0	0	0	0	0
			SA	12	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	0	0	0	0	0	0
			MC	32	1	1	0	2	1	1
			SA	12	2	2	0	0	0	0
	White	Asian	CR	4	0	0	0	0	0	0
			MC	32	5	1	4	0	0	0
			SA	12	0	0	0	0	0	0
		Black	CR	4	0	0	0	0	0	0
			MC	32	2	0	2	1	1	0
			SA	12	2	2	0	0	0	0
		Hispanic	CR	4	0	0	0	0	0	0
			MC	32	1	1	0	0	0	0
			SA	12	0	0	0	0	0	0
11	Male	Female	CR	4	1	0	1	0	0	0
			MC	24	1	1	0	1	1	0
			SA	18	1	1	0	0	0	0
	No Disability	Disability	CR	4	1	1	0	0	0	0
			MC	24	2	1	1	0	0	0
			SA	18	2	2	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
11	Non-EconDis	EconDis	CR	4	0	0	0	0	0	0
			MC	24	0	0	0	0	0	0
			SA	18	0	0	0	0	0	0
	Non-LEP	LEP	CR	4	1	1	0	0	0	0
			MC	24	8	4	4	0	0	0
			SA	18	3	2	1	2	2	0
	White	Asian	CR	4	0	0	0	0	0	0
			MC	24	1	1	0	0	0	0
			SA	18	2	0	2	0	0	0
		Black	CR	4	0	0	0	0	0	0
			MC	24	1	1	0	0	0	0
			SA	18	0	0	0	0	0	0
		Hispanic	CR	4	0	0	0	0	0	0
			MC	24	1	1	0	0	0	0
			SA	18	1	1	0	0	0	0

Table G-2. 2012–13 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored—Reading

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
3	Male	Female	CR	6	0	0	0	0	0	0
			MC	28	0	0	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	2	2	0	0	0	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	0	0	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
			MC	28	2	2	0	2	2	0
	White	Asian	CR	6	0	0	0	0	0	0
			MC	28	3	2	1	1	0	1

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
3	White	Black	CR	6	0	0	0	0	0	0
			MC	28	3	3	0	0	0	0
		Hispanic	CR	6	0	0	0	0	0	0
			MC	28	3	3	0	1	1	0
4	Male	Female	CR	6	0	0	0	0	0	0
			MC	28	2	2	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	7	5	2	1	1	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	1	1	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
			MC	28	5	4	1	1	1	0
	White	Asian	CR	6	0	0	0	0	0	0
			MC	28	5	3	2	0	0	0
		Black	CR	6	0	0	0	0	0	0
			MC	28	4	4	0	1	1	0
		Hispanic	CR	6	0	0	0	0	0	0
			MC	28	3	2	1	1	1	0
5	Male	Female	CR	6	2	0	2	0	0	0
			MC	28	4	4	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	4	4	0	0	0	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	0	0	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
			MC	28	9	9	0	2	2	0
	White	Asian	CR	6	0	0	0	0	0	0
			MC	28	4	4	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
			MC	28	7	7	0	1	1	0
	White	Hispanic	CR	6	0	0	0	0	0	0
			MC	28	8	8	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring		Total	Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
6	Male	Female	CR	6	0	0	0	0	0	0
			MC	28	3	3	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	3	3	0	0	0	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	2	2	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
			MC	28	14	14	0	3	3	0
	White	Asian	CR	6	0	0	0	0	0	0
			MC	28	6	6	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
			MC	28	7	7	0	2	2	0
	White	Hispanic	CR	6	0	0	0	0	0	0
			MC	28	9	9	0	1	1	0
7	Male	Female	CR	6	3	0	3	0	0	0
			MC	28	3	3	0	0	0	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	6	6	0	0	0	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	2	2	0	0	0	0
	Non-LEP	LEP	CR	6	1	0	1	0	0	0
			MC	28	11	11	0	3	3	0
	White	Asian	CR	6	1	0	1	0	0	0
			MC	28	6	6	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
			MC	28	8	8	0	2	2	0
	White	Hispanic	CR	6	0	0	0	0	0	0
			MC	28	6	6	0	1	1	0
8	Male	Female	CR	6	4	0	4	0	0	0
			MC	28	5	5	0	2	2	0
	No Disability	Disability	CR	6	1	1	0	0	0	0
			MC	28	1	1	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
8	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	1	1	0	0	0	0
	Non-LEP	LEP	CR	6	2	0	2	0	0	0
			MC	28	13	13	0	5	5	0
	White	Asian	CR	6	1	0	1	0	0	0
			MC	28	4	4	0	0	0	0
	White	Black	CR	6	0	0	0	0	0	0
			MC	28	9	9	0	2	2	0
	White	Hispanic	CR	6	0	0	0	0	0	0
			MC	28	7	7	0	4	4	0
11	Male	Female	CR	6	1	0	1	0	0	0
			MC	28	6	5	1	2	2	0
	No Disability	Disability	CR	6	0	0	0	0	0	0
			MC	28	3	3	0	0	0	0
	Non-EconDis	EconDis	CR	6	0	0	0	0	0	0
			MC	28	4	4	0	0	0	0
	Non-LEP	LEP	CR	6	0	0	0	0	0	0
			MC	28	10	10	0	9	9	0
	White	Asian	CR	6	1	0	1	0	0	0
			MC	28	8	8	0	1	1	0
	White	Black	CR	6	0	0	0	0	0	0
			MC	28	8	8	0	1	1	0
	White	Hispanic	CR	6	1	0	1	0	0	0
			MC	28	7	7	0	3	3	0

Table G-3. 2012–13 NECAP: Number of Items Classified as "Low" or "High" DIF Overall and by Grade and Group Favored—Writing

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
					Total	Favoring		Total	Favoring	
	Reference	Focal				Reference	Focal		Reference	Focal
5	Male	Female	CR	3	0	0	0	0	0	0
			MC	10	0	0	0	0	0	0
			WP	1	0	0	0	0	0	0
	No Disability	Disability	CR	3	0	0	0	0	0	0
			MC	10	9	9	0	0	0	0
			WP	1	0	0	0	0	0	0
	Non-EconDis	EconDis	CR	3	0	0	0	0	0	0
			MC	10	0	0	0	0	0	0
			WP	1	0	0	0	0	0	0
	Non-LEP	LEP	CR	3	0	0	0	0	0	0
			MC	10	4	4	0	1	1	0
			WP	1	0	0	0	0	0	0
	White	Asian	CR	3	0	0	0	0	0	0
			MC	10	1	1	0	0	0	0
			WP	1	0	0	0	0	0	0
	White	Black	CR	3	0	0	0	0	0	0
			MC	10	3	3	0	0	0	0
			WP	1	0	0	0	0	0	0
	White	Hispanic	CR	3	0	0	0	0	0	0
			MC	10	4	4	0	0	0	0
			WP	1	0	0	0	0	0	0
8	Male	Female	CR	3	0	0	0	0	0	0
			MC	10	0	0	0	0	0	0
			WP	1	0	0	0	0	0	0
	No Disability	Disability	CR	3	0	0	0	0	0	0
			MC	10	7	7	0	2	2	0
			WP	1	0	0	0	0	0	0
	Non-EconDis	EconDis	CR	3	0	0	0	0	0	0
			MC	10	2	2	0	0	0	0
			WP	1	0	0	0	0	0	0

continued

Grade	Group		Item Type	Number of Items	Number "Low"			Number "High"		
	Reference	Focal			Total	Favoring		Total	Favoring	
						Reference	Focal		Reference	Focal
8	Non-LEP	LEP	CR	3	0	0	0	0	0	0
			MC	10	6	6	0	0	0	0
			WP	1	0	0	0	0	0	0
	White	Asian	CR	3	0	0	0	0	0	0
			MC	10	1	0	1	0	0	0
			WP	1	0	0	0	0	0	0
	White	Black	CR	3	0	0	0	0	0	0
			MC	10	2	2	0	1	1	0
			WP	1	0	0	0	0	0	0
	White	Hispanic	CR	3	0	0	0	0	0	0
			MC	10	3	3	0	0	0	0
			WP	1	0	0	0	0	0	0
11	Male	Female	WP	1	1	0	1	0	0	0
	No Disability	Disability	WP	1	0	0	0	1	1	0
	Non-EconDis	EconDis	WP	1	1	1	0	0	0	0
	Non-LEP	LEP	WP	1	0	0	0	1	1	0
	White	Asian	WP	1	0	0	0	0	0	0
	White	Black	WP	1	0	0	0	0	0	0
	White	Hispanic	WP	1	0	0	0	0	0	0

APPENDIX H—ITEM RESPONSE THEORY CALIBRATION RESULTS

Table H-1. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 3

IREF	*Parameters and Measures of Standard Error*					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
123045	1.26743	0.03086	-1.61354	0.04181	0.37153	0.02385
139569	0.99543	0.01809	-1.56402	0.03535	0.10389	0.02187
119678	1.13619	0.01951	-0.69927	0.02059	0.17786	0.01103
124433	0.75657	0.01015	-1.69300	0.01855	0.00000	0.00000
119811	1.09075	0.02221	-1.32200	0.03534	0.24731	0.01984
139552	0.87599	0.01523	-1.28469	0.03349	0.07922	0.01822
119720	0.91803	0.01273	-0.71061	0.01704	0.02692	0.00798
122953	0.71004	0.02205	1.24763	0.02055	0.16886	0.00673
144607	1.09555	0.02404	0.28231	0.01769	0.31652	0.00702
119716	0.93535	0.01566	-1.53698	0.03203	0.06437	0.01885
119712	0.87131	0.02058	-0.57534	0.03793	0.34013	0.01446
124334	1.02727	0.01862	-0.91959	0.02694	0.17538	0.01428
203099	0.84419	0.01512	-0.74267	0.02856	0.09993	0.01372
203221	0.97323	0.01627	-0.13796	0.01678	0.10642	0.00786
139659	0.68760	0.01956	0.21538	0.03657	0.27232	0.01200
119776	0.48754	0.02197	0.93732	0.05247	0.24445	0.01468
168752	1.41443	0.02455	0.45858	0.00968	0.17988	0.00456
119808	0.86698	0.01612	-1.24216	0.03715	0.11317	0.01974
203182	1.30320	0.02434	0.94621	0.00910	0.11439	0.00338
168355	0.80281	0.01598	-1.03546	0.04007	0.13979	0.01920
119704	0.87556	0.01842	0.46323	0.01729	0.15811	0.00703
119721	0.84627	0.01414	-1.08253	0.02968	0.06304	0.01521
124370	0.83753	0.01308	-1.71661	0.03068	0.03971	0.01618
145674	0.38045	0.01249	-2.73705	0.22454	0.17461	0.06924
124436	0.95912	0.01652	-0.93151	0.02656	0.11283	0.01407

IREF	*Parameters and Measures of Standard Error*					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
119840	0.79977	0.01610	-0.69789	0.03473	0.15538	0.01551
203218	1.18848	0.01979	0.54927	0.00979	0.10261	0.00426
139658	0.69365	0.01557	0.14940	0.02722	0.11599	0.01053
139665	0.78437	0.01949	0.47515	0.02327	0.21677	0.00859
119891	0.65079	0.02623	1.33224	0.02813	0.27649	0.00822
168364	0.68942	0.00824	-0.78772	0.01210	0.00000	0.00000
119825	0.78172	0.01540	-0.61493	0.03273	0.12805	0.01452
168582	0.49408	0.01134	-1.79433	0.08848	0.08506	0.03378
168754	1.30101	0.02120	0.50134	0.00920	0.11553	0.00417
145733	0.81136	0.01312	-0.71266	0.02443	0.04758	0.01143
119747	0.83676	0.00917	-0.65048	0.00973	0.00000	0.00000
168409	0.57872	0.00753	-0.79399	0.01405	0.00000	0.00000
145250	0.74568	0.00837	-0.31520	0.00932	0.00000	0.00000
119852	0.88103	0.00979	0.68105	0.00829	0.00000	0.00000
145252	0.78355	0.00907	-0.94708	0.01183	0.00000	0.00000
203063	0.80996	0.00941	0.79541	0.00935	0.00000	0.00000
203129	0.49540	0.00700	-0.62459	0.01483	0.00000	0.00000
119701	0.46604	0.00761	-1.77348	0.02838	0.00000	0.00000
119731	0.92931	0.01103	-1.31825	0.01271	0.00000	0.00000
119784	0.44193	0.00772	-2.05013	0.03427	0.00000	0.00000

Table H-2. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 3

IREF	Parameters and Measures of Standard Error									
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)
119780	0.76713	0.00473	0.40416	0.00732	1.05817	0.00890	-1.05817	0.00989	0	0
119862	0.85919	0.00551	0.11266	0.00652	0.82087	0.00829	-0.82087	0.00809	0	0
119920	0.94678	0.00725	-0.43617	0.00676	0.37873	0.00794	-0.37873	0.00706	0	0
124305	0.55144	0.00469	-1.41453	0.01305	0.71306	0.01520	-0.71306	0.01148	0	0
139695	0.70199	0.00425	-0.61379	0.00877	1.84834	0.01584	-1.84834	0.00999	0	0
145505	0.82627	0.00630	1.23284	0.00820	0.63153	0.00797	-0.63153	0.01083	0	0
168416	0.72342	0.00516	-0.09214	0.00770	0.62777	0.00935	-0.62777	0.00878	0	0
203056	0.97458	0.00667	-0.15000	0.00599	0.57130	0.00764	-0.57130	0.00690	0	0
203198	0.71826	0.00548	0.56282	0.00805	0.51398	0.00874	-0.51398	0.00945	0	0
215701	0.78428	0.00605	1.56105	0.00908	0.81092	0.00845	-0.81092	0.01368	0	0

Table H-3. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 4

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
123105	0.71816	0.01614	-2.02778	0.05600	0.11181	0.03071
145802	0.44512	0.01449	-1.10123	0.09332	0.13606	0.03089
120173	0.67135	0.01632	-0.06692	0.02904	0.15994	0.01147
144634	0.82104	0.01862	0.29108	0.01869	0.18227	0.00780
201329	0.91403	0.01659	-0.82691	0.02155	0.10934	0.01131
124560	0.71236	0.01345	-0.90480	0.02839	0.06024	0.01327
145538	0.84811	0.01895	0.68456	0.01427	0.13490	0.00584
201339	0.92318	0.01760	-0.20608	0.01747	0.14658	0.00837
120187	0.75290	0.01691	-0.50835	0.03006	0.18704	0.01297
120251	1.08588	0.02535	-1.32418	0.03166	0.32622	0.01673
124620	0.74394	0.01269	-2.45570	0.02492	0.00000	0.00000
120272	0.61905	0.01496	-0.58828	0.03975	0.13074	0.01600
139437	0.67589	0.01583	-1.43910	0.05146	0.14412	0.02424
120045	0.47811	0.01584	0.22940	0.04784	0.12786	0.01557
120108	0.45471	0.01840	-0.58072	0.09743	0.28396	0.02649

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120193	0.73302	0.02203	1.22920	0.01766	0.17102	0.00600
169097	0.51644	0.01442	0.14056	0.03737	0.08700	0.01325
120033	0.78367	0.01609	-1.96768	0.04356	0.08498	0.02508
201358	0.68837	0.01592	1.01789	0.01446	0.05235	0.00515
201331	0.51006	0.01477	-0.15263	0.04680	0.11308	0.01628
145848	0.50809	0.01129	-0.68254	0.04080	0.04572	0.01512
201356	0.80884	0.02061	0.94784	0.01530	0.15761	0.00578
139511	0.88162	0.01701	-2.39910	0.03411	0.04798	0.01999
120178	0.80707	0.01767	-2.50378	0.05177	0.08826	0.03259
139466	0.64883	0.01660	-0.40366	0.03820	0.19535	0.01472
144648	0.45372	0.00720	-0.92398	0.01640	0.00000	0.00000
144646	0.90108	0.02284	-0.17273	0.02614	0.36314	0.00997
145539	0.95480	0.02011	-1.52446	0.03340	0.17452	0.01939
169107	0.85870	0.01569	0.49117	0.01176	0.05674	0.00492
120214	0.86022	0.01452	-1.60496	0.02508	0.03999	0.01353

continued

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
168431	0.51296	0.01715	0.68438	0.03476	0.13728	0.01176
145821	0.83824	0.01591	-0.62232	0.02269	0.11431	0.01107
139497	1.15667	0.02510	0.58440	0.01180	0.22747	0.00516
139491	0.85818	0.01733	-0.52629	0.02351	0.16860	0.01114
120255	1.15359	0.02269	-0.27781	0.01584	0.24548	0.00803
120131	0.70893	0.00937	-1.19744	0.01292	0.00000	0.00000
169111	0.81070	0.00972	0.52498	0.00777	0.00000	0.00000
120221	0.55567	0.00829	-1.44032	0.01775	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
201383	0.58422	0.00801	0.49655	0.01013	0.00000	0.00000
120159	0.69329	0.00886	-0.81652	0.01096	0.00000	0.00000
124737	0.83951	0.00980	-0.43523	0.00812	0.00000	0.00000
120199	0.43721	0.00704	0.36763	0.01272	0.00000	0.00000
168457	0.36843	0.00660	-0.32537	0.01557	0.00000	0.00000
120071	0.85488	0.00985	-0.23285	0.00755	0.00000	0.00000
145085	0.71265	0.00941	-1.20082	0.01289	0.00000	0.00000

Table H-4. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 4

IREF	Parameters and Measures of Standard Error									
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)
120083	0.61510	0.00488	-1.00393	0.00904	0.93990	0.01253	-0.93990	0.00923	0.00000	0.00000
120203	0.67445	0.00493	-0.50514	0.00767	0.98770	0.01061	-0.98770	0.00860	0.00000	0.00000
120222	0.90094	0.00606	-0.01153	0.00669	1.64674	0.00977	-1.64674	0.00877	0.00000	0.00000
120284	0.63792	0.00505	-0.67427	0.00841	0.79783	0.01089	-0.79783	0.00892	0.00000	0.00000
124723	0.73510	0.00580	-0.63977	0.00747	0.71427	0.00974	-0.71427	0.00795	0.00000	0.00000
139450	0.86503	0.00774	0.57949	0.00661	0.31269	0.00701	-0.31269	0.00740	0.00000	0.00000
139487	0.69056	0.00448	0.18786	0.00765	1.47083	0.01000	-1.47083	0.01021	0.00000	0.00000
139527	0.72026	0.00541	-0.92358	0.00763	1.06272	0.01188	-1.06272	0.00810	0.00000	0.00000
201364	0.98009	0.00836	-0.12695	0.00584	0.31249	0.00669	-0.31249	0.00636	0.00000	0.00000
201425	0.30431	0.00258	0.72179	0.01597	1.26306	0.01746	-1.26306	0.01896	0.00000	0.00000

Table H-5. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 5

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120654	0.60768	0.00799	-1.00564	0.01480	0.00000	0.00000
167859	0.84250	0.01595	0.21943	0.01789	0.10042	0.00769
167675	0.88674	0.01895	-0.45153	0.03034	0.24858	0.01315

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120762	1.07315	0.02790	0.65219	0.01785	0.37215	0.00617
124937	0.90268	0.02247	0.41078	0.02227	0.30827	0.00806
145584	0.86489	0.01665	0.17238	0.01861	0.12541	0.00805

continued

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
204285	0.90147	0.01934	0.46596	0.01757	0.18577	0.00713
120818	0.84327	0.01898	0.29315	0.02200	0.21062	0.00868
120657	1.22520	0.02022	0.04083	0.01243	0.15336	0.00623
139344	0.84286	0.01778	-0.04043	0.02469	0.19298	0.01032
217161	0.60421	0.01750	0.67568	0.03013	0.14016	0.01062
145890	0.75484	0.01372	-1.73416	0.04452	0.06621	0.02467
204683	0.85788	0.01398	-0.57785	0.02191	0.05427	0.01076
120749	1.09400	0.02250	0.50664	0.01419	0.22726	0.00594
120626	1.12557	0.02320	0.49472	0.01405	0.24084	0.00589
120683	0.86203	0.02090	0.75578	0.01768	0.20302	0.00668
120720	1.00818	0.01944	-0.03526	0.01895	0.21069	0.00854
120702	0.92383	0.02128	0.50111	0.01863	0.24389	0.00722
124806	0.88164	0.01735	0.03255	0.02070	0.15945	0.00900
120630	0.85527	0.01542	-1.09942	0.03284	0.08768	0.01778
120831	0.90085	0.01767	-0.36383	0.02537	0.18499	0.01162
167797	0.50136	0.01779	-0.18753	0.07859	0.22017	0.02309

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
204245	1.03778	0.01827	-1.03887	0.02597	0.12397	0.01529
120682	0.72228	0.01743	-0.25926	0.03786	0.21579	0.01452
145294	0.79840	0.01859	-0.29505	0.03411	0.25490	0.01343
145870	0.64821	0.01751	-0.33890	0.04981	0.23081	0.01766
145112	1.66570	0.04652	1.45133	0.01064	0.24594	0.00304
120678	0.62585	0.01333	-0.03797	0.03049	0.05815	0.01191
120801	1.27128	0.02038	0.20209	0.01064	0.13079	0.00524
124040	0.76340	0.01788	0.26134	0.02542	0.18723	0.00980
145871	1.06889	0.01748	-0.28020	0.01610	0.11736	0.00823
145597	1.71048	0.03448	0.75283	0.00882	0.25473	0.00388
145102	0.75228	0.00844	-0.22881	0.00890	0.00000	0.00000
167878	0.62749	0.00772	-0.42284	0.01099	0.00000	0.00000
120659	0.68993	0.00830	-0.72107	0.01153	0.00000	0.00000
124804	0.73974	0.00854	-0.58062	0.01023	0.00000	0.00000
120783	0.57194	0.00734	-0.10222	0.01080	0.00000	0.00000
120850	0.88825	0.00940	-0.25218	0.00789	0.00000	0.00000

Table H-6. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 5

IREF	Parameters and Measures of Standard Error													
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
120639	0.87719	0.00760	-0.41190	0.00744	0.23645	0.00797	-0.23645	0.00747	0	0				
120694	0.45183	0.00381	0.70825	0.01209	0.69469	0.01280	-0.69469	0.01384	0	0				
120746	0.90464	0.00636	0.71875	0.00644	0.64103	0.00719	-0.64103	0.00840	0	0				
120791	0.97268	0.00768	0.04761	0.00616	0.29866	0.00697	-0.29866	0.00680	0	0				
145571	0.84842	0.00687	0.21498	0.00685	0.32025	0.00761	-0.32025	0.00762	0	0				
167743	0.99735	0.00731	0.26940	0.00582	0.42561	0.00678	-0.42561	0.00684	0	0				
120857	0.99974	0.00542	-0.20356	0.00539	1.73802	0.01133	0.76689	0.00783	-1.10091	0.00703	-1.40400	0.00754	0	0
144664	1.34843	0.00715	0.03124	0.00396	1.70631	0.00905	0.29224	0.00557	-0.76385	0.00558	-1.23470	0.00638	0	0
145942	0.72452	0.00387	0.13019	0.00690	1.47707	0.01049	0.65341	0.00883	-0.65749	0.00871	-1.47299	0.01027	0	0
167794	0.88631	0.00509	0.10664	0.00595	1.07928	0.00847	0.87444	0.00805	-0.74134	0.00757	-1.21237	0.00839	0	0

Table H-7. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 6

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
119316	0.88574	0.01528	-0.20263	0.01548	0.07450	0.00739
139336	0.76582	0.01765	-0.11435	0.02554	0.20701	0.01056
119311	0.78032	0.02116	0.69725	0.01986	0.24839	0.00739
139269	0.46721	0.01854	0.65594	0.04872	0.19478	0.01475
119219	1.12030	0.02229	0.01121	0.01458	0.24203	0.00704
119385	0.92075	0.01928	-0.11535	0.01946	0.22142	0.00879
167709	1.01755	0.02051	0.23039	0.01435	0.20363	0.00660
125025	0.51098	0.00738	-0.68451	0.01336	0.00000	0.00000
145606	0.49860	0.01629	-1.11715	0.08862	0.20504	0.03064
119231	0.54432	0.01895	0.31368	0.04496	0.25810	0.01391
119283	1.15270	0.02360	0.48429	0.01176	0.22149	0.00539
119232	0.32502	0.00679	-1.86803	0.03406	0.00000	0.00000
119221	0.74813	0.01885	0.49909	0.02080	0.20281	0.00808
125135	0.61243	0.01656	0.16074	0.03205	0.16987	0.01181
119216	0.75009	0.01454	-0.42190	0.02327	0.08473	0.01064
119393	0.46377	0.01807	0.81300	0.04318	0.16009	0.01363
202514	1.20221	0.02676	0.48652	0.01274	0.29791	0.00558
202497	0.60459	0.01691	0.26824	0.03183	0.17597	0.01152
202530	1.29058	0.02544	0.40509	0.01067	0.22938	0.00517
167914	1.42673	0.03123	1.08858	0.00850	0.18477	0.00344
119377	0.85118	0.01965	0.52800	0.01665	0.19804	0.00685

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
139274	0.74012	0.01549	-0.07266	0.02215	0.11719	0.00957
145130	0.67523	0.01307	-0.85417	0.03036	0.05725	0.01391
167910	0.56568	0.01962	1.14718	0.02561	0.16874	0.00885
119175	0.61952	0.01506	-0.19510	0.03320	0.12503	0.01310
119299	0.86900	0.02065	0.53154	0.01707	0.22776	0.00690
145346	1.00077	0.02216	-0.05667	0.01913	0.29791	0.00834
145124	0.90611	0.01769	0.15101	0.01548	0.14509	0.00707
167684	1.06187	0.02424	0.57467	0.01378	0.26671	0.00580
145978	0.71988	0.02156	1.02732	0.01954	0.21873	0.00699
119326	0.78975	0.01902	0.73761	0.01634	0.16413	0.00652
124983	0.99887	0.01710	0.26789	0.01115	0.08607	0.00520
119227	0.65269	0.00822	-0.29248	0.00947	0.00000	0.00000
119248	0.69255	0.00882	-0.87198	0.01129	0.00000	0.00000
119263	0.42557	0.00747	1.40994	0.01852	0.00000	0.00000
119319	0.75619	0.00917	-0.63972	0.00955	0.00000	0.00000
202623	0.93062	0.01062	-0.51727	0.00781	0.00000	0.00000
139370	0.55476	0.00803	-1.23893	0.01603	0.00000	0.00000

Table H-8. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 6

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
119203	0.67015	0.00468	-0.01525	0.00733	1.11286	0.00976	-1.11286	0.00886	0	0				
119205	0.81839	0.00760	0.73845	0.00687	0.29229	0.00716	-0.29229	0.00755	0	0				
119229	0.89603	0.00694	0.05147	0.00580	0.55201	0.00713	-0.55201	0.00672	0	0				
119281	0.76231	0.00564	-0.27745	0.00667	0.85459	0.00901	-0.85459	0.00760	0	0				

continued

IREF	Parameters and Measures of Standard Error													
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
119341	0.55991	0.00429	0.47660	0.00863	0.90154	0.01001	-0.90154	0.01055	0	0				
125128	0.89653	0.00629	0.14581	0.00570	0.93888	0.00759	-0.93888	0.00716	0	0				
167831	1.17903	0.00764	0.95119	0.00435	1.00301	0.00555	0.28044	0.00545	-0.50432	0.00648	-0.77912	0.00725	0	0
167982	0.98940	0.00559	0.21450	0.00470	1.37282	0.00781	0.53132	0.00637	-0.38703	0.00614	-1.51713	0.00796	0	0
202642	0.88678	0.00587	0.34712	0.00546	0.90507	0.00722	0.54825	0.00682	-0.56952	0.00692	-0.88382	0.00736	0	0
202876	0.99507	0.00554	0.22505	0.00469	1.45690	0.00796	0.58806	0.00639	-0.32009	0.00611	-1.72488	0.00867	0	0

Table H-9. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 7

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
119378	1.09602	0.02619	-0.28194	0.02593	0.42756	0.01038
204917	0.98501	0.02165	0.16426	0.02053	0.29236	0.00854
139925	0.99408	0.01952	0.45575	0.01483	0.17708	0.00653
120446	1.08141	0.01979	0.57933	0.01182	0.14121	0.00529
204888	1.11015	0.02083	0.76427	0.01094	0.13887	0.00468
205033	0.79805	0.01788	0.10085	0.02572	0.20709	0.01051
144729	1.47728	0.02365	0.14809	0.00937	0.14998	0.00518
145375	1.05625	0.02333	0.24920	0.01866	0.31663	0.00772
145364	0.84333	0.01729	0.34841	0.01882	0.15019	0.00803
120351	0.85238	0.01614	-0.08006	0.02106	0.12438	0.00975
120334	0.64308	0.01152	-0.45597	0.02810	0.03684	0.01189
125310	0.71715	0.01571	-0.76365	0.04187	0.14040	0.01917
120514	0.80013	0.01901	0.51969	0.02177	0.21162	0.00843
120419	0.71582	0.01611	-0.13024	0.03180	0.15818	0.01312
119242	1.17298	0.02055	0.45710	0.01116	0.14732	0.00527
125349	0.97051	0.02649	1.22729	0.01545	0.26131	0.00517
120421	0.84619	0.01510	-0.55955	0.02445	0.08408	0.01235
204998	1.25347	0.02560	0.80933	0.01103	0.21798	0.00466
204939	0.66791	0.01404	-0.85841	0.04315	0.08883	0.01979
120440	0.96147	0.01936	0.41526	0.01610	0.18708	0.00701
120511	0.83135	0.02122	-0.13695	0.03373	0.35394	0.01230

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
120496	0.28739	0.01177	0.81844	0.09034	0.05198	0.02066
145151	0.77593	0.02026	1.08695	0.01793	0.17529	0.00649
125286	0.42650	0.00705	-1.10004	0.02130	0.00000	0.00000
146126	1.05467	0.02118	0.62708	0.01335	0.19084	0.00574
205049	0.53292	0.01303	-0.83478	0.06189	0.08529	0.02406
204932	0.92127	0.01770	-0.06602	0.02020	0.16869	0.00944
120462	0.80796	0.01521	0.21552	0.01833	0.08791	0.00807
145370	0.64470	0.01933	0.01803	0.04690	0.30661	0.01509
205054	1.16515	0.01915	0.50899	0.00996	0.10053	0.00461
145166	0.95684	0.01725	-0.48765	0.02182	0.13473	0.01139
120400	0.77660	0.01724	0.53013	0.02002	0.14769	0.00808
169533	0.77792	0.00885	0.51420	0.00807	0.00000	0.00000
140280	0.58757	0.00780	-0.66539	0.01294	0.00000	0.00000
145157	0.79806	0.00901	-0.24301	0.00848	0.00000	0.00000
169224	0.92571	0.01009	-0.35062	0.00785	0.00000	0.00000
204925	0.59538	0.00768	0.59865	0.01023	0.00000	0.00000
120468	0.71549	0.00851	0.68424	0.00897	0.00000	0.00000

Table H-10. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 7

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
120413	0.52372	0.00360	1.89814	0.01052	2.02906	0.01103	-2.02906	0.02325	0	0				
145144	0.90386	0.00644	0.71405	0.00615	0.65059	0.00701	-0.65059	0.00796	0	0				
150893	0.85740	0.00859	1.39321	0.00913	0.21462	0.00794	-0.21462	0.00882	0	0				
169486	0.97713	0.00697	0.14742	0.00564	0.56057	0.00697	-0.56057	0.00670	0	0				
204567	0.87364	0.00782	0.80039	0.00708	0.23763	0.00722	-0.23763	0.00762	0	0				
204667	1.16547	0.00813	0.19185	0.00486	0.56747	0.00617	-0.56747	0.00597	0	0				
120458	1.02089	0.00646	1.09111	0.00544	0.88566	0.00633	0.21147	0.00653	-0.16802	0.00705	-0.92911	0.00931	0	0
125344	0.78945	0.00437	1.42440	0.00643	1.92152	0.00820	0.48244	0.00799	-0.88973	0.01162	-1.51423	0.01543	0	0
140292	1.19806	0.00660	0.77198	0.00435	1.69968	0.00702	0.27487	0.00566	-0.87044	0.00734	-1.10411	0.00816	0	0
169226	0.99100	0.00552	0.80275	0.00509	1.34253	0.00705	0.27211	0.00643	-0.48665	0.00725	-1.12800	0.00907	0	0

Table H-11. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 8

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
120969	1.02914	0.01781	-0.51406	0.02579	0.18435	0.01262
169411	0.64487	0.01559	-0.18597	0.04830	0.15607	0.01732
145386	0.42876	0.01565	-0.20714	0.11092	0.13502	0.03001
205021	0.85890	0.01604	0.05538	0.02378	0.14112	0.00999
120998	1.23857	0.02014	0.72540	0.00993	0.11116	0.00425
169406	1.08793	0.01861	0.19707	0.01627	0.17634	0.00735
121056	0.46690	0.01261	-0.17304	0.06903	0.06871	0.02115
205064	0.76293	0.02123	0.72507	0.02771	0.26847	0.00923
125542	1.34033	0.01979	-0.16179	0.01362	0.15032	0.00726
145197	1.01443	0.01614	-0.28646	0.02038	0.11882	0.00989
121025	1.27737	0.02341	-0.15521	0.01930	0.31689	0.00885
145643	1.20961	0.02666	1.10160	0.01205	0.20619	0.00440
139845	0.37052	0.00658	-1.28926	0.03128	0.00000	0.00000
169368	1.32124	0.02460	0.85785	0.01054	0.18318	0.00438
120961	1.19611	0.02438	0.54932	0.01541	0.29303	0.00620

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
139782	0.71232	0.02666	1.56079	0.02510	0.23120	0.00711
121024	0.70807	0.01433	-1.16456	0.05890	0.11057	0.02703
125592	0.61075	0.01954	-0.18234	0.07205	0.33252	0.02017
120894	1.10339	0.02183	0.51304	0.01604	0.24617	0.00657
205073	0.84687	0.01095	-1.02829	0.02115	0.02196	0.00923
120910	1.23965	0.02901	1.09938	0.01272	0.25479	0.00456
145641	0.86024	0.01717	-0.63846	0.03896	0.20929	0.01711
169404	1.58527	0.04124	1.42267	0.01135	0.24499	0.00332
122522	0.81956	0.01478	-0.06968	0.02514	0.10183	0.01071
120971	1.40469	0.02583	-0.25580	0.01864	0.34561	0.00888
121033	1.52709	0.02525	0.48400	0.01020	0.21216	0.00488
139756	0.99890	0.01477	0.01595	0.01536	0.06619	0.00707
120879	0.68435	0.01596	0.30099	0.03210	0.13643	0.01176
120950	1.05658	0.02618	1.28474	0.01401	0.19021	0.00457
121020	0.52959	0.01362	0.23581	0.04597	0.06666	0.01524

continued

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
205045	0.33296	0.00613	0.05483	0.01842	0.00000	0.00000
139848	0.86931	0.01891	0.59218	0.02015	0.19032	0.00779
204961	1.20235	0.01133	0.50707	0.00621	0.00000	0.00000
139865	1.06155	0.01032	-0.39065	0.00810	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
121040	0.44766	0.00678	-0.94597	0.02190	0.00000	0.00000
120932	0.64917	0.00815	-1.11808	0.01750	0.00000	0.00000
121061	0.99590	0.01130	1.25455	0.01010	0.00000	0.00000
125498	1.15055	0.01072	0.17648	0.00643	0.00000	0.00000

Table H-12. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 8

IREF	Parameters and Measures of Standard Error													
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
120876	1.08134	0.00657	-0.02276	0.00563	0.76698	0.00791	-0.76698	0.00673	0	0				
120936	1.20595	0.00885	-0.00912	0.00566	0.23559	0.00642	-0.23559	0.00607	0	0				
121031	1.24504	0.00773	-0.02810	0.00505	0.63221	0.00701	-0.63221	0.00599	0	0				
139885	0.79786	0.00518	-0.05862	0.00736	0.67023	0.00934	-0.67023	0.00837	0	0				
145393	1.27805	0.01154	1.45684	0.00775	0.22536	0.00669	-0.22536	0.00809	0	0				
146483	1.12018	0.00664	0.25643	0.00545	0.83742	0.00731	-0.83742	0.00702	0	0				
121002	1.31676	0.00689	0.69137	0.00438	0.93130	0.00604	0.18710	0.00559	-0.33406	0.00605	-0.78434	0.00716	0	0
121016	1.43658	0.00743	0.72685	0.00405	0.96655	0.00573	0.19111	0.00527	-0.35987	0.00581	-0.79778	0.00700	0	0
125528	1.30279	0.00730	-0.18708	0.00460	0.62464	0.00710	0.28018	0.00631	-0.13497	0.00578	-0.76985	0.00574	0	0
139894	1.14182	0.00531	0.68007	0.00485	1.56503	0.00786	0.59969	0.00628	-0.47848	0.00695	-1.68624	0.01269	0	0

Table H-13. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Mathematics Grade 11

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
140086	0.92393	0.02144	-1.39169	0.04797	0.13436	0.02907
140042	1.32060	0.02491	0.31752	0.01191	0.13221	0.00584
119573	0.68914	0.01560	-1.12153	0.05218	0.07866	0.02505
145233	0.98049	0.01883	0.12888	0.01678	0.07958	0.00782
205802	1.55422	0.04003	0.91145	0.01203	0.28741	0.00480
206305	0.98963	0.02497	0.53153	0.01965	0.22775	0.00796
205785	0.99181	0.02118	-0.31474	0.02493	0.16200	0.01233

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
140013	1.59549	0.03741	1.12644	0.00982	0.16093	0.00366
145230	1.97961	0.05109	1.52164	0.00906	0.09968	0.00238
125758	1.02494	0.02933	1.32323	0.01515	0.15685	0.00490
144783	2.25019	0.07529	1.54948	0.01019	0.24205	0.00311
144797	1.34531	0.04465	1.75326	0.01530	0.14813	0.00325
119419	0.64466	0.02079	-1.13872	0.09317	0.24275	0.03714
169568	1.39279	0.02362	0.24897	0.01008	0.08021	0.00496

continued

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
144820	1.03748	0.02075	-0.72135	0.02615	0.11454	0.01475
119576	0.59493	0.02280	0.52108	0.04978	0.23585	0.01568
125769	1.18425	0.03737	1.27846	0.01559	0.27041	0.00503
145469	0.36631	0.00788	-1.16062	0.03074	0.00000	0.00000
140181	1.37712	0.02815	0.25141	0.01335	0.20477	0.00662
119437	1.05879	0.02519	0.35134	0.01923	0.23543	0.00821
119625	1.02725	0.01935	0.04005	0.01656	0.08536	0.00797
169559	0.79470	0.02098	0.39761	0.02645	0.17103	0.01047
169577	1.45896	0.02644	0.63801	0.00907	0.09095	0.00399
205745	0.92428	0.02274	0.94229	0.01551	0.11760	0.00597
145461	0.89421	0.01109	-0.09715	0.00905	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	SE (a)	b	SE (b)	c	SE (c)
140089	1.37068	0.01674	0.86406	0.00710	0.00000	0.00000
119562	0.99714	0.01298	0.99880	0.00953	0.00000	0.00000
125854	1.23215	0.01609	1.14317	0.00869	0.00000	0.00000
125822	1.21097	0.01564	1.09523	0.00859	0.00000	0.00000
145234	0.92677	0.01179	0.74022	0.00912	0.00000	0.00000
210417	1.13677	0.01331	0.00675	0.00747	0.00000	0.00000
119603	0.97063	0.01353	1.28836	0.01140	0.00000	0.00000
205985	1.14678	0.01403	0.77985	0.00785	0.00000	0.00000
169588	0.63121	0.00909	0.33448	0.01146	0.00000	0.00000
119553	0.99375	0.01218	0.56336	0.00826	0.00000	0.00000
119464	1.18768	0.01410	0.58206	0.00728	0.00000	0.00000

Table H-14. 2012–13 NECAP: IRT Parameters for Polytomous Items
Mathematics Grade 11

IREF	Parameters and Measures of Standard Error													
	a	SE (a)	b	SE (b)	D0	SE (D0)	D1	SE (D1)	D2	SE (D2)	D3	SE (D3)	D4	SE (D4)
119620	0.80455	0.01072	1.60831	0.01426	0.14856	0.01087	-0.14856	0.01182	0	0				
125790	1.06187	0.01067	1.45680	0.00877	0.42166	0.00815	-0.42166	0.01071	0	0				
140033	1.10087	0.01132	1.25886	0.00825	0.29396	0.00784	-0.29396	0.00921	0	0				
169595	1.29853	0.01256	0.19043	0.00605	0.16632	0.00663	-0.16632	0.00656	0	0				
169711	1.00307	0.00791	1.00501	0.00713	0.96722	0.00797	-0.96722	0.01151	0	0				
169792	1.19455	0.01283	0.67063	0.00695	0.09688	0.00706	-0.09688	0.00721	0	0				
119460	1.39971	0.00977	1.48984	0.00507	1.37427	0.00627	0.39011	0.00684	-0.56922	0.01051	-1.19516	0.01696	0	0
119489	1.61137	0.01319	1.15339	0.00504	0.55252	0.00574	0.40174	0.00582	-0.29268	0.00706	-0.66157	0.00859	0	0
119623	1.21449	0.00805	1.30315	0.00549	1.44031	0.00702	0.66126	0.00690	-0.63148	0.01034	-1.47009	0.01804	0	0
141322	1.37826	0.00915	0.29730	0.00483	1.08913	0.00752	0.97566	0.00726	-0.56123	0.00659	-1.50355	0.00918	0	0

Table H-15. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 3

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
171719	0.69229	0.00835	-0.80298	0.01187	0.00000	0.00000
171765	0.83519	0.01527	-0.91521	0.03042	0.11262	0.01464
171325	0.80837	0.01623	-0.90168	0.03577	0.16528	0.01629
171327	1.08647	0.01848	-1.04639	0.02310	0.13962	0.01295
171329	0.87818	0.01783	-0.92577	0.03449	0.21564	0.01592
171334	0.47236	0.01308	-0.92048	0.08030	0.09557	0.02642
117787	0.47784	0.00720	-1.17077	0.01957	0.00000	0.00000
117793	0.61214	0.01605	-0.88442	0.06266	0.19880	0.02279
117789	1.04361	0.01866	-0.84070	0.02387	0.18360	0.01237
117790	0.40768	0.00683	-1.22511	0.02315	0.00000	0.00000
117794	0.37107	0.00667	-1.30771	0.02628	0.00000	0.00000
117798	0.88869	0.01844	-0.83255	0.03374	0.24281	0.01509
117796	1.12008	0.02236	0.63326	0.01175	0.16836	0.00495
117797	0.88999	0.01998	0.35267	0.01910	0.21535	0.00761

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
208039	0.59875	0.00775	-0.84389	0.01365	0.00000	0.00000
208046	1.03440	0.01911	0.10600	0.01533	0.17141	0.00696
208067	0.58685	0.00820	-1.47217	0.01933	0.00000	0.00000
208072	1.37640	0.02150	0.17641	0.00939	0.12496	0.00473
208076	1.26729	0.02567	0.44260	0.01230	0.25225	0.00539
208088	0.85839	0.01526	-0.30296	0.02089	0.10218	0.00945
208086	0.72207	0.01825	0.26725	0.02749	0.20275	0.01007
208093	1.29027	0.02234	-0.32165	0.01495	0.22465	0.00777
171725	0.51020	0.01265	-0.10590	0.04262	0.05595	0.01445
148224	0.72535	0.00991	-1.76330	0.01931	0.00000	0.00000
171511	0.98096	0.01544	-1.33567	0.02429	0.05555	0.01344
171515	1.04526	0.01745	-1.26900	0.02504	0.09691	0.01442
171525	0.29589	0.00650	-1.84506	0.04199	0.00000	0.00000
171508	0.37499	0.01669	0.73005	0.07311	0.09664	0.01991

Table H-16. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 3

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
117801	0.64263	0.00324	-1.34301	0.00841	1.98763	0.02207	1.10409	0.01553	-0.19370	0.01078	-2.89804	0.01140	0	0
117802	0.47358	0.00285	-3.44779	0.01168	2.74831	0.06097	1.94287	0.04484	-0.82082	0.01812	-3.87035	0.01250	0	0
171335	0.70277	0.00344	1.16585	0.00740	2.74002	0.01138	0.95418	0.00876	-0.85028	0.01242	-2.84390	0.03183	0	0
171554	0.75252	0.00372	1.28291	0.00713	3.03450	0.01166	0.78595	0.00843	-1.07227	0.01390	-2.74819	0.03389	0	0
208096	0.71635	0.00511	-1.83611	0.00984	1.59833	0.02358	0.18514	0.01289	-0.59019	0.01023	-1.19328	0.00916	0	0
208103	1.09339	0.00559	0.89895	0.00493	1.76543	0.00725	0.71895	0.00617	-0.62274	0.00833	-1.86164	0.01757	0	0

Table H-17. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 4

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
208865	1.24339	0.02252	-0.26173	0.01609	0.23290	0.00848
208860	0.82216	0.01643	-0.77065	0.03289	0.15762	0.01551
171911	0.76320	0.01443	-1.04761	0.03561	0.08539	0.01729
171904	0.62795	0.00950	-1.81705	0.02355	0.00000	0.00000
171908	0.84011	0.01888	0.38955	0.02075	0.20457	0.00841
171970	0.75207	0.01549	-1.01490	0.04075	0.13069	0.01916
118014	1.64634	0.02843	-0.96410	0.01394	0.13873	0.00958
118015	0.77547	0.01868	0.68077	0.02000	0.17389	0.00780
118017	1.03523	0.01873	0.23496	0.01439	0.15127	0.00679
118019	1.40941	0.02359	-0.30344	0.01300	0.18904	0.00753
118020	1.11423	0.01999	-0.02592	0.01553	0.18793	0.00771
118021	0.85936	0.01870	0.08866	0.02338	0.22453	0.00970
118022	0.69299	0.01447	-0.80522	0.04020	0.10925	0.01766
118023	1.06740	0.01853	-0.76611	0.02131	0.14131	0.01171

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
208641	1.01171	0.01542	-0.24580	0.01431	0.06136	0.00708
208667	0.70165	0.01770	0.94892	0.01909	0.11955	0.00723
208672	0.79721	0.02940	1.75386	0.02170	0.22995	0.00540
208683	0.49278	0.01768	-0.03914	0.07299	0.23469	0.02092
208686	0.93834	0.02494	0.19482	0.02685	0.41802	0.00912
208701	1.16487	0.02004	-0.25052	0.01570	0.17483	0.00828
208713	0.87161	0.02768	1.45231	0.01791	0.24558	0.00545
208714	1.45555	0.02581	0.15051	0.01122	0.22933	0.00600
208849	0.79668	0.01545	-1.52900	0.04263	0.09091	0.02280
172219	0.89344	0.01734	-1.43943	0.03770	0.11773	0.02113
208048	0.89552	0.01453	-0.74814	0.02171	0.05907	0.01091
208052	0.89371	0.01622	-1.61851	0.03499	0.07150	0.02008
208061	0.20271	0.00782	-1.47146	0.18716	0.06598	0.03052
208065	0.55729	0.00916	-2.05487	0.02923	0.00000	0.00000

Table H-18. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 4

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
118026	0.69367	0.00402	-0.88740	0.00747	1.52840	0.01579	0.50656	0.01109	-0.35977	0.00919	-1.67518	0.00888	0	0
118028	0.69228	0.00548	-2.77589	0.01112	1.83968	0.04517	0.76482	0.02590	-0.57914	0.01450	-2.02537	0.00968	0	0
171975	0.92307	0.00475	0.01142	0.00549	1.52925	0.01021	0.81823	0.00803	-0.44732	0.00688	-1.90015	0.00958	0	0
208080	0.70741	0.00358	1.30103	0.00738	3.24289	0.01324	1.18560	0.00859	-1.22721	0.01360	-3.20127	0.03631	0	0
208727	0.73070	0.00463	-1.76566	0.00786	2.19212	0.03288	0.53885	0.01492	-0.79070	0.00968	-1.94028	0.00835	0	0
208740	0.66727	0.00331	1.01572	0.00768	3.24876	0.01494	0.89352	0.00900	-1.13298	0.01212	-3.00929	0.02705	0	0

Table H-19. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 5

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
209705	0.85934	0.01303	-0.54668	0.02221	0.03891	0.01047
118120	0.75995	0.01686	-0.07059	0.03340	0.17979	0.01308
209539	0.69826	0.00937	-1.48988	0.02009	0.00000	0.00000
209546	0.61319	0.00828	-1.24950	0.01924	0.00000	0.00000
209551	0.82710	0.01076	-1.53806	0.01825	0.00000	0.00000
209552	0.45111	0.01583	1.04979	0.04468	0.07055	0.01392
209591	0.45848	0.01452	-1.15660	0.13351	0.14497	0.04182
209593	0.60358	0.00806	-1.12147	0.01807	0.00000	0.00000
209595	0.98681	0.01768	0.12657	0.01785	0.14833	0.00808
209599	0.86719	0.01710	-0.81224	0.03870	0.16538	0.01853
209616	0.83990	0.01021	-1.27163	0.01519	0.00000	0.00000
209617	0.66218	0.00849	-1.13015	0.01682	0.00000	0.00000
209618	0.52216	0.00901	-0.48870	0.03399	0.02632	0.01133
209631	0.84507	0.01019	-1.23901	0.01481	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
209325	0.99999	0.01599	-0.57966	0.02222	0.08711	0.01154
209352	0.57733	0.01240	-0.56918	0.05097	0.06341	0.01916
209359	1.05212	0.01817	-0.91578	0.02834	0.13435	0.01587
209363	1.06576	0.01856	-0.73405	0.02607	0.15984	0.01396
209368	0.50612	0.00699	0.26255	0.01233	0.00000	0.00000
209370	1.05320	0.01815	-1.42195	0.03428	0.08930	0.02182
209372	0.97790	0.01634	-0.74508	0.02640	0.10035	0.01394
209377	0.65267	0.01423	0.21200	0.03006	0.07357	0.01146
209700	0.66657	0.00873	-1.27980	0.01831	0.00000	0.00000
149127	0.92001	0.01708	-0.52521	0.02908	0.15746	0.01383
171822	0.71674	0.00878	-1.04310	0.01497	0.00000	0.00000
171825	1.00204	0.01435	-1.87207	0.02018	0.00000	0.00000
171837	0.90414	0.01657	-0.99155	0.03660	0.12062	0.01922
171842	0.71621	0.01500	-1.36775	0.06321	0.11389	0.03054

Table H-20. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 5

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE(D4)*
171848	0.74319	0.00338	0.68999	0.00723	2.38283	0.01251	0.66411	0.00872	-0.71502	0.01014	-2.33191	0.01909	0	0
209380	1.06178	0.00516	1.24755	0.00541	2.25543	0.00832	0.62381	0.00675	-0.77568	0.01085	-2.10355	0.02607	0	0
209383	1.33490	0.00666	1.06659	0.00443	2.04141	0.00733	0.55609	0.00566	-0.72665	0.00888	-1.87085	0.02040	0	0
209558	0.67407	0.00303	-0.06928	0.00794	2.56815	0.01854	0.80405	0.01052	-0.93471	0.00984	-2.43749	0.01455	0	0
209634	0.84743	0.00408	1.20181	0.00689	3.33058	0.01443	0.77312	0.00792	-1.10250	0.01330	-3.00121	0.04079	0	0
209640	0.96268	0.00460	1.20183	0.00586	2.31423	0.00907	0.65061	0.00719	-0.80472	0.01109	-2.16012	0.02505	0	0

Table H-21. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 6

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
172264	0.72139	0.01335	-1.31866	0.03615	0.05963	0.01800
118387	0.48074	0.01185	-1.05998	0.06633	0.07189	0.02402
172111	0.62644	0.00964	-1.91485	0.02257	0.00000	0.00000
172131	0.70185	0.01435	-0.92203	0.03718	0.09339	0.01721
172129	0.60334	0.01322	-0.76186	0.04258	0.07721	0.01764
172134	0.77090	0.01128	-1.88232	0.01911	0.00000	0.00000
209118	0.46276	0.01542	0.08671	0.06058	0.12911	0.01883
209119	1.04502	0.01801	-2.00202	0.02540	0.03814	0.01517
209127	0.79785	0.01439	-0.60684	0.02401	0.07460	0.01155
209137	0.34449	0.00648	-0.88169	0.02161	0.00000	0.00000
209142	0.48029	0.00746	-1.22334	0.01910	0.00000	0.00000
209143	0.80773	0.01169	-1.86183	0.01827	0.00000	0.00000
209146	0.91526	0.01499	-1.63042	0.02569	0.04020	0.01446
209148	0.68629	0.01122	-1.46424	0.02959	0.03254	0.01343

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
208965	0.95139	0.01604	-1.52019	0.02592	0.05121	0.01523
208967	0.56809	0.00991	-0.99187	0.03169	0.03021	0.01236
208972	0.60111	0.00931	-1.88504	0.02287	0.00000	0.00000
208974	0.85068	0.01172	-1.68496	0.01582	0.00000	0.00000
208981	1.16402	0.02140	-0.90988	0.02046	0.18732	0.01201
208986	0.65552	0.01381	-0.75360	0.03782	0.08579	0.01647
208987	0.44885	0.01110	-0.37131	0.05085	0.04818	0.01670
208989	0.71898	0.01631	-1.27636	0.04988	0.15711	0.02369
208944	0.31068	0.00699	-2.27788	0.04644	0.00000	0.00000
172242	0.41629	0.01744	0.60159	0.06696	0.15752	0.01873
209267	0.60877	0.01082	-2.50494	0.03242	0.00000	0.00000
209284	0.33888	0.00680	-1.67944	0.03259	0.00000	0.00000
209288	0.28475	0.01163	0.27598	0.10745	0.06159	0.02405
209293	0.53026	0.01572	0.51786	0.03611	0.11635	0.01236

Table H-22. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 6

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
172140	0.74656	0.00370	-0.21665	0.00669	2.95364	0.02012	0.98244	0.00948	-1.02406	0.00835	-2.91201	0.01476	0	0
208991	1.12940	0.00593	0.77543	0.00458	2.23296	0.00840	0.73019	0.00583	-0.76026	0.00773	-2.20290	0.01746	0	0
208995	0.93214	0.00482	0.91724	0.00549	2.58872	0.00991	0.79354	0.00672	-0.95872	0.00969	-2.42354	0.02131	0	0
209151	1.05024	0.00548	0.34716	0.00502	2.61125	0.01256	0.96304	0.00667	-0.92807	0.00736	-2.64622	0.01747	0	0
209153	1.01038	0.00525	0.75029	0.00510	2.28545	0.00906	1.00296	0.00645	-0.83195	0.00834	-2.45646	0.01999	0	0
209299	1.06366	0.00558	0.00753	0.00509	2.47829	0.01399	1.12141	0.00757	-1.04786	0.00687	-2.55183	0.01327	0	0

Table H-23. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 7

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
173338	0.27816	0.00659	-1.86520	0.03794	0.00000	0.00000
207560	0.61590	0.01273	-1.02774	0.03630	0.06167	0.01650
147203	0.55708	0.01039	-2.52692	0.03059	0.00000	0.00000
147210	0.32920	0.00700	-1.92247	0.03364	0.00000	0.00000
147215	0.34927	0.00708	-1.80854	0.03036	0.00000	0.00000
147213	0.53677	0.01173	-1.48839	0.05047	0.06202	0.02185
207625	0.54963	0.01135	-2.94296	0.03866	0.00000	0.00000
207630	0.49572	0.00860	-1.98198	0.02491	0.00000	0.00000
207633	0.60301	0.00989	-2.00946	0.02188	0.00000	0.00000
207636	0.75217	0.01155	-1.88560	0.01746	0.00000	0.00000
207640	0.36748	0.00680	-1.05748	0.02040	0.00000	0.00000
207643	0.46577	0.00749	-1.13077	0.01713	0.00000	0.00000
207644	0.58136	0.01051	-1.33410	0.03133	0.03310	0.01337
207647	0.71007	0.01227	-2.33553	0.02360	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
173246	0.50177	0.01688	-0.88508	0.08002	0.24109	0.02657
173249	0.66811	0.01433	-0.66536	0.03151	0.09518	0.01432
173247	0.32577	0.00878	-1.37879	0.07964	0.05035	0.02222
173255	0.42096	0.01365	-0.07566	0.05937	0.08375	0.01903
173254	0.89988	0.01771	-1.27838	0.02751	0.11813	0.01608
173260	0.39253	0.00699	-1.14164	0.02000	0.00000	0.00000
173261	1.03304	0.01942	-0.88365	0.01966	0.14534	0.01150
173269	0.39390	0.00688	-0.85484	0.01746	0.00000	0.00000
207521	0.39576	0.00686	-0.76882	0.01674	0.00000	0.00000
147686	0.58587	0.01354	0.00159	0.02775	0.06784	0.01108
147626	0.35655	0.01092	-1.19994	0.09500	0.07463	0.02859
147623	0.69404	0.01264	-1.40690	0.03016	0.04466	0.01509
147625	0.54934	0.01270	-1.09393	0.04796	0.07431	0.02029
147633	0.35955	0.00661	-0.61956	0.01715	0.00000	0.00000

Table H-24. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 7

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
147224	0.88433	0.00481	0.34588	0.00522	2.16364	0.00998	0.84167	0.00694	-0.78460	0.00725	-2.22071	0.01199	0	0
147637	0.82405	0.00448	-0.25290	0.00565	2.83082	0.01788	0.80400	0.00815	-1.06075	0.00724	-2.57406	0.01097	0	0
173286	1.10676	0.00616	0.21168	0.00429	2.06983	0.00938	0.75375	0.00604	-0.79431	0.00611	-2.02927	0.00962	0	0
173293	0.92212	0.00548	0.20184	0.00499	1.18603	0.00747	0.53016	0.00654	-0.26648	0.00639	-1.44972	0.00813	0	0
207649	1.02221	0.00574	-0.46152	0.00465	2.20795	0.01434	0.88475	0.00785	-0.82203	0.00597	-2.27067	0.00819	0	0
207654	1.04740	0.00582	0.42064	0.00450	1.98664	0.00845	0.80786	0.00609	-0.77554	0.00660	-2.01897	0.01083	0	0

Table H-25. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 8

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
172713	0.43001	0.01083	-1.73658	0.08819	0.07947	0.03293
208253	0.55500	0.01498	-0.66169	0.05979	0.13723	0.02362
172356	0.52851	0.00843	-1.42833	0.02065	0.00000	0.00000
172361	0.71861	0.01805	-0.96653	0.05093	0.22671	0.02336
172359	0.62786	0.02048	-0.84767	0.07580	0.37705	0.02541
172353	0.75829	0.01321	-2.14459	0.02444	0.00000	0.00000
207706	0.50776	0.01235	-0.57438	0.05208	0.06725	0.02004
207708	0.56967	0.01637	-0.00438	0.04684	0.18183	0.01682
207710	0.77451	0.01608	-0.99149	0.03645	0.11391	0.01943
207711	0.51011	0.00819	-1.36755	0.02057	0.00000	0.00000
207715	0.66077	0.00957	-1.32176	0.01633	0.00000	0.00000
207717	0.40107	0.01022	-0.52348	0.06141	0.04947	0.01961
207742	0.41523	0.00796	-1.92339	0.03231	0.00000	0.00000
207746	0.34748	0.00769	-2.32346	0.04473	0.00000	0.00000

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
207829	0.75365	0.01446	-0.54824	0.02729	0.07535	0.01354
207831	0.65468	0.01741	0.45006	0.02970	0.18943	0.01112
207832	0.34148	0.01348	0.74588	0.07280	0.06362	0.01976
207833	0.55688	0.01435	-0.40086	0.04954	0.10936	0.01949
207835	0.60913	0.01357	-0.85644	0.04585	0.08217	0.02069
207840	0.60471	0.00898	-1.29963	0.01729	0.00000	0.00000
207844	0.43323	0.01479	0.26594	0.06284	0.10324	0.01976
207850	0.48474	0.01352	0.33032	0.04209	0.06800	0.01470
118749	0.22351	0.00621	-1.57553	0.04655	0.00000	0.00000
208279	0.36415	0.00936	-1.67183	0.09001	0.06418	0.02837
207763	0.34734	0.00953	-0.32448	0.06762	0.04510	0.01895
207769	0.59418	0.01271	-1.35028	0.05202	0.07129	0.02462
207772	0.50725	0.01211	-1.06535	0.06169	0.07346	0.02493
207777	0.49098	0.00756	-0.86719	0.01638	0.00000	0.00000

Table H-26. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 8

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
172373	0.99322	0.00527	0.18136	0.00486	2.29402	0.01234	0.76303	0.00690	-0.76239	0.00647	-2.29464	0.01060	0	0
207754	0.99253	0.00533	0.83074	0.00483	2.03579	0.00832	0.75172	0.00625	-0.67133	0.00726	-2.11618	0.01356	0	0
207757	1.15947	0.00629	0.04909	0.00422	2.01236	0.01104	0.70886	0.00640	-0.64818	0.00565	-2.07304	0.00850	0	0
207782	0.80898	0.00421	0.51501	0.00585	2.50380	0.01221	0.84997	0.00759	-0.82283	0.00797	-2.53094	0.01454	0	0
207852	1.19264	0.00655	0.79665	0.00417	2.18171	0.00826	0.66250	0.00552	-0.78742	0.00672	-2.05678	0.01260	0	0
207854	1.11623	0.00603	0.68862	0.00435	1.99867	0.00821	0.67278	0.00580	-0.69806	0.00656	-1.97341	0.01134	0	0

Table H-27. 2012–13 NECAP: IRT Parameters for Dichotomous Items
Reading Grade 11

IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
147908	0.39314	0.01577	-0.78567	0.12879	0.12563	0.03775
172855	0.66556	0.02179	0.23670	0.04294	0.26784	0.01488
208337	0.52167	0.01756	-1.77656	0.12636	0.17538	0.05038
208345	0.36907	0.00794	-0.89476	0.02739	0.00000	0.00000
208480	0.63535	0.01959	-0.49069	0.05879	0.22070	0.02211
208482	0.47117	0.01020	-2.18593	0.04168	0.00000	0.00000
208486	0.77123	0.01924	-0.68268	0.04113	0.16102	0.01893
208478	0.61301	0.01550	-0.86445	0.05280	0.08451	0.02243
208502	0.75739	0.01769	-0.46469	0.03421	0.11187	0.01561
208519	0.94079	0.01995	-0.46430	0.02503	0.12499	0.01274
208520	0.67160	0.01135	-1.44784	0.02198	0.00000	0.00000
208524	0.62268	0.01127	-1.71353	0.02673	0.00000	0.00000
208684	0.33452	0.00777	-1.03445	0.03198	0.00000	0.00000
208682	1.03993	0.02308	-0.72021	0.02723	0.17261	0.01476

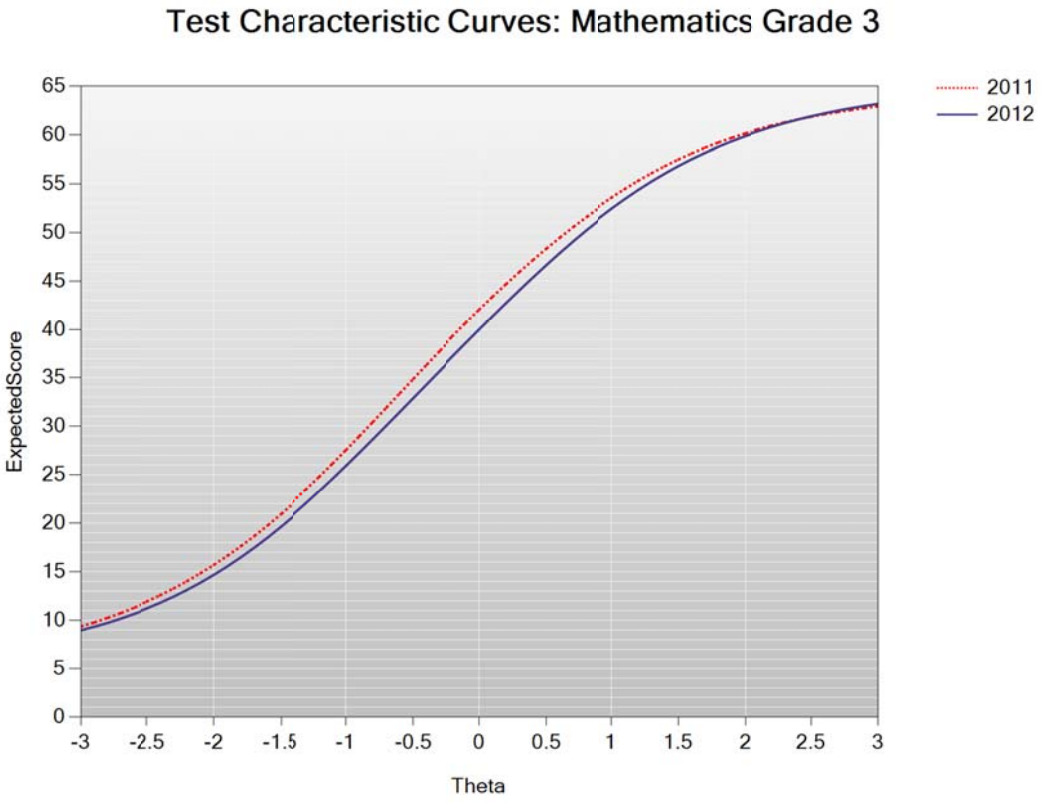
IREF	Parameters and Measures of Standard Error					
	a	*SE (a)*	*b*	*SE (b)*	*c*	*SE (c)*
208680	0.45886	0.01613	-0.16965	0.07478	0.10366	0.02419
208689	0.51596	0.01061	-2.08466	0.03713	0.00000	0.00000
147381	0.79105	0.02070	0.28681	0.02708	0.19198	0.01122
147382	0.73783	0.01698	-0.64418	0.03679	0.09248	0.01700
147383	0.66561	0.01039	-0.86631	0.01644	0.00000	0.00000
147385	0.99399	0.02130	-0.56905	0.02536	0.14384	0.01336
208704	0.45015	0.01103	-2.76186	0.05689	0.00000	0.00000
208710	0.71429	0.01533	-1.00758	0.03847	0.05898	0.01808
208721	0.59286	0.01462	-0.77388	0.04972	0.06944	0.02049
208726	0.55136	0.01476	-0.88985	0.06199	0.08203	0.02458
208729	0.64521	0.01071	-1.26557	0.02062	0.00000	0.00000
208730	0.63810	0.01071	-1.31307	0.02131	0.00000	0.00000
208742	0.64061	0.01465	-1.00082	0.04584	0.06340	0.02027
226071	0.30281	0.00749	0.73590	0.02166	0.00000	0.00000

Table H-28. 2012–13 NECAP: IRT Parameters for Polytomous Items
Reading Grade 11

IREF	Parameters and Measures of Standard Error													
	a	*SE (a)*	*b*	*SE (b)*	*D0*	*SE (D0)*	*D1*	*SE (D1)*	*D2*	*SE (D2)*	*D3*	*SE (D3)*	*D4*	*SE (D4)*
147388	1.22148	0.00778	0.42019	0.00500	2.01350	0.01171	0.84734	0.00741	-0.80326	0.00695	-2.05758	0.01192	0	0
208528	1.11865	0.00703	0.02900	0.00530	2.02634	0.01455	0.75221	0.00848	-0.66287	0.00685	-2.11567	0.01022	0	0
208545	1.10397	0.00696	0.31018	0.00530	1.56510	0.01041	0.76710	0.00786	-0.52782	0.00696	-1.80437	0.01011	0	0
208700	1.13071	0.00709	0.51305	0.00532	2.07316	0.01181	0.83715	0.00755	-0.77920	0.00742	-2.13111	0.01341	0	0
208744	1.10583	0.00695	0.20429	0.00543	2.19441	0.01455	0.84351	0.00834	-0.83675	0.00717	-2.20118	0.01173	0	0
208747	1.20374	0.00770	0.27096	0.00506	1.83586	0.01157	0.93449	0.00801	-0.73385	0.00673	-2.03650	0.01083	0	0

APPENDIX I—TEST CHARACTERISTIC CURVES AND TEST INFORMATION FUNCTION PLOTS

Figure I-1. 2012–13 NECAP: Mathematics Grade 3 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



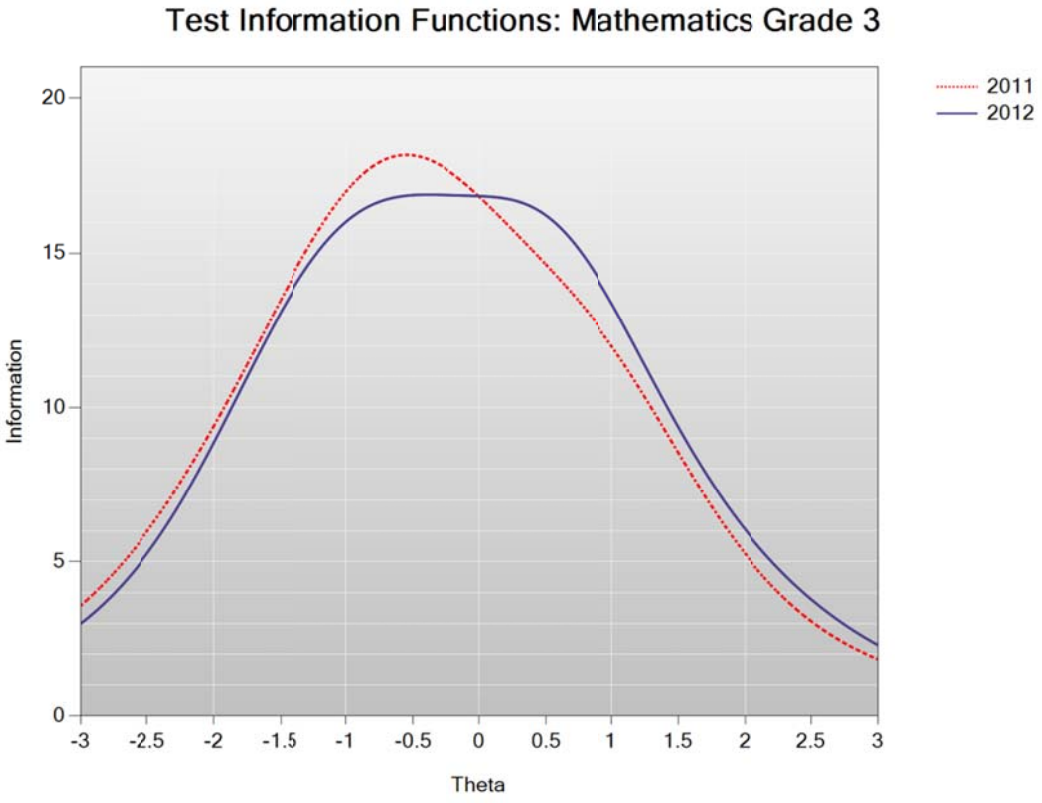
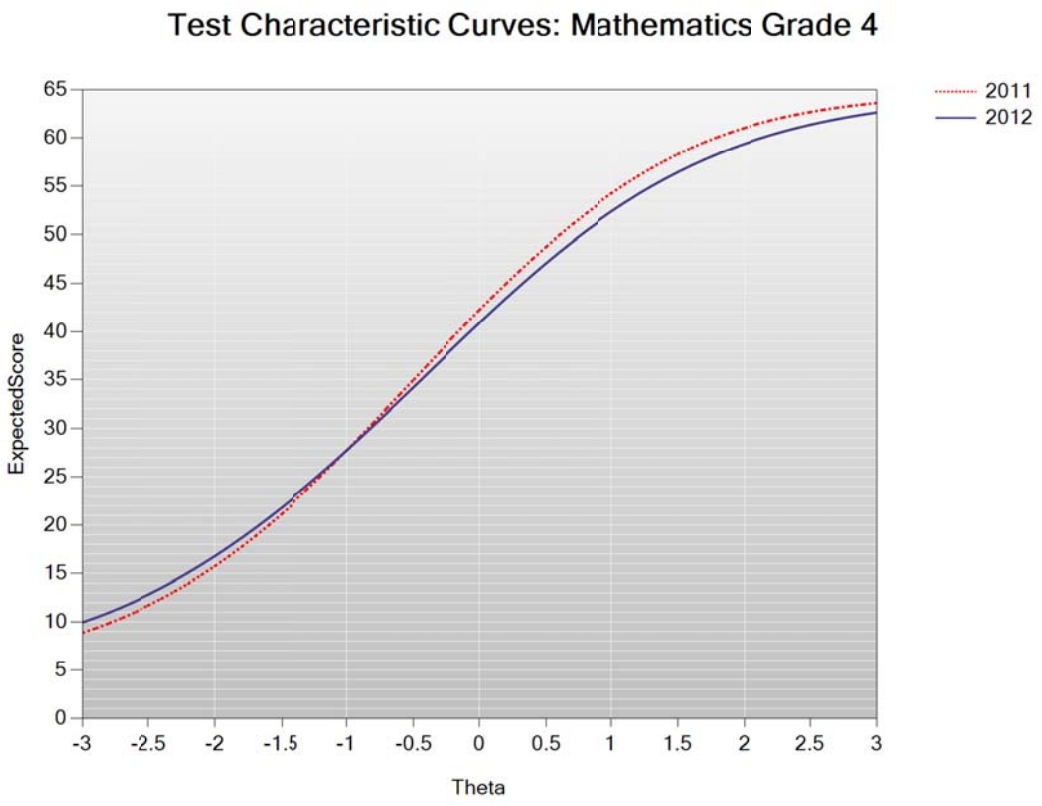


Figure I-2. 2012–13 NECAP: Mathematics Grade 4 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



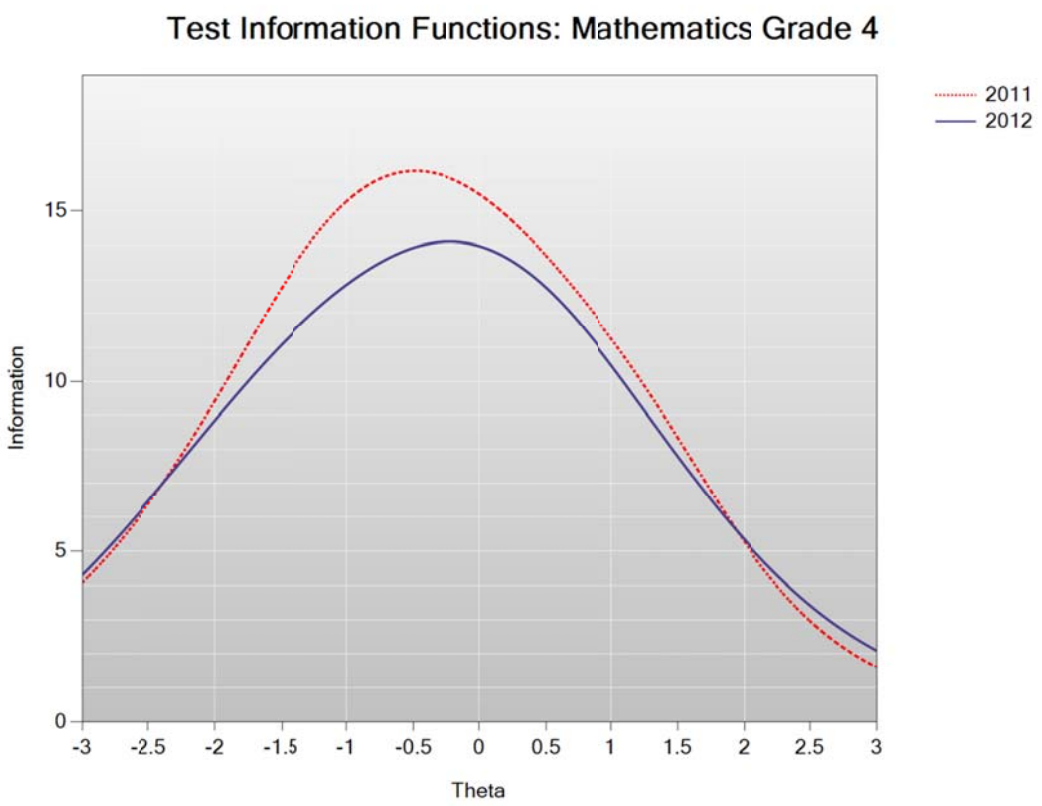
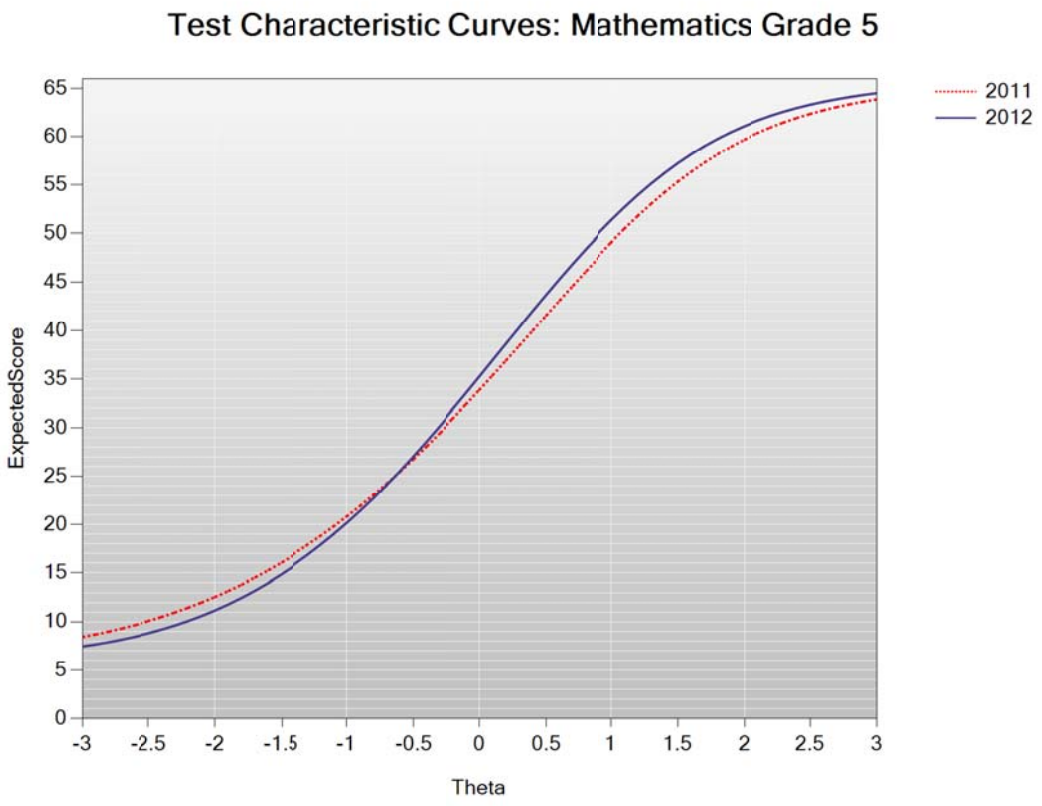


Figure I-3. 2012–13 NECAP: Mathematics Grade 5 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



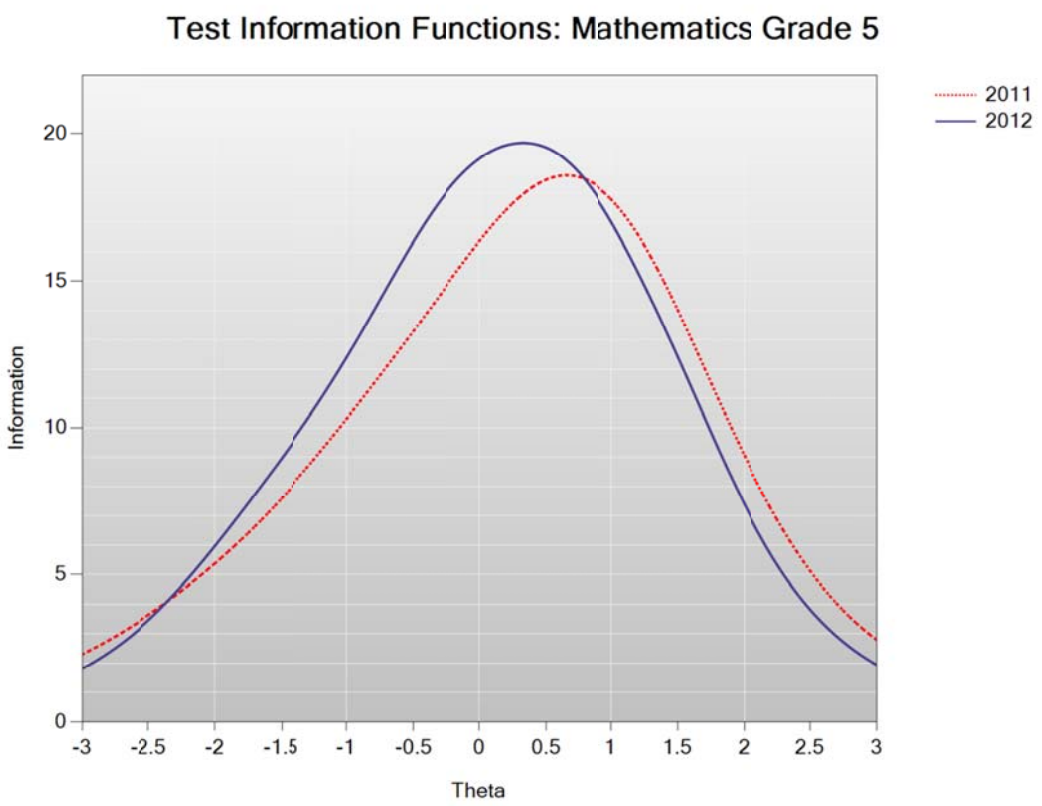
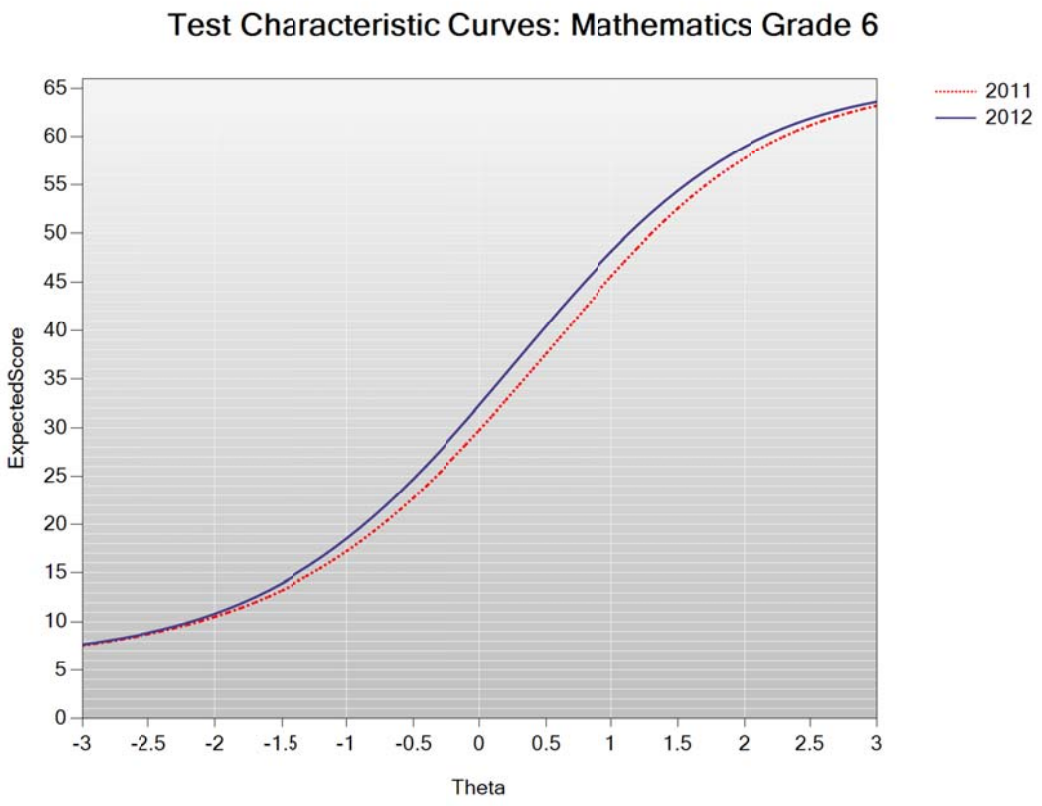


Figure I-4. 2012–13 NECAP: Mathematics Grade 6 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



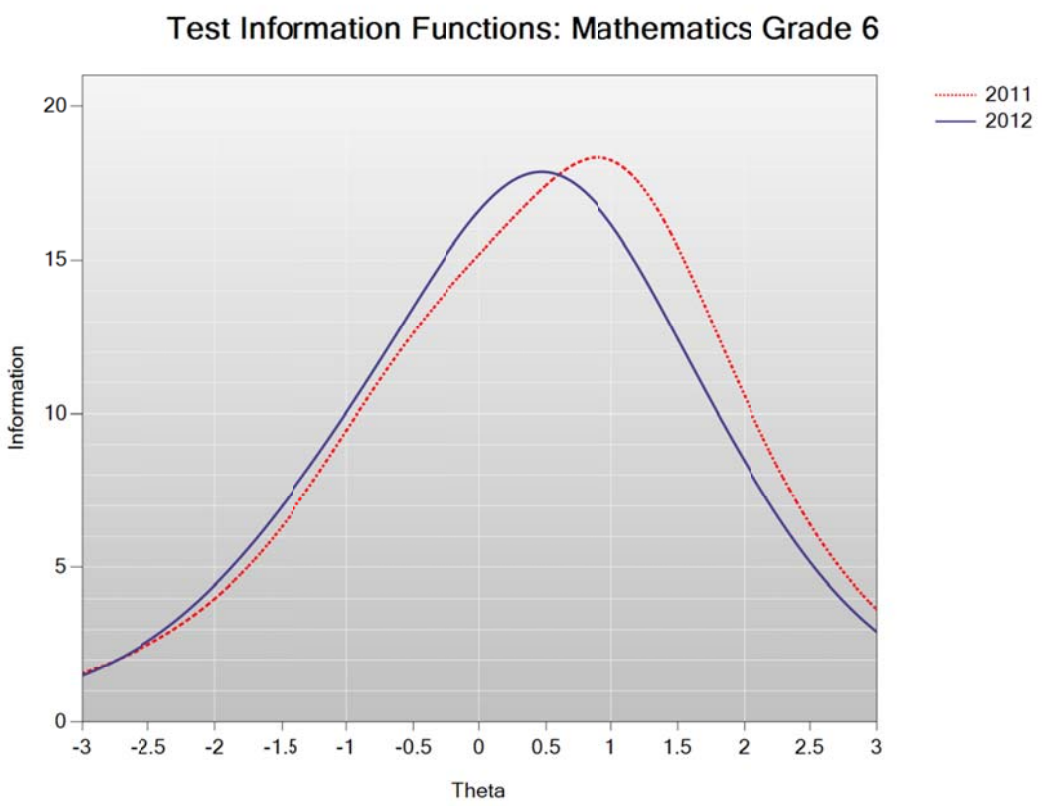
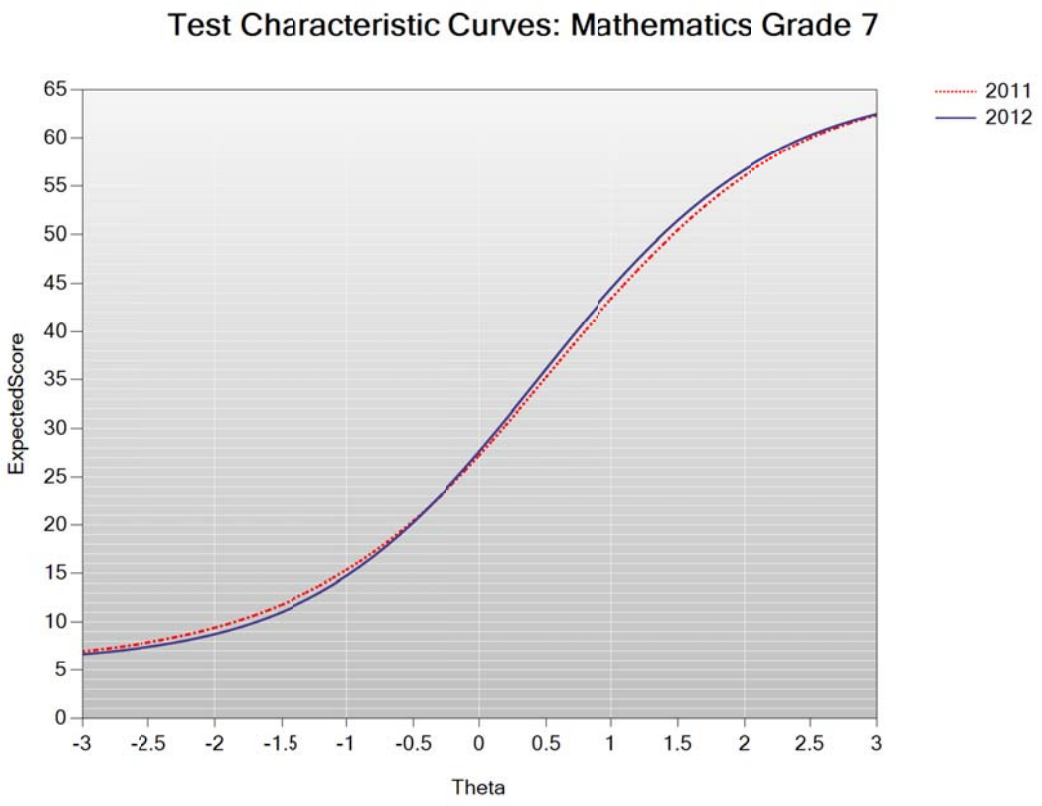


Figure I-5. 2012–13 NECAP: Mathematics Grade 7 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



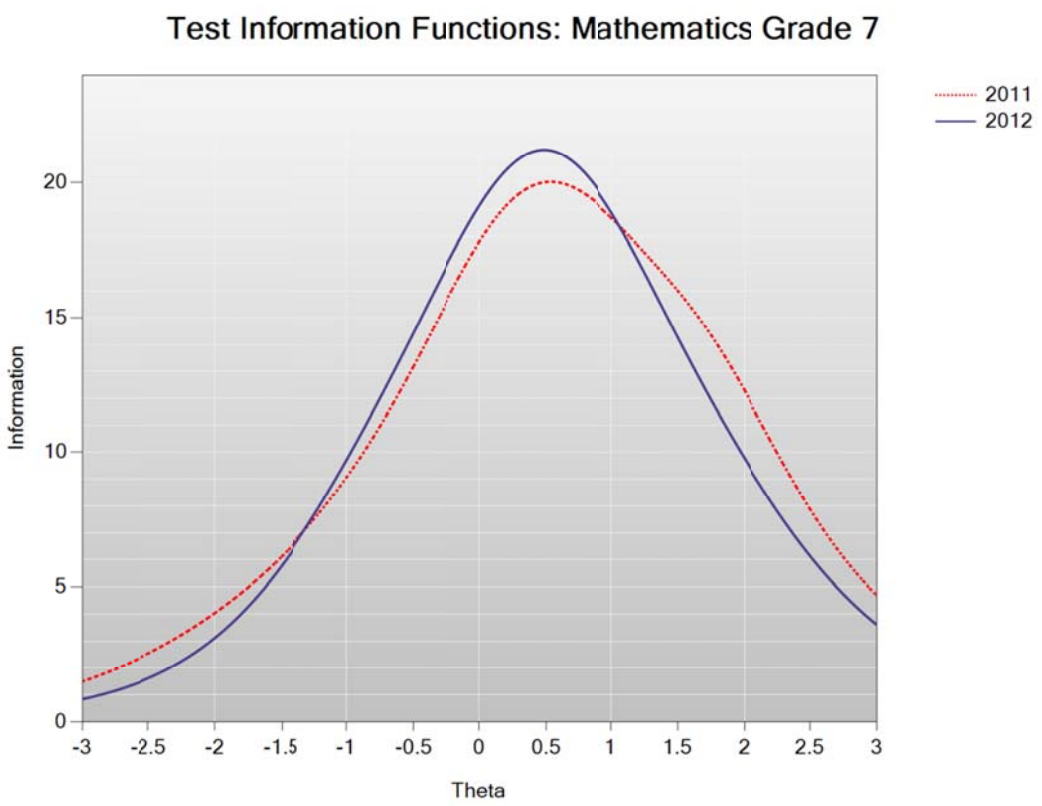
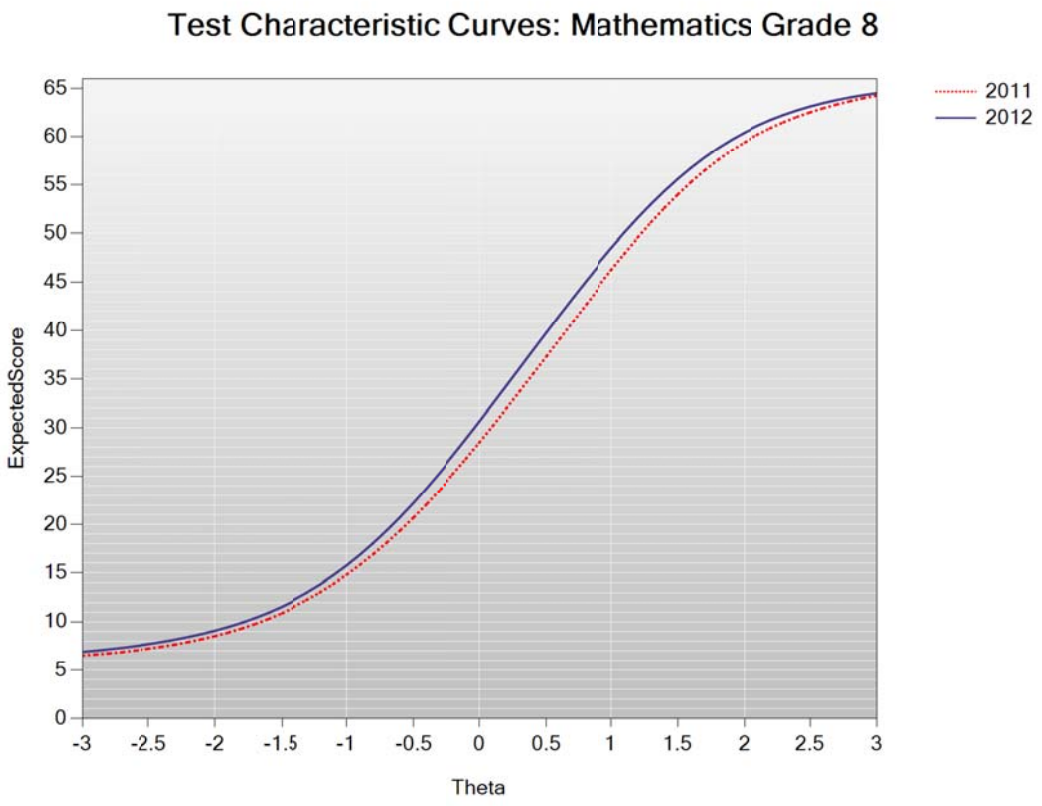


Figure I-6. 2012–13 NECAP: Mathematics Grade 8 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



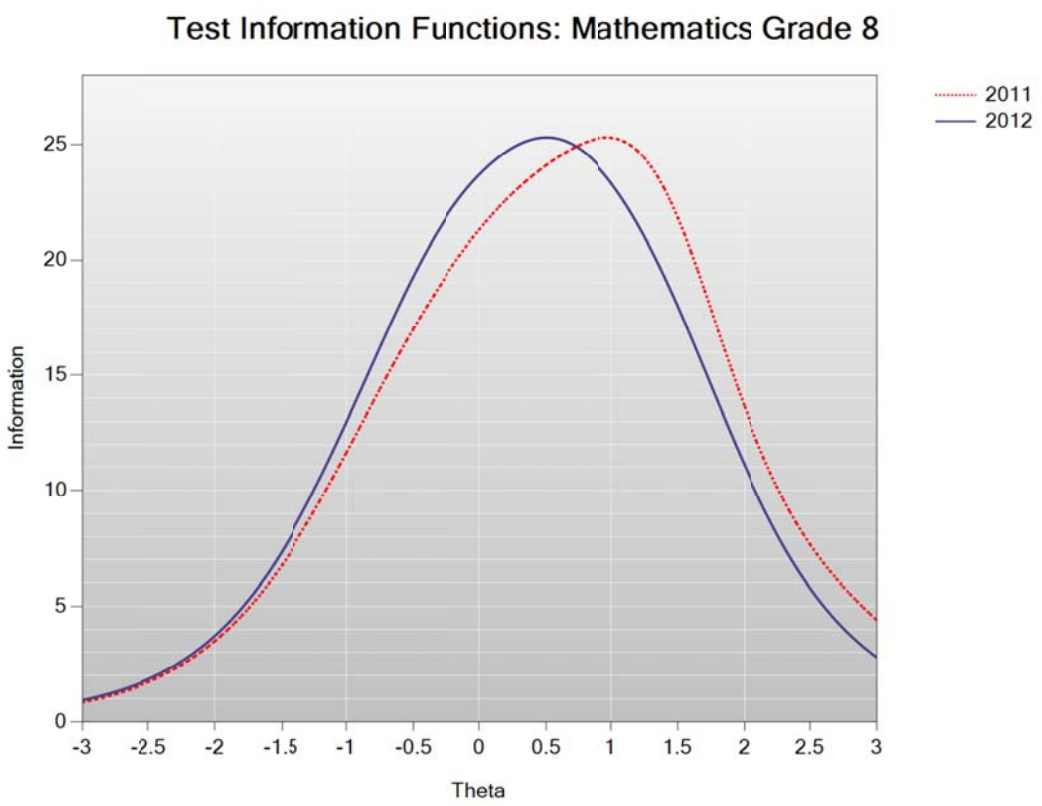
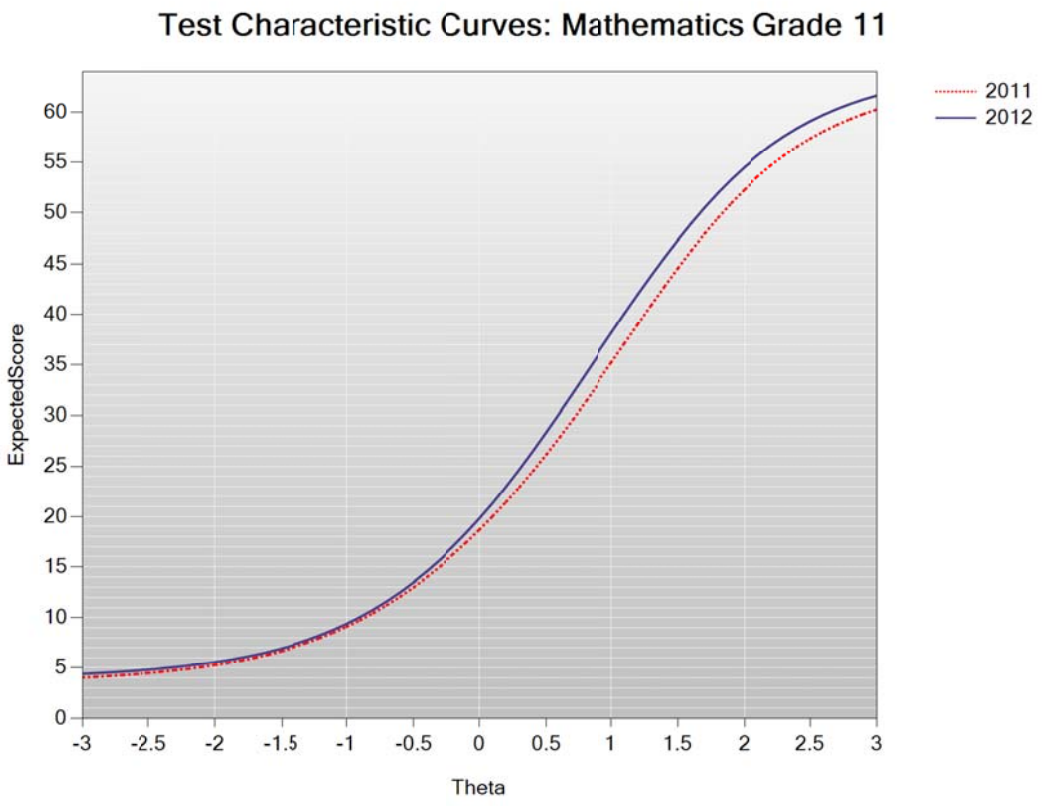


Figure I-7. 2012–13 NECAP: Mathematics Grade 11 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



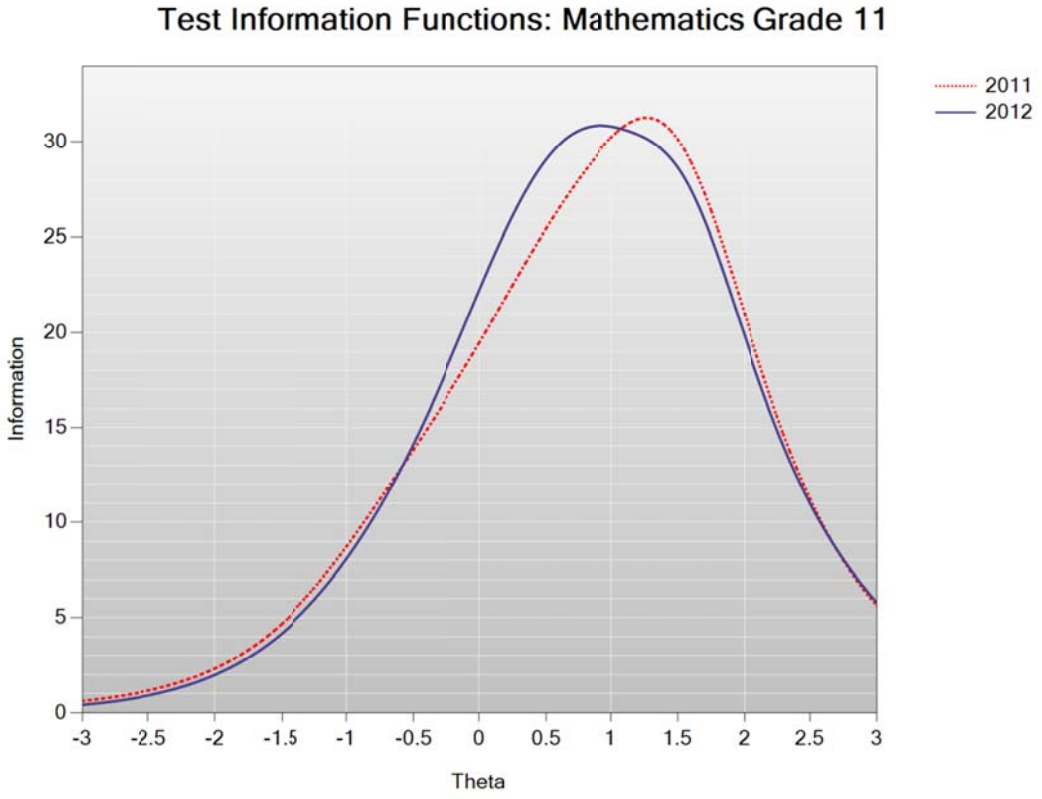
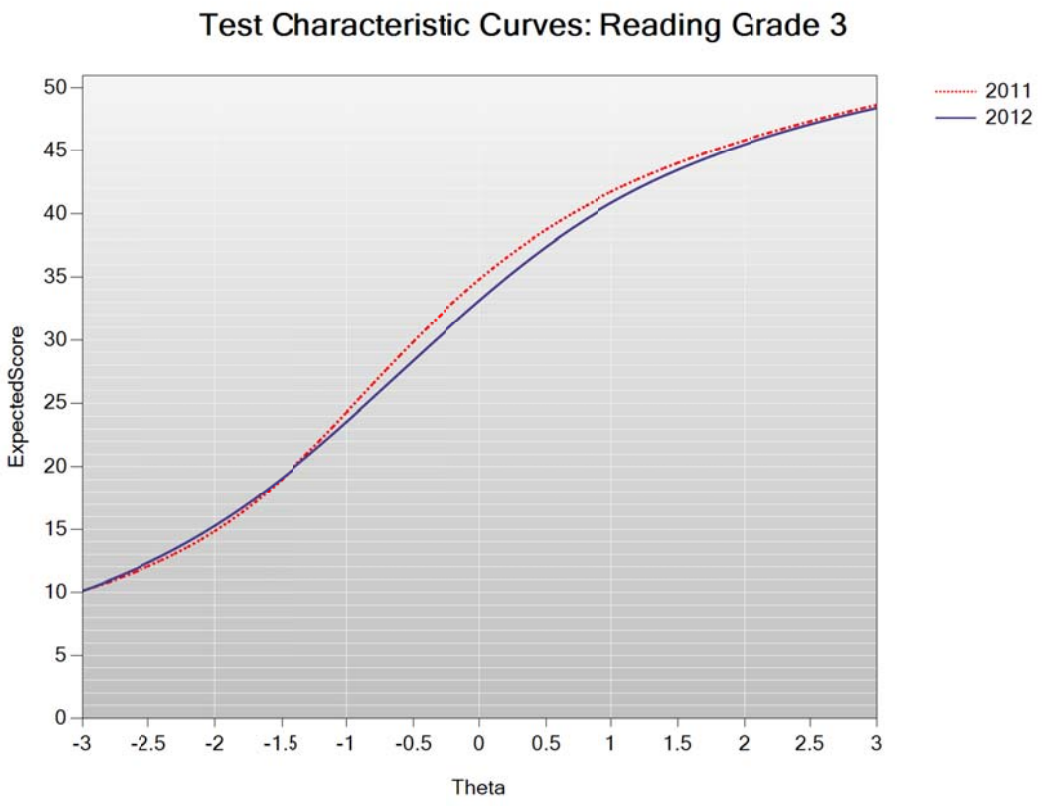


Figure I-8. 2012–13 NECAP: Reading Grade 3 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



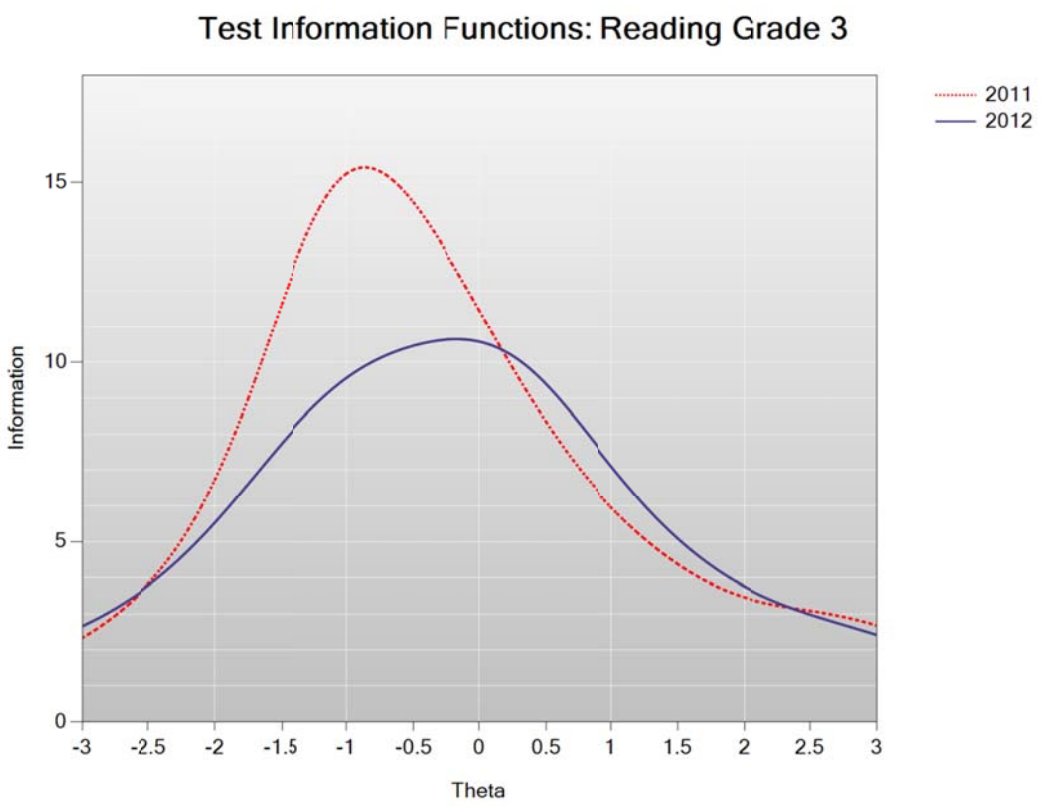
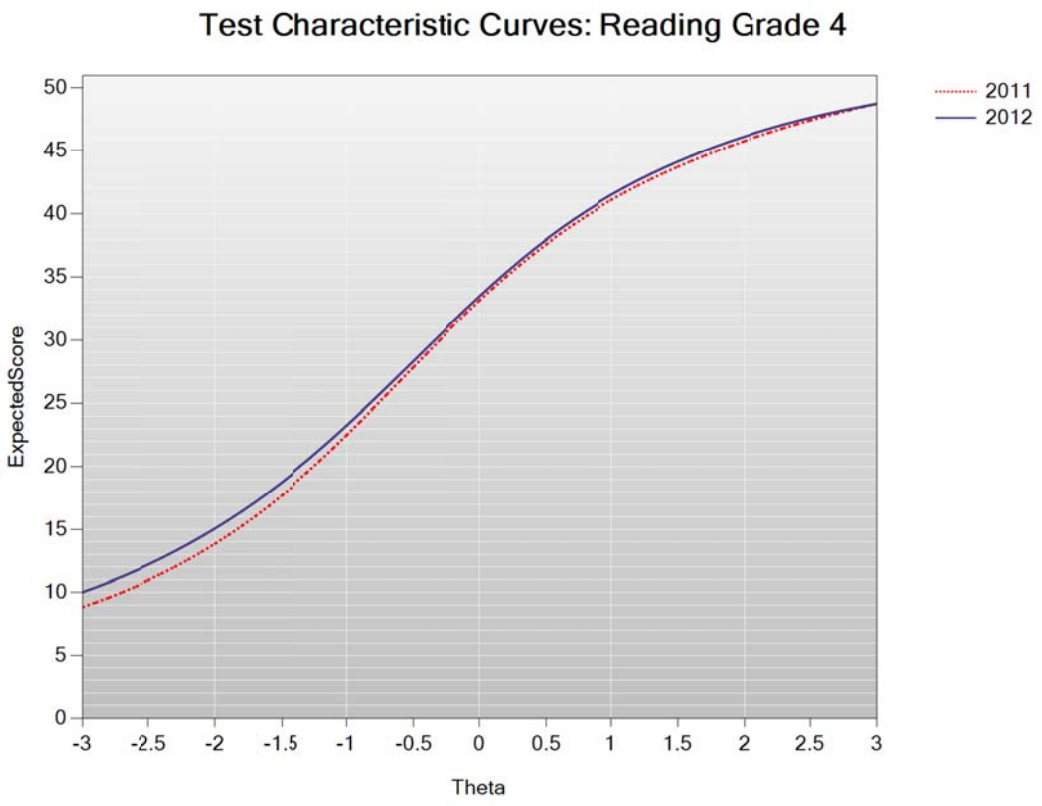


Figure I-9. 2012–13 NECAP: Reading Grade 4 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



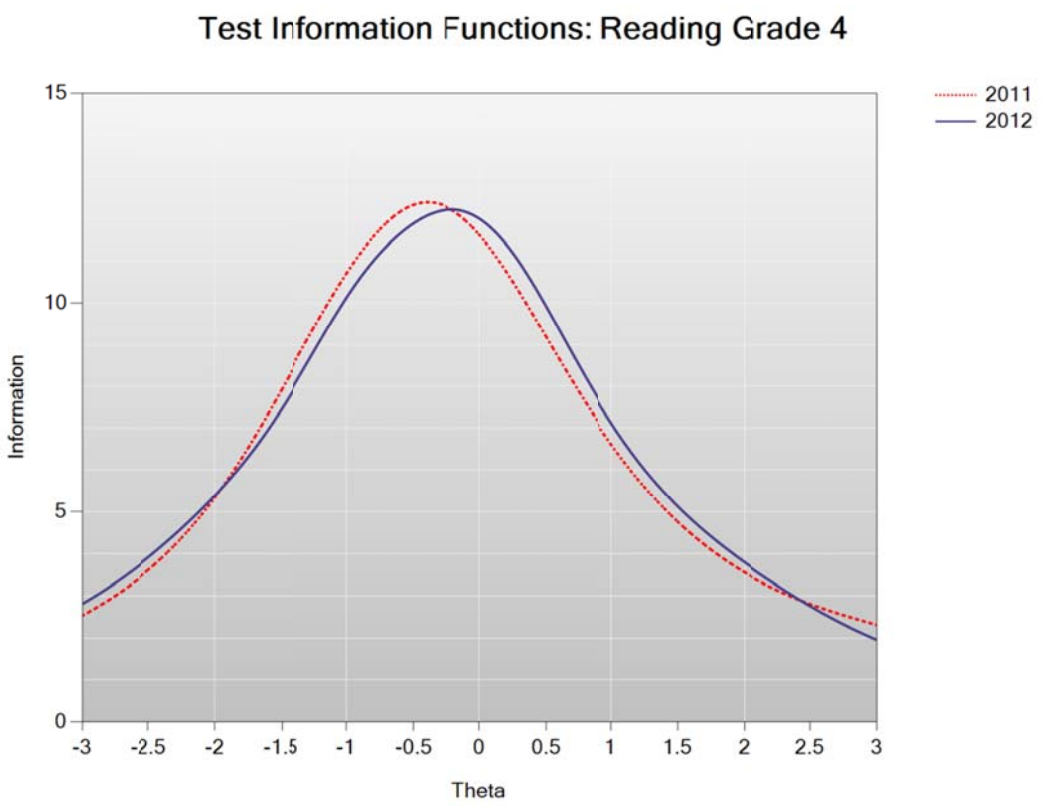
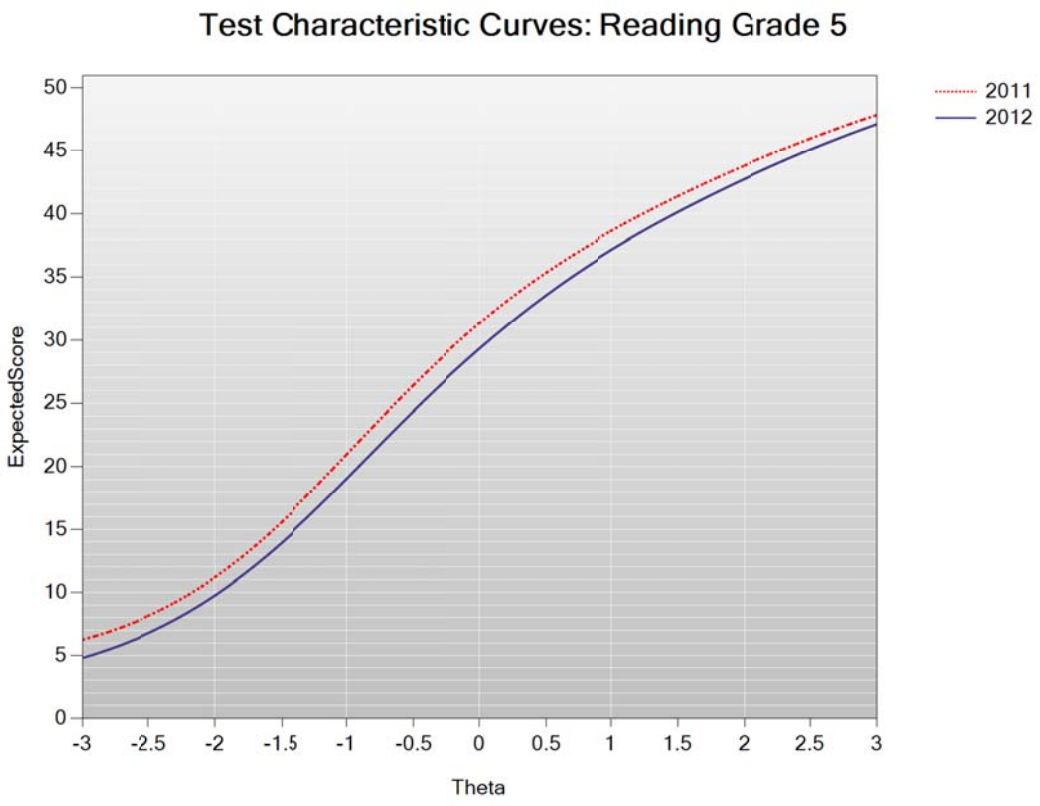


Figure I-10. 2012–13 NECAP: Reading Grade 5 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



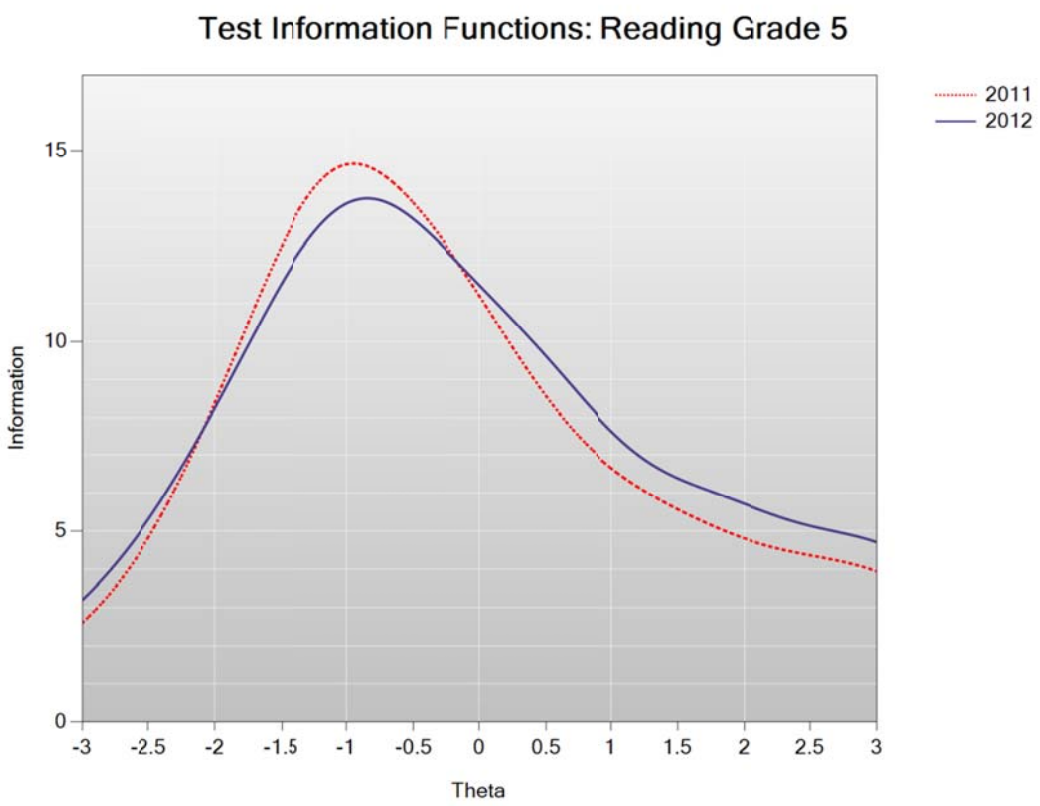
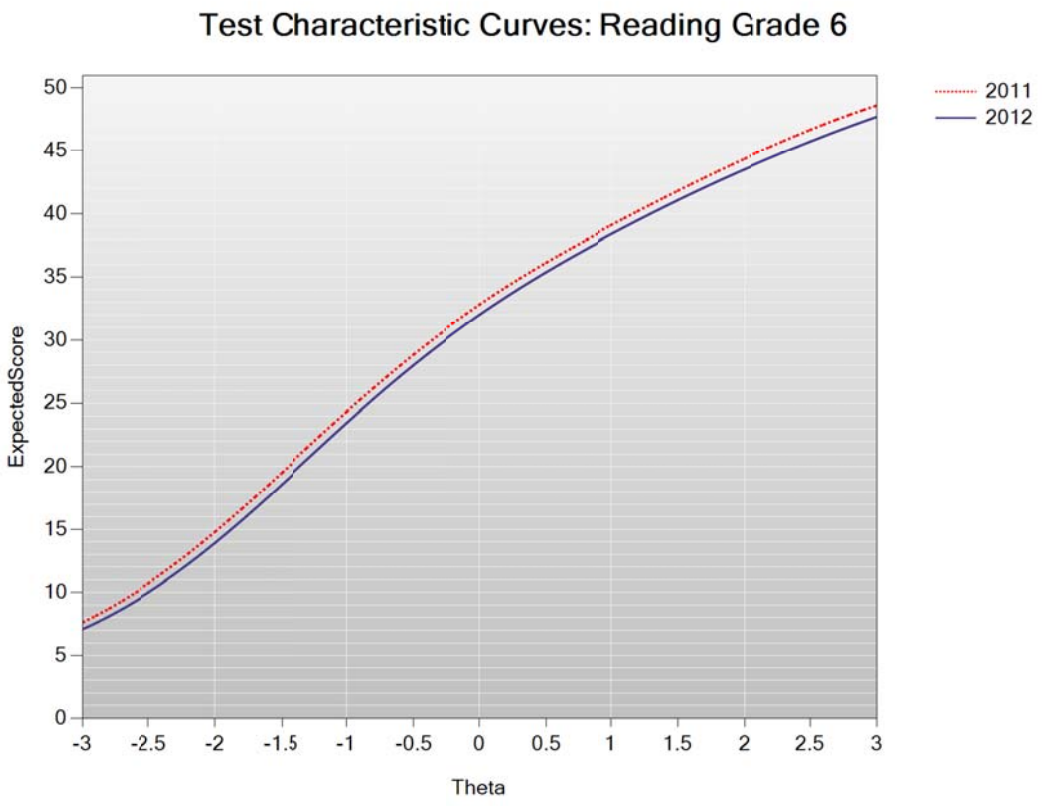


Figure I-11. 2012–13 NECAP: Reading Grade 6 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



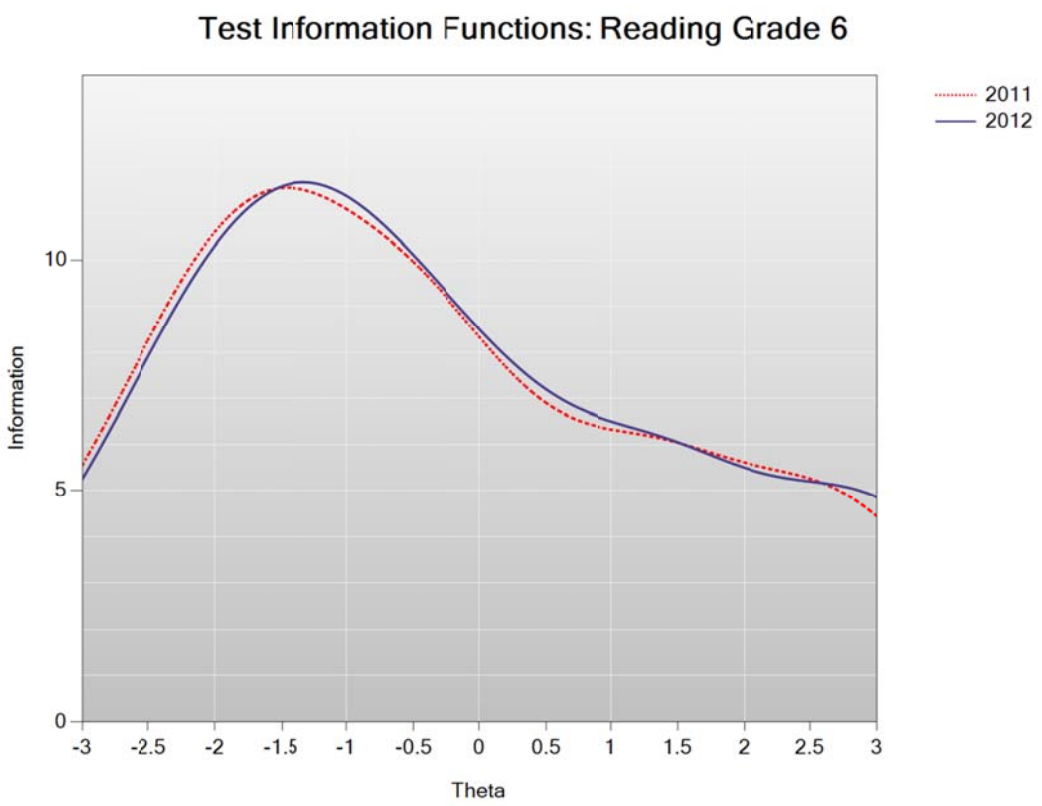
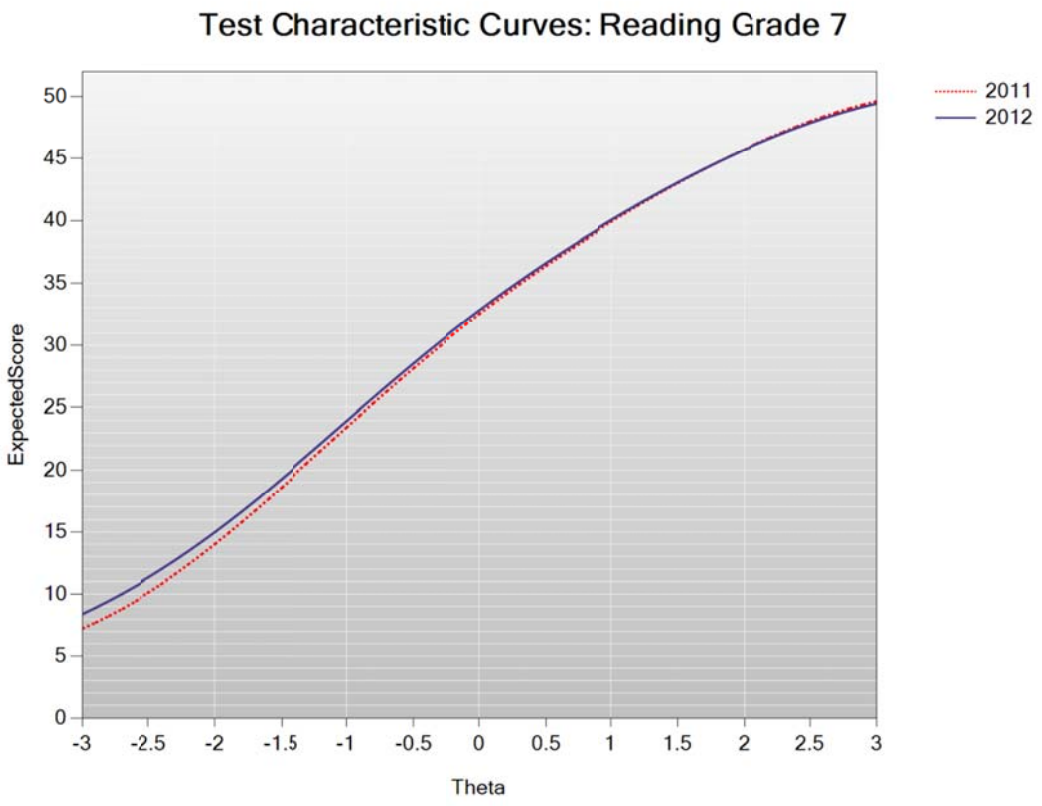


Figure I-12. 2012–13 NECAP: Reading Grade 7 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



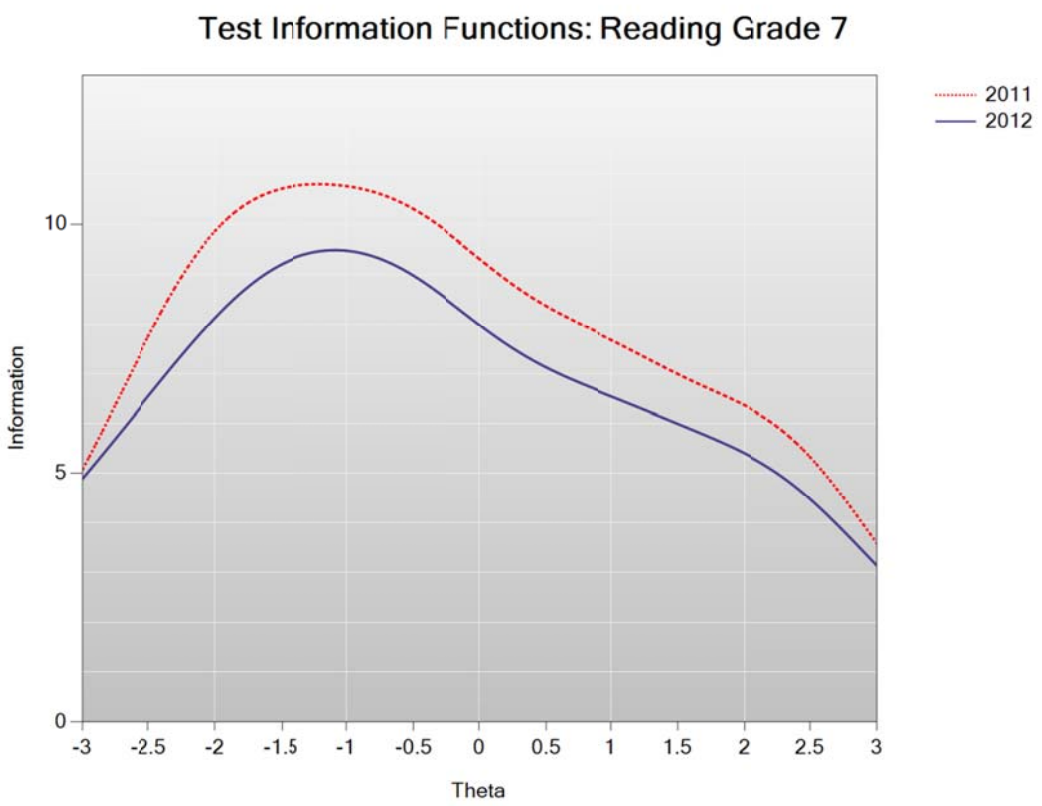
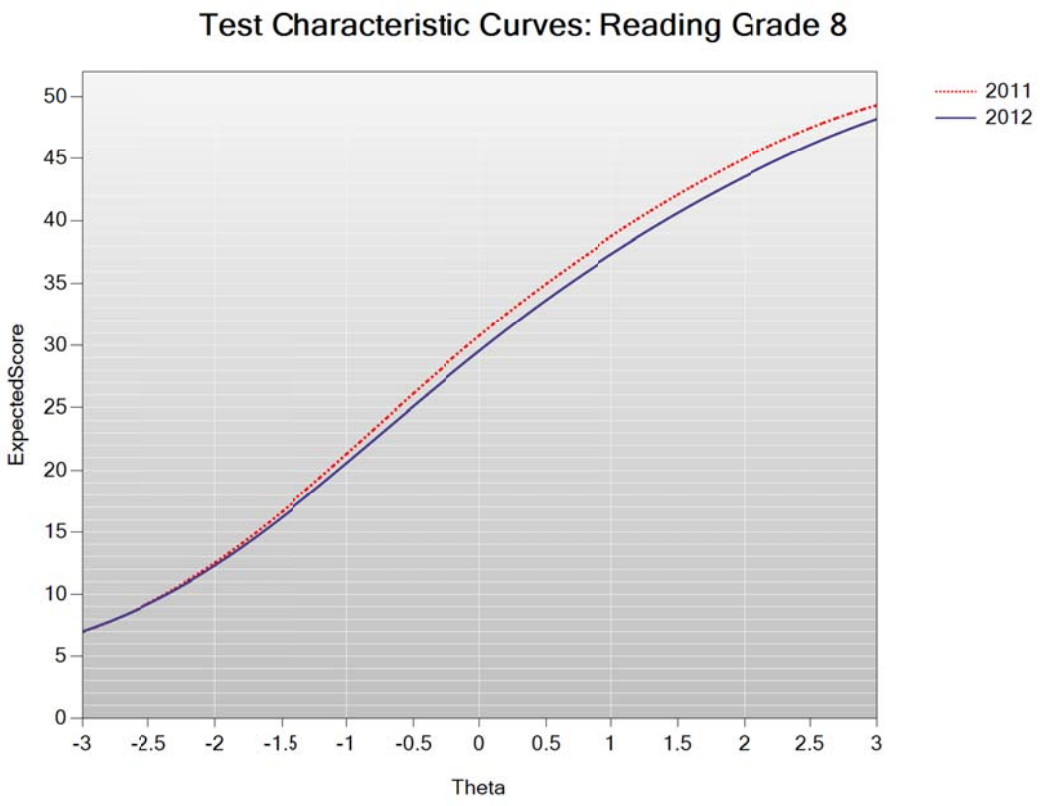


Figure I-13. 2012–13 NECAP: Reading Grade 8 Plots
Top: Test Characteristic Curve Bottom: Test Information Function



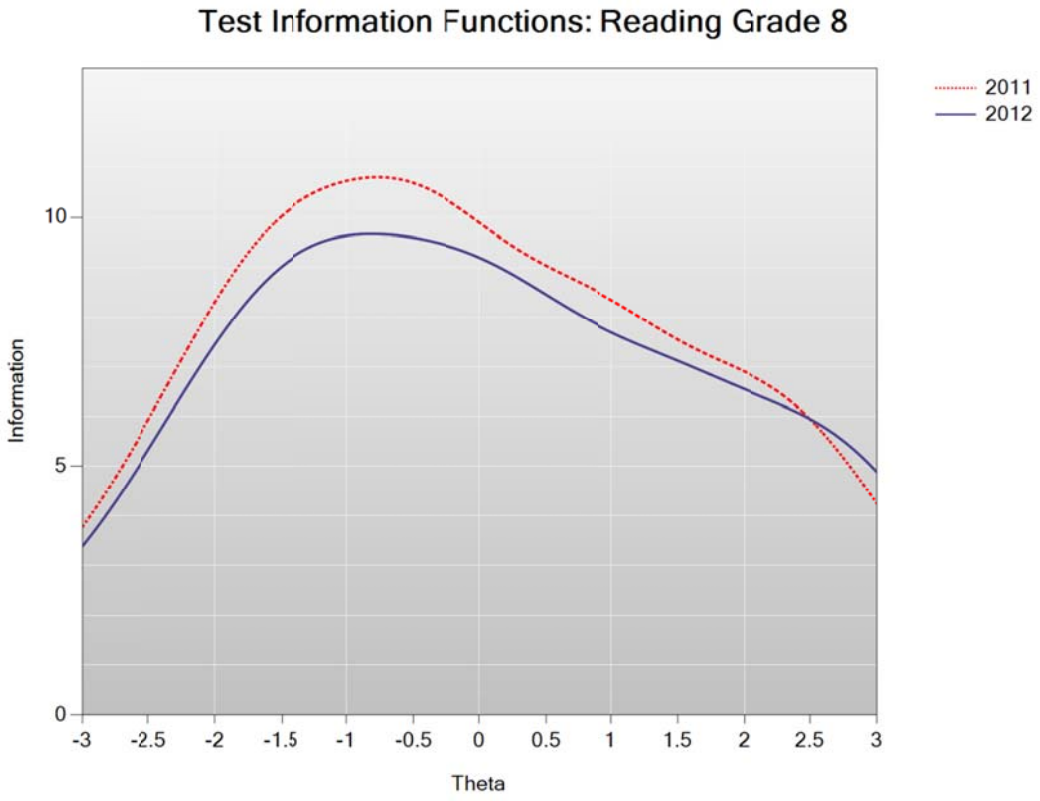
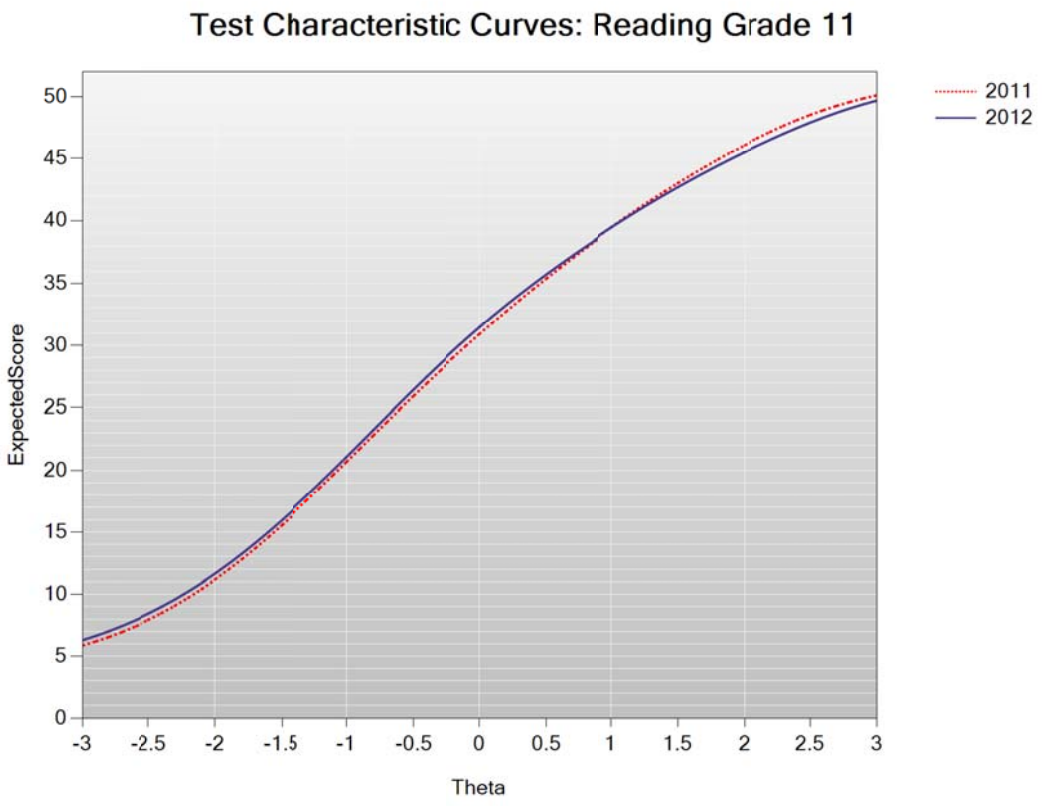
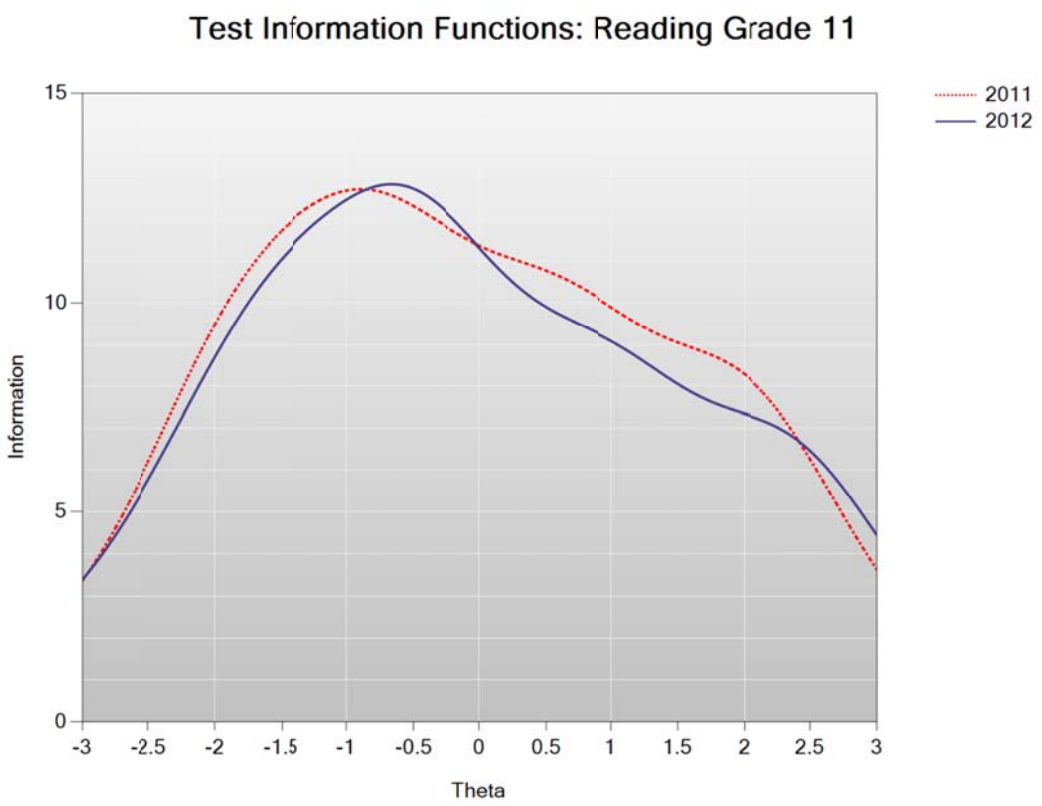


Figure I-14. 2012–13 NECAP: Reading Grade 11 Plots
Top: Test Characteristic Curve Bottom: Test Information Function





APPENDIX J—DELTA ANALYSES AND RESCORE ANALYSES

Figure J-1. 2012–13 NECAP: Delta Plot
Mathematics Grade 3

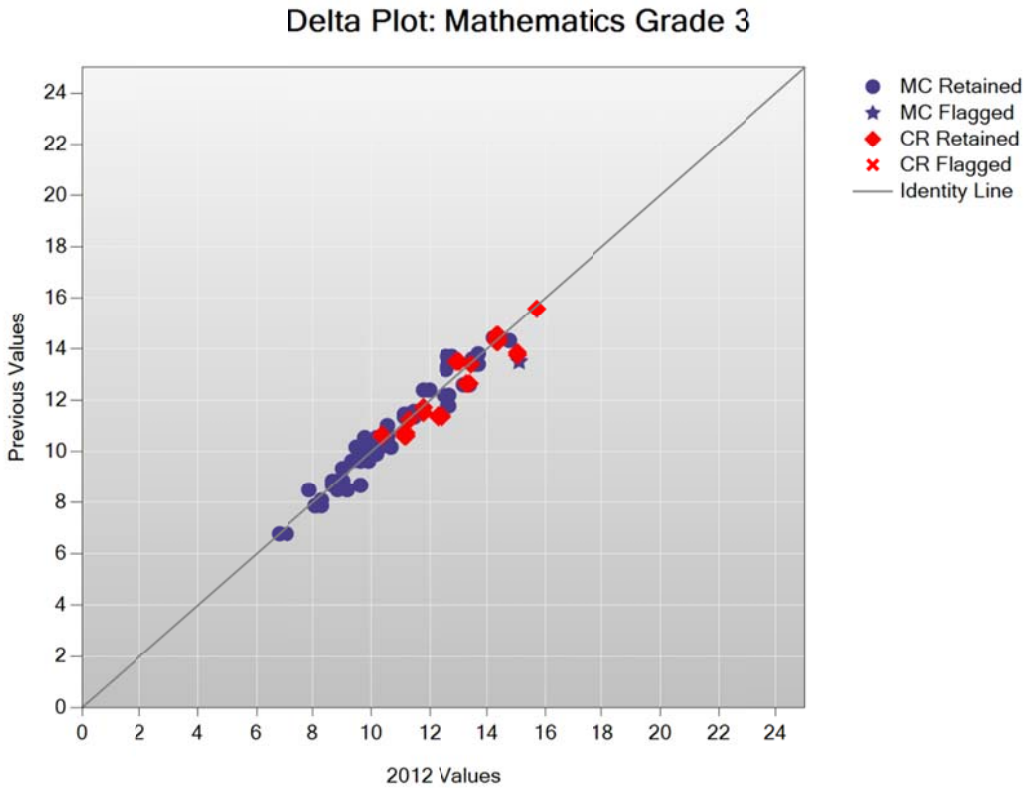


Table J-1. 2012–13 NECAP Delta Analysis
Mathematics Grade 3

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
119681	0.90000	0.89000	7.87379	8.09389	False	-0.63973
119681	0.90000	0.89000	7.87379	8.09389	False	-0.63973
119681	0.90000	0.89000	7.87379	8.09389	False	-0.63973
119681	0.90000	0.89000	7.87379	8.09389	False	-0.63973
119685	0.80000	0.82000	9.63352	9.33854	False	-0.14342
119708	0.71500	0.67500	10.72779	11.18495	False	-0.15711
119708	0.72500	0.67500	10.60896	11.18495	False	0.20888
119708	0.72500	0.68000	10.60896	11.12920	False	0.04166
119708	0.71500	0.68000	10.72779	11.12920	False	-0.32433
119729	0.78000	0.76000	9.91123	10.17479	False	-0.67246
119729	0.78000	0.76000	9.91123	10.17479	False	-0.67246
119752	0.87000	0.85000	8.49444	8.85427	False	-0.27024
119752	0.87000	0.83000	8.49444	9.18334	False	0.71689
119772	0.85000	0.84000	8.85427	9.02217	False	-0.87479
119775	0.65000	0.68000	11.45872	11.12920	False	0.10632
119775	0.66000	0.65000	11.35015	11.45872	False	-1.21651
119775	0.66000	0.68000	11.35015	11.12920	False	-0.22806
119775	0.65000	0.65000	11.45872	11.45872	False	-0.88213

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
119810	0.78000	0.77000	9.91123	10.04461	False	-1.06296
119810	0.78000	0.78000	9.91123	9.91123	False	-1.00604
119832	0.66000	0.67000	11.35015	11.24035	False	-0.56146
119832	0.65000	0.67000	11.45872	11.24035	False	-0.22708
119869	0.58000	0.53000	12.19243	12.69892	False	-0.12637
119869	0.58000	0.55000	12.19243	12.49735	False	-0.73102
119877	0.89000	0.88000	8.09389	8.30005	False	-0.69913
119882	0.73000	0.79000	10.54875	9.77432	False	1.36810
119907	0.46000	0.45500	13.40173	13.45215	False	-0.87781
119909	0.42500	0.30500	13.75647	15.04029	False	2.08017
119909	0.41500	0.30500	13.85881	15.04029	False	1.76501
119916	0.43000	0.52000	13.70550	12.79939	False	2.01585
119916	0.43000	0.54000	13.70550	12.59827	False	2.61916
119918	0.62500	0.62000	11.72544	11.77808	False	-1.01867
119918	0.64500	0.62000	11.51258	11.77808	False	-0.79486
119930	0.48000	0.54000	13.20061	12.59827	False	1.06422
119930	0.47000	0.54000	13.30108	12.59827	False	1.37363
119930	0.48000	0.54000	13.20061	12.59827	False	1.06422
119930	0.47000	0.54000	13.30108	12.59827	False	1.37363
122862	0.63000	0.62000	11.67259	11.77808	False	-1.18145
122981	0.53500	0.46000	12.64862	13.40173	False	0.57689
122981	0.53500	0.47000	12.64862	13.30108	False	0.27495
124292	0.76000	0.81000	10.17479	9.48841	False	1.07401
124292	0.76000	0.81000	10.17479	9.48841	False	1.07401
124490	0.45000	0.49500	13.50265	13.05013	False	0.63893
124490	0.45000	0.50500	13.50265	12.94987	False	0.93971
139548	0.76000	0.72000	10.17479	10.66863	False	-0.00278
139560	0.90000	0.88000	7.87379	8.30005	False	-0.02128
139572	0.44000	0.45000	13.60388	13.50265	False	-0.40671
139580	0.72000	0.73000	10.66863	10.54875	False	-0.58577
139673	0.67000	0.66500	11.24035	11.29541	False	-1.06479
139677	0.73000	0.73000	10.54875	10.54875	False	-0.95499
139677	0.75000	0.73000	10.30204	10.54875	False	-0.75431
139689	0.54000	0.46000	12.59827	13.40173	False	0.73198
139689	0.54000	0.48000	12.59827	13.20061	False	0.12867
144609	0.64000	0.63000	11.56616	11.67259	False	-1.19277
144609	0.64000	0.65000	11.56616	11.45872	False	-0.55122
144611	0.69000	0.73000	11.01660	10.54875	False	0.48590
144612	0.26000	0.25000	15.57338	15.69796	False	-0.92639
144617	0.82000	0.84000	9.33854	9.02217	False	-0.10286
145244	0.37000	0.33000	14.32741	14.75965	False	-0.52006
145244	0.37000	0.33000	14.32741	14.75965	False	-0.52006
145259	0.37500	0.37000	14.27456	14.32741	False	-0.81524
145259	0.35000	0.36500	14.54128	14.38050	False	-0.15303
145259	0.35000	0.37000	14.54128	14.32741	False	0.00622
145259	0.37500	0.36500	14.27456	14.38050	False	-0.97449
145504	0.36500	0.37000	14.38050	14.32741	False	-0.48895
145504	0.36500	0.36500	14.38050	14.38050	False	-0.64820
145526	0.73000	0.76000	10.54875	10.17479	False	0.16678

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
145526	0.73000	0.74000	10.54875	10.42662	False	-0.58864
145672	0.94000	0.93000	6.78091	7.09684	False	-0.26474
145672	0.94000	0.94000	6.78091	6.78091	False	-1.21245
145720	0.78000	0.78000	9.91123	9.91123	False	-1.00604
145729	0.42000	0.43000	13.80757	13.70550	False	-0.38787
168404	0.45000	0.30000	13.50265	15.09760	True	3.03382
168404	0.45000	0.30000	13.50265	15.09760	True	3.03382
168521	0.77000	0.78000	10.04461	9.91123	False	-0.59524
168521	0.77000	0.79000	10.04461	9.77432	False	-0.18453
168602	0.56000	0.60000	12.39612	11.98661	False	0.42135
168602	0.56000	0.62000	12.39612	11.77808	False	1.04690
198283	0.86000	0.85000	8.67872	8.85427	False	-0.83781
198283	0.86000	0.86000	8.67872	8.67872	False	-1.10472
198551	0.85000	0.84000	8.85427	9.02217	False	-0.87479
201619	0.86000	0.80000	8.67872	9.63352	False	1.49973
201619	0.86000	0.80000	8.67872	9.63352	False	1.49973
201811	0.75000	0.75000	10.30204	10.30204	False	-0.97475
201811	0.76000	0.75000	10.17479	10.30204	False	-1.10246
203044	0.80000	0.78000	9.63352	9.91123	False	-0.60778
203044	0.80000	0.80000	9.63352	9.63352	False	-1.02827
203231	0.62000	0.54000	11.77808	12.59827	False	0.84780
203231	0.62000	0.53000	11.77808	12.69892	False	1.14974
209937	0.36000	0.38000	14.43384	14.22192	False	-0.00825
223883	0.76000	0.79000	10.17479	9.77432	False	0.21639
223883	0.76000	0.80000	10.17479	9.63352	False	0.63875
231019	0.72500	0.74500	10.60896	10.36465	False	-0.21730
231019	0.72000	0.74500	10.66863	10.36465	False	-0.03352
242311	0.65500	0.55500	11.40458	12.44678	False	1.54369
242311	0.66000	0.57000	11.35015	12.29450	False	1.25453
242311	0.66000	0.55500	11.35015	12.44678	False	1.71133
242311	0.65500	0.57000	11.40458	12.29450	False	1.08689
255679	0.85000	0.85000	8.85427	8.85427	False	-1.09066
255679	0.85000	0.86000	8.85427	8.67872	False	-0.56408
255932	0.46000	0.44000	13.40173	13.60388	False	-1.13618
255932	0.46000	0.43000	13.40173	13.70550	False	-0.83134
269219	0.87000	0.90000	8.49444	7.87379	False	0.74229

Table J-2. 2012–13 NECAP Rescore Analysis
Mathematics Grade 3

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
122981	2	1.18049	1.22439	0.91206	0.93155	0.04814	False
242311	2	1.29268	1.29756	0.67192	0.68062	0.00726	False
124490	2	0.98049	0.99512	0.84997	0.81798	0.01722	False
119708	2	1.38049	1.39024	0.71380	0.72216	0.01367	False
145504	2	0.80976	0.83902	0.73821	0.75168	0.03965	False
145259	2	0.75122	0.74634	0.80993	0.79930	-0.00602	False
119909	2	0.78537	0.60000	0.81066	0.79436	-0.22866	False

continued

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
144612	2	0.56098	0.53171	0.86299	0.82998	-0.03391	False
119907	2	0.88780	0.93659	0.86280	0.87336	0.05654	False
231019	2	1.46341	1.47805	0.77499	0.76922	0.01888	False
139673	2	1.36585	1.36098	0.62336	0.63009	-0.00783	False
119918	2	1.37745	1.38235	0.87483	0.87551	0.00560	False

Figure J-2. 2012–13 NECAP: Delta Plot
Mathematics Grade 4

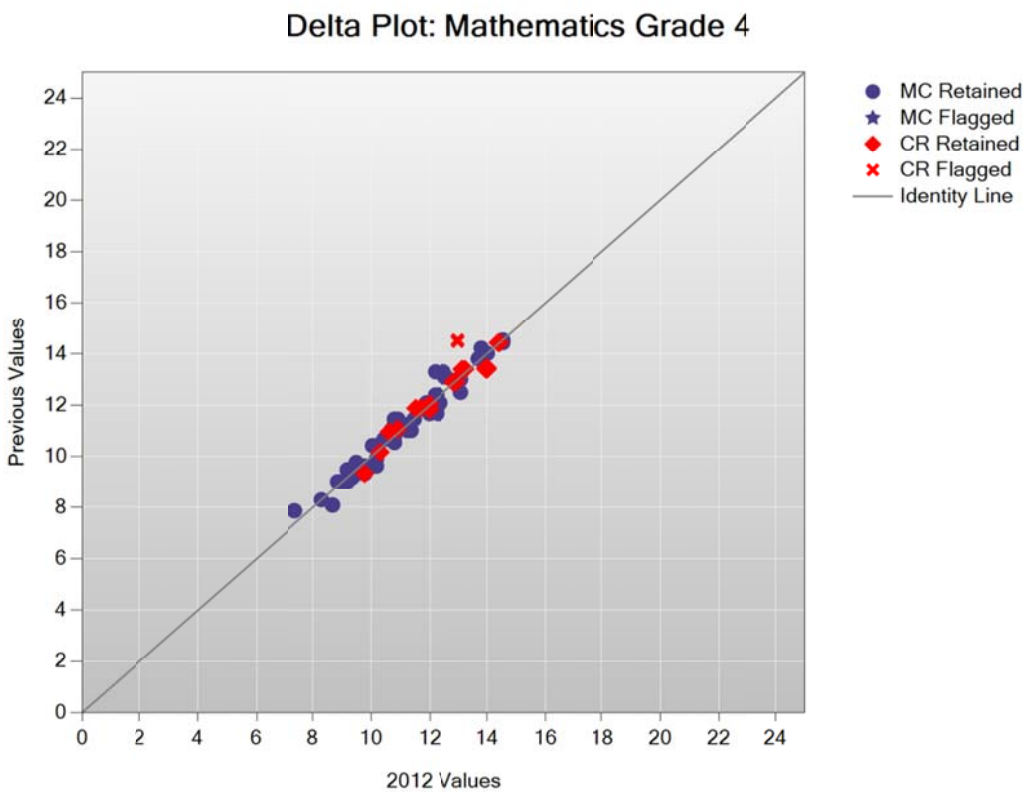


Table J-3. 2012–13 NECAP Delta Analysis
Mathematics Grade 4

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
120028	0.83000	0.82000	9.18334	9.33854	False	-0.69366
120028	0.83000	0.82000	9.18334	9.33854	False	-0.69366
120028	0.83000	0.82000	9.18334	9.33854	False	-0.69366
120028	0.83000	0.82000	9.18334	9.33854	False	-0.69366
120072	0.84000	0.84000	9.02217	9.02217	False	-0.87806
120074	0.67000	0.67000	11.24035	11.24035	False	-0.82585
120074	0.67000	0.67000	11.24035	11.24035	False	-0.82585
120175	0.50000	0.50000	13.00000	13.00000	False	-0.46602

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
120175	0.50000	0.49000	13.00000	13.10028	False	-0.10883
120197	0.59000	0.61000	12.08982	11.88272	False	-0.76768
120197	0.59000	0.61000	12.08982	11.88272	False	-0.76768
120209	0.80000	0.78000	9.63352	9.91123	False	-0.16521
120209	0.80000	0.76000	9.63352	10.17479	False	0.77362
120220	0.84000	0.85000	9.02217	8.85427	False	-0.27997
120224	0.81000	0.83000	9.48841	9.18334	False	0.11330
120253	0.84000	0.84000	9.02217	9.02217	False	-0.87806
120253	0.84000	0.83000	9.02217	9.18334	False	-0.70536
120253	0.84000	0.84000	9.02217	9.02217	False	-0.87806
120253	0.84000	0.83000	9.02217	9.18334	False	-0.70536
120261	0.81000	0.81000	9.48841	9.48841	False	-0.97340
120261	0.81000	0.81000	9.48841	9.48841	False	-0.97340
120267	0.65000	0.71000	11.45872	10.78646	False	1.01832
120267	0.65000	0.70000	11.45872	10.90240	False	0.60534
120267	0.67000	0.70000	11.24035	10.90240	False	-0.12786
120267	0.67000	0.71000	11.24035	10.78646	False	0.28512
120274	0.81000	0.80000	9.48841	9.63352	False	-0.66726
120274	0.81000	0.79000	9.48841	9.77432	False	-0.16572
120290	0.68500	0.70000	11.07309	10.90240	False	-0.68943
124700	0.59500	0.60000	12.03830	11.98661	False	-0.84678
139424	0.72000	0.74000	10.66863	10.42662	False	-0.35267
139426	0.40000	0.40000	14.01339	14.01339	False	-0.25878
139432	0.65000	0.65000	11.45872	11.45872	False	-0.78120
139455	0.63000	0.60000	11.67259	11.98661	False	0.38112
139455	0.63000	0.57000	11.67259	12.29450	False	1.47785
139455	0.61000	0.57000	11.88272	12.29450	False	0.77230
139455	0.61000	0.60000	11.88272	11.98661	False	-0.32444
139465	0.69000	0.72000	11.01660	10.66863	False	-0.04643
139477	0.82000	0.79000	9.33854	9.77432	False	0.33750
139477	0.82000	0.79000	9.33854	9.77432	False	0.33750
139477	0.80000	0.79000	9.63352	9.77432	False	-0.65290
139477	0.80000	0.79000	9.63352	9.77432	False	-0.65290
139482	0.61500	0.60000	11.83050	11.98661	False	-0.14909
139517	0.47000	0.58000	13.30108	12.19243	False	2.19604
139517	0.47000	0.55000	13.30108	12.49735	False	1.10986
144635	0.76000	0.75000	10.17479	10.30204	False	-0.59048
144637	0.36000	0.36500	14.43384	14.38050	False	-0.36278
144637	0.36000	0.36000	14.43384	14.43384	False	-0.17280
144639	0.61000	0.64500	11.88272	11.51258	False	-0.14452
144639	0.61000	0.63000	11.88272	11.67259	False	-0.71450
144641	0.89000	0.86000	8.09389	8.67872	False	0.61394
144641	0.89000	0.86000	8.09389	8.67872	False	0.61394
144645	0.35000	0.50000	14.54128	13.00000	True	3.48349
144645	0.35000	0.50000	14.54128	13.00000	True	3.48349
144645	0.35500	0.50000	14.48742	13.00000	True	3.30266
144645	0.35500	0.50000	14.48742	13.00000	True	3.30266
145070	0.90000	0.92000	7.87379	7.37971	False	1.11673
145070	0.90000	0.92000	7.87379	7.37971	False	1.11673

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
145087	0.88000	0.88000	8.30005	8.30005	False	-0.73039
145268	0.81000	0.81000	9.48841	9.48841	False	-0.97340
145281	0.74000	0.76000	10.42662	10.17479	False	-0.26822
145281	0.74000	0.77000	10.42662	10.04461	False	0.19548
145281	0.74000	0.76000	10.42662	10.17479	False	-0.26822
145281	0.74000	0.77000	10.42662	10.04461	False	0.19548
145290	0.51000	0.50500	12.89972	12.94987	False	-0.30791
145290	0.51000	0.51500	12.89972	12.84957	False	-0.66519
145545	0.78000	0.76000	9.91123	10.17479	False	-0.15882
145780	0.71000	0.71000	10.78646	10.78646	False	-0.91867
145788	0.79000	0.81000	9.77432	9.48841	False	-0.01347
145803	0.80000	0.81000	9.63352	9.48841	False	-0.48621
145815	0.49000	0.55000	13.10028	12.49735	False	0.43565
145815	0.49000	0.54000	13.10028	12.59827	False	0.07620
145851	0.36000	0.35000	14.43384	14.54128	False	0.20993
145851	0.35000	0.35000	14.54128	14.54128	False	-0.15083
168384	0.42000	0.43000	13.80757	13.70550	False	-0.66448
168395	0.68000	0.70000	11.12920	10.90240	False	-0.50103
168395	0.68000	0.70000	11.12920	10.90240	False	-0.50103
168410	0.73000	0.71000	10.54875	10.78646	False	-0.12053
168439	0.59000	0.56000	12.08982	12.39612	False	0.43893
168483	0.46000	0.48500	13.40173	13.15043	False	-0.87849
168483	0.46000	0.47500	13.40173	13.25083	False	-0.92141
169080	0.81000	0.82000	9.48841	9.33854	False	-0.43953
169090	0.38000	0.42000	14.22192	13.80757	False	-0.46543
169095	0.82000	0.81000	9.33854	9.48841	False	-0.68090
169095	0.82000	0.81000	9.33854	9.48841	False	-0.68090
169105	0.69000	0.67000	11.01660	11.24035	False	-0.07460
169105	0.69000	0.66000	11.01660	11.35015	False	0.31652
169109	0.55000	0.49000	12.49735	13.10028	False	1.57885
169117	0.46000	0.40000	13.40173	14.01339	False	1.79490
169117	0.46000	0.40500	13.40173	13.96170	False	1.61079
169117	0.45500	0.40000	13.45215	14.01339	False	1.62561
169117	0.45500	0.40500	13.45215	13.96170	False	1.44151
198442	0.82000	0.79000	9.33854	9.77432	False	0.33750
201375	0.69000	0.66000	11.01660	11.35015	False	0.31652
201375	0.69000	0.67000	11.01660	11.24035	False	-0.07460
227082	0.69500	0.72500	10.95971	10.60896	False	-0.02488
227082	0.69500	0.72500	10.95971	10.60896	False	-0.02488
232535	0.56000	0.58000	12.39612	12.19243	False	-0.84242
232535	0.56000	0.57000	12.39612	12.29450	False	-0.95148
232535	0.56000	0.58000	12.39612	12.19243	False	-0.84242
232535	0.56000	0.57000	12.39612	12.29450	False	-0.95148
255685	0.83000	0.83000	9.18334	9.18334	False	-0.91101
255694	0.68000	0.68000	11.12920	11.12920	False	-0.84858

Table J-4. 2012–13 NECAP Rescore Analysis
Mathematics Grade 4

IREF	*Max*	*Mean*		*St Dev*		*Effect Size*	*Discard*
		Old	*New*	*Old*	*New*		
169117	2	0.93171	0.91220	0.84166	0.82814	-0.02318	False
168483	2	0.93171	0.98049	0.92462	0.92165	0.05276	False
144637	2	0.68780	0.70244	0.60126	0.60462	0.02434	False
145290	2	1.15610	1.19512	0.78759	0.80352	0.04955	False
144639	2	1.21463	1.20976	0.80461	0.80287	-0.00606	False
144645	2	0.70732	0.68780	0.90700	0.88405	-0.02151	False
144635	2	1.60000	1.59024	0.75009	0.76422	-0.01301	False
198442	2	1.65854	1.61463	0.61760	0.63498	-0.07109	False
139482	2	1.16176	1.15686	0.77906	0.78320	-0.00629	False
124700	2	1.21463	1.19512	0.52557	0.51448	-0.03713	False
120290	2	1.37561	1.38537	0.77216	0.77367	0.01263	False
227082	2	1.34146	1.33659	0.57671	0.59207	-0.00846	False

Figure J-3. 2012–13 NECAP: Delta Plot
Mathematics Grade 5

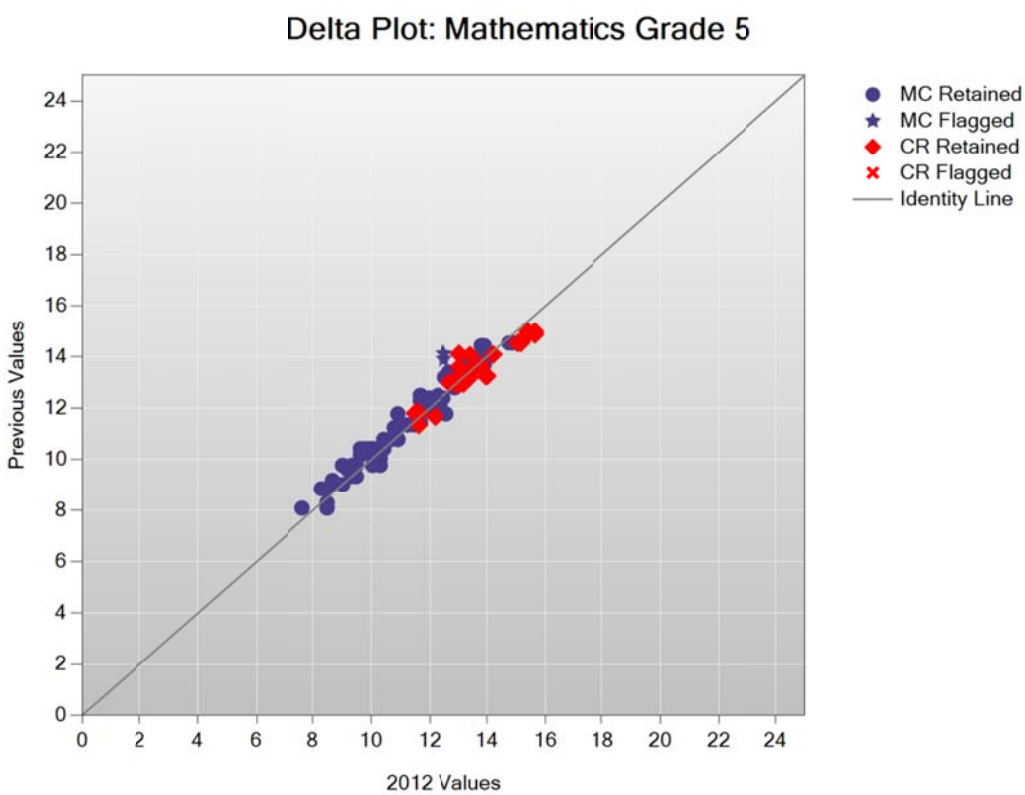


Table J-5. 2012–13 NECAP Delta Analysis
Mathematics Grade 5

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
120640	0.82000	0.81000	9.33854	9.48841	False	0.33209
120640	0.82000	0.82000	9.33854	9.33854	False	-0.19159
120642	0.84000	0.84000	9.02217	9.02217	False	-0.12369
120645	0.85000	0.87000	8.85427	8.49444	False	-1.00332
120645	0.85000	0.88000	8.85427	8.30005	False	-0.32412
120648	0.89000	0.91000	8.09389	7.63698	False	-0.82731
120650	0.43000	0.41000	13.70550	13.91018	False	-0.41361
120658	0.79000	0.82000	9.77432	9.33854	False	-0.54051
120658	0.79000	0.81000	9.77432	9.48841	False	-1.06419
120667	0.40000	0.42000	14.01339	13.80757	False	-0.43426
120667	0.36000	0.42000	14.43384	13.80757	False	1.12507
120667	0.36000	0.41000	14.43384	13.91018	False	0.76655
120667	0.40000	0.41000	14.01339	13.91018	False	-0.79278
120686	0.57000	0.60000	12.29450	11.98661	False	-0.44649
120686	0.56000	0.60000	12.39612	11.98661	False	-0.06960
120704	0.62000	0.63500	11.77808	11.61950	False	-1.07904
120704	0.62000	0.64500	11.77808	11.51258	False	-0.70544
120704	0.61500	0.63500	11.83050	11.61950	False	-0.88461
120704	0.61500	0.64500	11.83050	11.51258	False	-0.51101
120724	0.65000	0.63000	11.45872	11.67259	False	0.10068
120725	0.50000	0.53000	13.00000	12.69892	False	-0.31888
120725	0.50000	0.52500	13.00000	12.74917	False	-0.49447
120727	0.74000	0.74000	10.42662	10.42662	False	-0.42511
120728	0.66000	0.67000	11.35015	11.24035	False	-1.00697
120736	0.48000	0.55000	13.20061	12.49735	False	1.12945
120751	0.50750	0.50250	12.92480	12.97493	False	-0.78606
120799	0.89000	0.87000	8.09389	8.49444	False	1.47510
120799	0.88000	0.87000	8.30005	8.49444	False	0.71048
120803	0.75000	0.78000	10.30204	9.91123	False	-0.58435
120803	0.75000	0.75000	10.30204	10.30204	False	-0.39837
120807	0.52000	0.51000	12.79939	12.89972	False	-0.58374
120807	0.52000	0.51000	12.79939	12.89972	False	-0.58374
120819	0.77000	0.76000	10.04461	10.17479	False	0.11173
120819	0.77000	0.75000	10.04461	10.30204	False	0.55637
120822	0.67000	0.71000	11.24035	10.78646	False	-0.16260
120822	0.67000	0.71000	11.24035	10.78646	False	-0.16260
120822	0.67000	0.70000	11.24035	10.90240	False	-0.56770
120822	0.67000	0.70000	11.24035	10.90240	False	-0.56770
120838	0.62000	0.57000	11.77808	12.29450	False	1.08932
120838	0.62000	0.54000	11.77808	12.59827	False	2.15071
120846	0.48250	0.46250	13.17552	13.37655	False	-0.31263
120846	0.48250	0.46000	13.17552	13.40173	False	-0.22463
121791	0.39250	0.38000	14.09124	14.22192	False	-0.75494
121805	0.59000	0.57000	12.08982	12.29450	False	-0.06686
121805	0.59000	0.56000	12.08982	12.39612	False	0.28821
124768	0.66000	0.65000	11.35015	11.45872	False	-0.24395
124768	0.66000	0.67000	11.35015	11.24035	False	-1.00697

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
124887	0.76000	0.80000	10.17479	9.63352	False	-0.08593
124887	0.74000	0.77000	10.42662	10.04461	False	-0.58839
124887	0.74000	0.80000	10.42662	9.63352	False	0.84804
124887	0.76000	0.77000	10.17479	10.04461	False	-0.82592
139141	0.44000	0.44000	13.60388	13.60388	False	-1.10699
139181	0.71000	0.70000	10.78646	10.90240	False	-0.09723
139181	0.71000	0.70000	10.78646	10.90240	False	-0.09723
139210	0.35000	0.30500	14.54128	15.04029	False	0.43545
139210	0.35000	0.29500	14.54128	15.15534	False	0.83745
139267	0.39500	0.45750	14.06524	13.42694	False	1.08805
139338	0.74000	0.78000	10.42662	9.91123	False	-0.12233
139338	0.74000	0.77000	10.42662	10.04461	False	-0.58839
145293	0.42000	0.41000	13.80757	13.91018	False	-0.79218
145293	0.42000	0.43000	13.80757	13.70550	False	-0.84091
145576	0.43500	0.42000	13.65463	13.80757	False	-0.58349
145576	0.43500	0.42000	13.65463	13.80757	False	-0.58349
145585	0.43000	0.48000	13.70550	13.20061	False	0.54464
145585	0.43000	0.47000	13.70550	13.30108	False	0.19361
145877	0.35000	0.33000	14.54128	14.75965	False	-0.54515
145877	0.35000	0.32000	14.54128	14.87080	False	-0.15680
145888	0.63000	0.58000	11.67259	12.19243	False	1.12388
145888	0.63000	0.58000	11.67259	12.19243	False	1.12388
145893	0.79000	0.75000	9.77432	10.30204	False	1.55883
145893	0.79000	0.77000	9.77432	10.04461	False	0.65934
145896	0.61000	0.62000	11.88272	11.77808	False	-1.10326
150816	0.39000	0.49500	14.11728	13.05013	False	2.59763
167664	0.55000	0.57000	12.49735	12.29450	False	-0.76998
167664	0.55000	0.57000	12.49735	12.29450	False	-0.76998
167682	0.49000	0.49000	13.10028	13.10028	False	-0.99891
167693	0.46000	0.53000	13.40173	12.69892	False	1.17106
167740	0.60000	0.61000	11.98661	11.88272	False	-1.12290
167762	0.50500	0.48000	12.94987	13.20061	False	-0.09048
167762	0.50500	0.49250	12.94987	13.07520	False	-0.52868
167875	0.75000	0.79000	10.30204	9.77432	False	-0.10596
167875	0.75000	0.77000	10.30204	10.04461	False	-1.05042
167875	0.74000	0.79000	10.42662	9.77432	False	0.35606
167875	0.74000	0.77000	10.42662	10.04461	False	-0.58839
167886	0.47500	0.40500	13.25083	13.96170	False	1.45268
167886	0.47500	0.40000	13.25083	14.01339	False	1.63328
203941	0.79000	0.84000	9.77432	9.02217	False	0.56494
203977	0.56000	0.55000	12.39612	12.49735	False	-0.49407
203978	0.62000	0.70000	11.77808	10.90240	False	1.42661
203997	0.39000	0.55000	14.11728	12.49735	True	4.52912
203997	0.41000	0.55000	13.91018	12.49735	True	3.76105
204291	0.71000	0.74000	10.78646	10.42662	False	-0.58861
204291	0.71000	0.72000	10.78646	10.66863	False	-0.91404
204719	0.80000	0.83000	9.63352	9.18334	False	-0.52041
204936	0.44000	0.49500	13.60388	13.05013	False	0.69356
225389	0.32000	0.27500	14.87080	15.39104	False	0.43892

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
225389	0.31000	0.25500	14.98340	15.63535	False	0.87495
225389	0.31000	0.27500	14.98340	15.39104	False	0.02130
225389	0.32000	0.25500	14.87080	15.63535	False	1.29258
225421	0.57000	0.63000	12.29450	11.67259	False	0.65076
225421	0.55000	0.63000	12.49735	11.67259	False	1.40309
255127	0.84000	0.86000	9.02217	8.67872	False	-1.02454
255127	0.83000	0.86000	9.18334	8.67872	False	-0.42680
255766	0.50000	0.51500	13.00000	12.84957	False	-0.84526
255766	0.50000	0.51000	13.00000	12.89972	False	-1.02052
269140	0.85000	0.87000	8.85427	8.49444	False	-1.00332
269376	0.66000	0.63500	11.35015	11.61950	False	0.31784
272113	0.48500	0.48000	13.15043	13.20061	False	-0.83433
272113	0.47000	0.48000	13.30108	13.20061	False	-0.95524

Table J-6. 2012–13 NECAP Rescore Analysis
Mathematics Grade 5

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
167762	4	2.09804	2.10784	1.47608	1.43834	0.00664	False
120846	4	2.07353	2.00000	1.25286	1.28782	-0.05869	False
145576	4	1.83902	1.82927	1.37587	1.37469	-0.00709	False
139267	4	1.55610	1.79024	1.19530	1.18992	0.19589	False
121791	4	1.49756	1.49756	1.21684	1.22882	0.00000	False
120751	4	2.11220	2.14634	1.66891	1.66772	0.02046	False

Figure J-4. 2012–13 NECAP: Delta Plot
Mathematics Grade 6

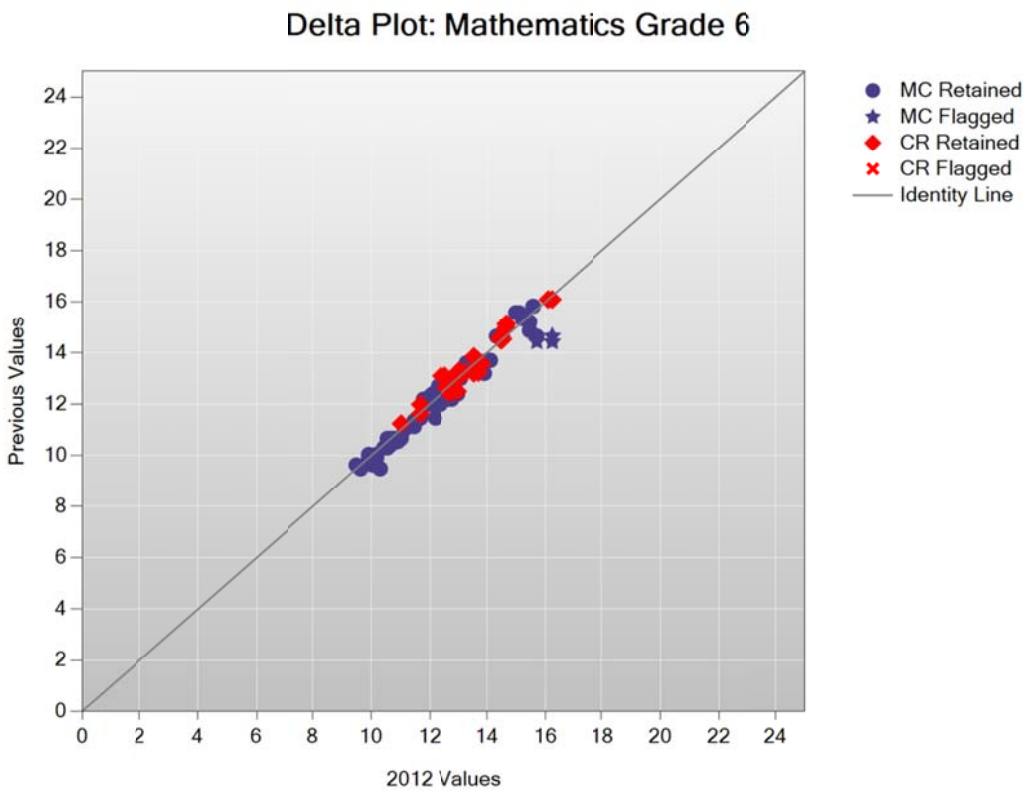


Table J-7. 2012–13 NECAP Delta Analysis
Mathematics Grade 6

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
119176	0.78000	0.78000	9.91123	9.91123	False	-0.73310
119176	0.77000	0.78000	10.04461	9.91123	False	-0.24948
119176	0.77000	0.76000	10.04461	10.17479	False	-0.94899
119176	0.78000	0.76000	9.91123	10.17479	False	-0.46537
119178	0.48000	0.49000	13.20061	13.10028	False	-0.41097
119183	0.72000	0.73000	10.66863	10.54875	False	-0.30676
119183	0.72000	0.72000	10.66863	10.66863	False	-0.74300
119184	0.81000	0.75000	9.48841	10.30204	False	1.53067
119184	0.81000	0.75000	9.48841	10.30204	False	1.53067
119184	0.81000	0.80000	9.48841	9.63352	False	-0.90196
119184	0.81000	0.80000	9.48841	9.63352	False	-0.90196
119193	0.78000	0.76000	9.91123	10.17479	False	-0.46537
119200	0.68000	0.65000	11.12920	11.45872	False	-0.20948
119200	0.66000	0.65000	11.35015	11.45872	False	-1.01056
119214	0.55000	0.56000	12.49735	12.39612	False	-0.39853
119214	0.56000	0.56000	12.39612	12.39612	False	-0.76557
119215	0.54000	0.53000	12.59827	12.69892	False	-1.02305
119215	0.54000	0.53000	12.59827	12.69892	False	-1.02305

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
119230	0.67000	0.69000	11.24035	11.01660	False	0.06371
119274	0.72000	0.71000	10.66863	10.78646	False	-0.98578
119274	0.72000	0.69000	10.66863	11.01660	False	-0.14836
119288	0.55000	0.53000	12.49735	12.69892	False	-0.65718
119288	0.55000	0.56000	12.49735	12.39612	False	-0.39853
119289	0.53500	0.54500	12.64862	12.54785	False	-0.40217
119289	0.53500	0.54500	12.64862	12.54785	False	-0.40217
119306	0.75000	0.73000	10.30204	10.54875	False	-0.52160
119306	0.75000	0.74000	10.30204	10.42662	False	-0.96601
119315	0.50000	0.52000	13.00000	12.79939	False	-0.04346
119315	0.51000	0.52000	12.89972	12.79939	False	-0.40704
119315	0.51000	0.54000	12.89972	12.59827	False	0.32480
119315	0.50000	0.54000	13.00000	12.59827	False	0.68837
119328	0.29000	0.27000	15.21354	15.45125	False	-0.49016
119328	0.32000	0.27000	14.87080	15.45125	False	0.75254
119351	0.69000	0.68000	11.01660	11.12920	False	-1.00023
119354	0.44000	0.47000	13.60388	13.30108	False	0.32047
119364	0.28000	0.29000	15.33137	15.21354	False	-0.37517
119364	0.27000	0.29000	15.45125	15.21354	False	0.05950
119365	0.56000	0.59000	12.39612	12.08982	False	0.34901
119365	0.57000	0.59000	12.29450	12.08982	False	-0.01944
119369	0.55500	0.53000	12.44678	12.69892	False	-0.47382
119369	0.55500	0.52500	12.44678	12.74917	False	-0.29096
119381	0.63500	0.63000	11.61950	11.67259	False	-0.94860
119381	0.60000	0.63000	11.98661	11.67259	False	0.38245
119400	0.41500	0.44500	13.85881	13.55322	False	0.32730
123448	0.34000	0.35500	14.64985	14.48742	False	-0.20397
123448	0.35500	0.35500	14.48742	14.48742	False	-0.79289
123513	0.36000	0.21000	14.43384	16.22568	True	5.15483
123513	0.34000	0.21000	14.64985	16.22568	True	4.37161
123513	0.34000	0.25000	14.64985	15.69796	False	2.45133
123513	0.36000	0.25000	14.43384	15.69796	True	3.23455
125065	0.44500	0.41750	13.55322	13.83317	False	-0.35814
139219	0.74000	0.72000	10.42662	10.66863	False	-0.53705
139310	0.44000	0.42000	13.60388	13.80757	False	-0.63497
139310	0.44000	0.42000	13.60388	13.80757	False	-0.63497
139312	0.43000	0.40000	13.70550	14.01339	False	-0.25450
139312	0.43000	0.39000	13.70550	14.11728	False	0.12353
139333	0.48000	0.41000	13.20061	13.91018	False	1.20051
139333	0.48000	0.41000	13.20061	13.91018	False	1.20051
139358	0.80000	0.81000	9.63352	9.48841	False	-0.20148
139358	0.80000	0.77000	9.63352	10.04461	False	0.06785
139360	0.55000	0.52000	12.49735	12.79939	False	-0.29161
144698	0.57000	0.60000	12.29450	11.98661	False	0.35611
144698	0.60000	0.60000	11.98661	11.98661	False	-0.76022
144698	0.57000	0.56000	12.29450	12.39612	False	-1.02351
144698	0.60000	0.56000	11.98661	12.39612	False	0.09282
144710	0.60000	0.57000	11.98661	12.29450	False	-0.27695
145135	0.29750	0.33750	15.12642	14.67718	False	0.83343

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
145135	0.29750	0.34000	15.12642	14.64985	False	0.93288
145140	0.26000	0.31000	15.57338	14.98340	False	1.33973
145140	0.26000	0.30000	15.57338	15.09760	False	0.92418
145333	0.58000	0.52000	12.19243	12.79939	False	0.81398
145333	0.58000	0.58000	12.19243	12.19243	False	-0.76291
145352	0.65000	0.63000	11.45872	11.67259	False	-0.62598
145604	0.55000	0.53000	12.49735	12.69892	False	-0.65718
145604	0.58000	0.53000	12.19243	12.69892	False	0.44841
145630	0.55000	0.50000	12.49735	13.00000	False	0.43839
145632	0.35000	0.35000	14.54128	14.54128	False	-0.79360
145632	0.35000	0.36000	14.54128	14.43384	False	-0.40262
145988	0.53000	0.57000	12.69892	12.29450	False	0.70207
145988	0.53000	0.57000	12.69892	12.29450	False	0.70207
145988	0.54000	0.57000	12.59827	12.29450	False	0.33712
145988	0.54000	0.57000	12.59827	12.29450	False	0.33712
145995	0.50000	0.49000	13.00000	13.10028	False	-1.01919
146022	0.52000	0.52000	12.79939	12.79939	False	-0.77084
167817	0.49000	0.54500	13.10028	12.54785	False	1.23541
167817	0.49000	0.56000	13.10028	12.39612	False	1.78750
167923	0.65000	0.63000	11.45872	11.67259	False	-0.62598
167923	0.65000	0.58000	11.45872	12.19243	False	1.26560
167946	0.65000	0.64000	11.45872	11.56616	False	-1.01323
167950	0.73000	0.72000	10.54875	10.66863	False	-0.97985
167950	0.73000	0.70000	10.54875	10.90240	False	-0.12924
167962	0.43000	0.43000	13.70550	13.70550	False	-0.78268
167962	0.43000	0.43500	13.70550	13.65463	False	-0.59760
167963	0.47000	0.49500	13.30108	13.05013	False	0.13575
167998	0.49750	0.51000	13.02507	12.89972	False	-0.31769
167998	0.49750	0.49750	13.02507	13.02507	False	-0.77379
198706	0.31000	0.34000	14.98340	14.64985	False	0.41434
198715	0.34000	0.37000	14.64985	14.32741	False	0.37828
198718	0.47750	0.47500	13.22572	13.25083	False	-0.86778
198718	0.47750	0.43000	13.22572	13.70550	False	0.36470
202604	0.47000	0.45000	13.30108	13.50265	False	-0.64668
202604	0.47000	0.43000	13.30108	13.70550	False	0.09146
202625	0.22000	0.21000	16.08877	16.22568	False	-0.84552
202625	0.22000	0.22000	16.08877	16.08877	False	-0.81381
202660	0.47000	0.48000	13.30108	13.20061	False	-0.41182
225287	0.48000	0.44500	13.20061	13.55322	False	-0.09840
225287	0.46500	0.44500	13.35138	13.55322	False	-0.64503
225313	0.58000	0.62000	12.19243	11.77808	False	0.74482
225313	0.60000	0.62000	11.98661	11.77808	False	-0.00140
225376	0.47000	0.49000	13.30108	13.10028	False	-0.04671
225376	0.47000	0.49000	13.30108	13.10028	False	-0.04671
225376	0.47000	0.49000	13.30108	13.10028	False	-0.04671
225376	0.47000	0.49000	13.30108	13.10028	False	-0.04671
256905	0.56000	0.50000	12.39612	13.00000	False	0.80542
270584	0.24000	0.26000	15.82521	15.57338	False	0.10598
270584	0.24000	0.26000	15.82521	15.57338	False	0.10598

Table J-8. 2012–13 NECAP Rescore Analysis
Mathematics Grade 6

IREF	*Max*	*Mean*		*St Dev*		*Effect Size*	*Discard*
		Old	*New*	*Old*	*New*		
198718	4	2.09756	2.04390	1.33708	1.33444	-0.04013	False
167998	4	2.05911	2.08374	0.98644	0.95671	0.02497	False
145135	4	1.19024	1.16585	0.96224	0.91216	-0.02535	False
125065	4	1.76585	1.75122	1.29762	1.29677	-0.01128	False
198706	4	1.31373	1.34314	1.49259	1.49269	0.01971	False
123448	4	1.44828	1.40887	1.26843	1.29318	-0.03107	False

Figure J-5. 2012–13 NECAP: Delta Plot
Mathematics Grade 7

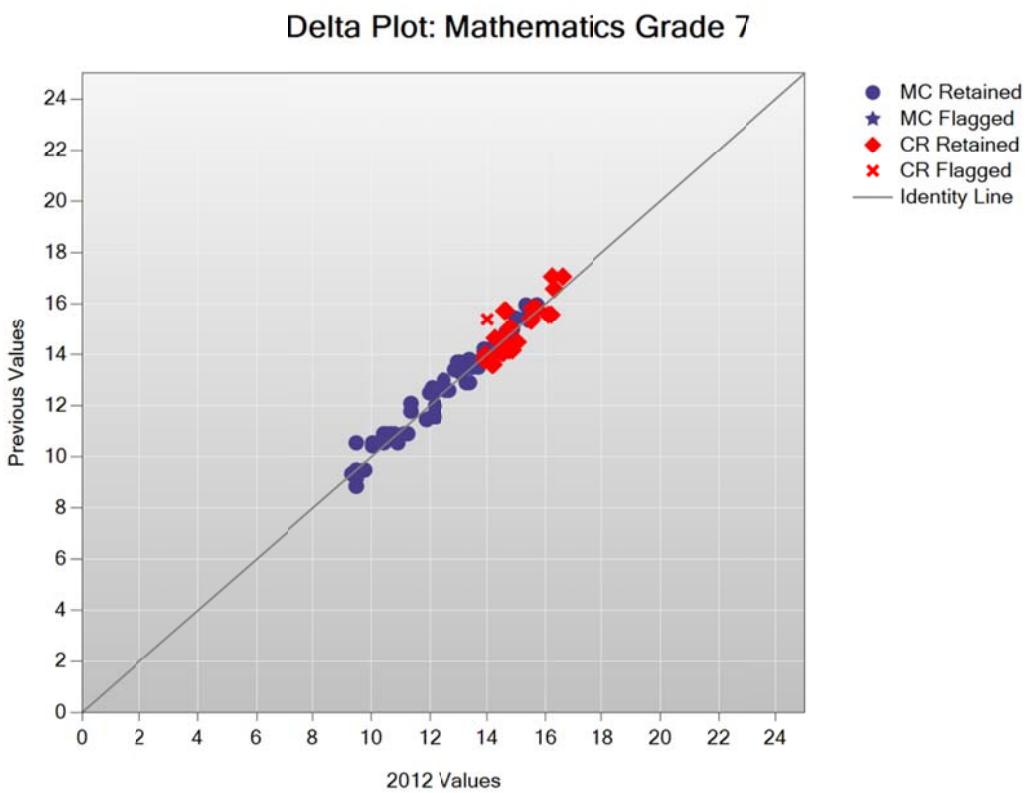


Table J-9. 2012–13 NECAP Delta Analysis
Mathematics Grade 7

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
120326	0.81000	0.79000	9.48841	9.77432	False	-0.35261
120326	0.81000	0.81000	9.48841	9.48841	False	-1.14182
120329	0.64000	0.59000	11.56616	12.08982	False	0.82869
120329	0.64000	0.58000	11.56616	12.19243	False	1.23155
120329	0.64000	0.59000	11.56616	12.08982	False	0.82869
120329	0.64000	0.58000	11.56616	12.19243	False	1.23155

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
120340	0.73000	0.70000	10.54875	10.90240	False	0.03986
120340	0.73000	0.74000	10.54875	10.42662	False	-0.78878
120353	0.60000	0.58000	11.98661	12.19243	False	-0.36907
120353	0.60000	0.58000	11.98661	12.19243	False	-0.36907
120358	0.44000	0.47000	13.60388	13.30108	False	-0.44383
120358	0.44000	0.46000	13.60388	13.40173	False	-0.83903
120363	0.55000	0.59000	12.49735	12.08982	False	0.09936
120363	0.53000	0.59000	12.69892	12.08982	False	0.86672
120363	0.53000	0.59000	12.69892	12.08982	False	0.86672
120363	0.55000	0.59000	12.49735	12.08982	False	0.09936
120364	0.73000	0.77000	10.54875	10.04461	False	0.71106
120364	0.74000	0.77000	10.42662	10.04461	False	0.24612
120379	0.43000	0.45000	13.70550	13.50265	False	-0.84836
120379	0.43000	0.43000	13.70550	13.70550	False	-0.97213
120379	0.45000	0.45000	13.50265	13.50265	False	-0.99632
120379	0.45000	0.43000	13.50265	13.70550	False	-0.19988
120381	0.73000	0.81000	10.54875	9.48841	False	2.89481
120391	0.26000	0.21250	15.57338	16.19111	False	1.67599
120391	0.26000	0.22000	15.57338	16.08877	False	1.27420
120404	0.62000	0.66000	11.77808	11.35015	False	0.26523
120404	0.59000	0.66000	12.08982	11.35015	False	1.45202
120467	0.85000	0.81000	8.85427	9.48841	False	0.93906
120467	0.83000	0.81000	9.18334	9.48841	False	-0.31370
120467	0.83000	0.81000	9.18334	9.48841	False	-0.31370
120467	0.85000	0.81000	8.85427	9.48841	False	0.93906
120470	0.30500	0.32500	15.04029	14.81505	False	-0.91965
120470	0.30500	0.33000	15.04029	14.75965	False	-0.70215
120481	0.43000	0.49000	13.70550	13.10028	False	0.73143
120481	0.43000	0.50000	13.70550	13.00000	False	1.12513
120482	0.44000	0.44000	13.60388	13.60388	False	-0.98425
120485	0.44000	0.44000	13.60388	13.60388	False	-0.98425
120503	0.18500	0.20500	16.58589	16.29557	False	-0.84850
120503	0.18500	0.20500	16.58589	16.29557	False	-0.84850
120517	0.38000	0.35000	14.22192	14.54128	False	0.34334
120517	0.38000	0.36000	14.22192	14.43384	False	-0.07852
120525	0.34500	0.36000	14.59542	14.43384	False	-1.11653
120531	0.70000	0.68000	10.90240	11.12920	False	-0.41597
120531	0.70000	0.67000	10.90240	11.24035	False	0.02040
121799	0.38500	0.32000	14.16950	14.87080	False	1.83665
121799	0.38500	0.33000	14.16950	14.75965	False	1.40028
122346	0.65000	0.61000	11.45872	11.88272	False	0.42463
122346	0.65000	0.61000	11.45872	11.88272	False	0.42463
125382	0.40000	0.40500	14.01339	13.96170	False	-1.13833
139917	0.38000	0.41000	14.22192	13.91018	False	-0.48243
139923	0.39000	0.41000	14.11728	13.91018	False	-0.88081
139924	0.45000	0.45000	13.50265	13.50265	False	-0.99632
139924	0.45000	0.47000	13.50265	13.30108	False	-0.82922
139926	0.31000	0.33000	14.98340	14.75965	False	-0.91874
139926	0.31000	0.32000	14.98340	14.87080	False	-1.26183

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
139971	0.34000	0.34000	14.64985	14.64985	False	-0.85949
139980	0.35500	0.30500	14.48742	15.04029	False	1.29182
139989	0.24000	0.26250	15.82521	15.54263	False	-0.78815
140233	0.70000	0.74000	10.90240	10.42662	False	0.55754
140233	0.70000	0.71000	10.90240	10.78646	False	-0.85528
140251	0.39500	0.34750	14.06524	14.56832	False	1.04595
140251	0.39500	0.35500	14.06524	14.48742	False	0.72836
140261	0.54000	0.54000	12.59827	12.59827	False	-1.10419
140261	0.54000	0.53000	12.59827	12.69892	False	-0.70900
144733	0.32000	0.34000	14.87080	14.64985	False	-0.91632
144735	0.15500	0.21000	17.06089	16.22568	False	1.23418
144735	0.15500	0.18500	17.06089	16.58589	False	-0.18008
145365	0.46000	0.51000	13.40173	12.89972	False	0.36243
145367	0.44000	0.38500	13.60388	14.16950	False	1.23651
145367	0.44000	0.38000	13.60388	14.22192	False	1.44234
145378	0.27000	0.28000	15.45125	15.33137	False	-1.23461
145378	0.27000	0.28000	15.45125	15.33137	False	-1.23461
146208	0.40000	0.39000	14.01339	14.11728	False	-0.52752
146266	0.55000	0.60000	12.49735	11.98661	False	0.50458
169205	0.70000	0.72000	10.90240	10.66863	False	-0.39266
169229	0.42000	0.46000	13.80757	13.40173	False	-0.06357
169401	0.23000	0.28000	15.95539	15.33137	False	0.53689
169401	0.23000	0.25000	15.95539	15.69796	False	-0.90243
169537	0.47000	0.49000	13.30108	13.10028	False	-0.80817
169538	0.38000	0.39000	14.22192	14.11728	False	-1.29553
181112	0.71000	0.74000	10.78646	10.42662	False	0.11618
181112	0.70000	0.74000	10.90240	10.42662	False	0.55754
181112	0.70000	0.73000	10.90240	10.54875	False	0.07804
181112	0.71000	0.73000	10.78646	10.54875	False	-0.36333
199932	0.27500	0.40000	15.39104	14.01339	True	3.56312
204661	0.35000	0.35000	14.54128	14.54128	False	-0.87244
204722	0.42250	0.41000	13.78201	13.91018	False	-0.45977
204874	0.82000	0.82000	9.33854	9.33854	False	-1.12394
204874	0.82000	0.81000	9.33854	9.48841	False	-0.90454
204892	0.27000	0.31000	15.45125	14.98340	False	-0.01614
205005	0.39000	0.39000	14.11728	14.11728	False	-0.92301
206179	0.23000	0.25000	15.95539	15.69796	False	-0.90243
206179	0.23000	0.25000	15.95539	15.69796	False	-0.90243
206190	0.24000	0.25000	15.82521	15.69796	False	-1.21892
206190	0.24000	0.26000	15.82521	15.57338	False	-0.90889
206202	0.50000	0.55000	13.00000	12.49735	False	0.41284
224856	0.34000	0.37500	14.64985	14.27456	False	-0.28395
225091	0.71000	0.73000	10.78646	10.54875	False	-0.36333
225091	0.71000	0.72000	10.78646	10.66863	False	-0.83403
225137	0.28000	0.26500	15.33137	15.51202	False	-0.06890
225137	0.27000	0.26500	15.45125	15.51202	False	-0.52530
227303	0.24750	0.34500	15.72951	14.59542	False	2.56648
227303	0.24750	0.34000	15.72951	14.64985	False	2.35276
234445	0.51000	0.46000	12.89972	13.40173	False	0.90277

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
234445	0.51000	0.47000	12.89972	13.30108	False	0.50757
234445	0.51000	0.46000	12.89972	13.40173	False	0.90277
234445	0.51000	0.47000	12.89972	13.30108	False	0.50757
256152	0.28000	0.27000	15.33137	15.45125	False	-0.30751

Table J-10. 2012–13 NECAP Rescore Analysis
Mathematics Grade 7

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
140251	4	1.60784	1.41176	1.26225	1.27921	-0.15534	False
120391	4	1.02439	0.93659	1.34587	1.28896	-0.06524	False
227303	4	1.01970	1.05911	0.78179	0.75408	0.05041	False
204661	4	1.39216	1.39216	1.04530	1.00215	0.00000	False
204722	4	1.74634	1.75122	1.21611	1.24293	0.00401	False
139989	4	0.93171	0.91220	1.17159	1.16191	-0.01665	False

Figure J-6. 2012–13 NECAP: Delta Plot
Mathematics Grade 8

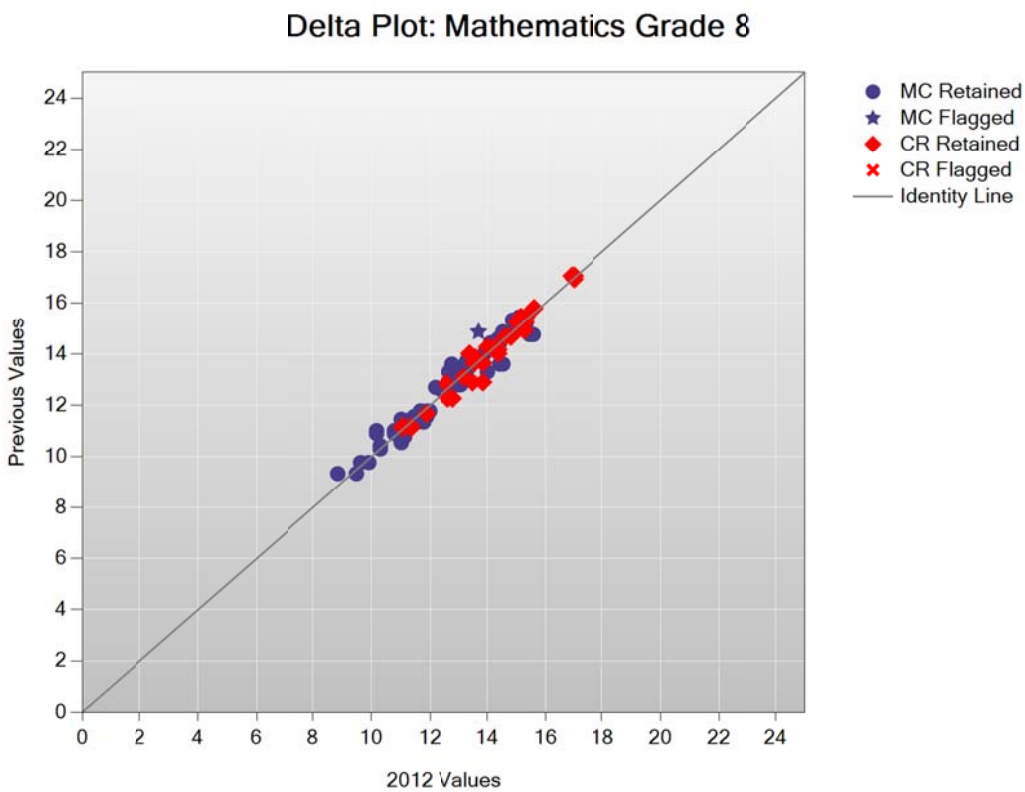


Table J-11. 2012–13 NECAP Delta Analysis
Mathematics Grade 8

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
120875	0.52000	0.49000	12.79939	13.10028	False	0.23033
120875	0.52000	0.50000	12.79939	13.00000	False	-0.18196
120877	0.66000	0.62000	11.35015	11.77808	False	0.74844
120877	0.66000	0.64000	11.35015	11.56616	False	-0.12285
120882	0.82000	0.81000	9.33854	9.48841	False	-0.40067
120882	0.82000	0.85000	9.33854	8.85427	False	0.80099
120884	0.73000	0.69000	10.54875	11.01660	False	0.91024
120886	0.79000	0.80000	9.77432	9.63352	False	-0.61251
120886	0.79000	0.78000	9.77432	9.91123	False	-0.45270
120889	0.69000	0.71000	11.01660	10.78646	False	-0.24881
120889	0.70000	0.76000	10.90240	10.17479	False	1.79692
120889	0.69000	0.76000	11.01660	10.17479	False	2.26613
120889	0.70000	0.71000	10.90240	10.78646	False	-0.71802
120965	0.27000	0.30000	15.45125	15.09760	False	0.24609
120965	0.27000	0.28000	15.45125	15.33137	False	-0.71505
120977	0.31000	0.31000	14.98340	14.98340	False	-1.00043
120985	0.40000	0.46000	14.01339	13.40173	False	1.31109
120993	0.51000	0.45000	12.89972	13.50265	False	1.47245
120993	0.51000	0.41500	12.89972	13.85881	False	2.93684
121027	0.62000	0.61000	11.77808	11.88272	False	-0.57952
121027	0.62000	0.60000	11.77808	11.98661	False	-0.15237
121032	0.57250	0.51750	12.26903	12.82448	False	1.27543
121032	0.57250	0.53500	12.26903	12.64862	False	0.55237
121039	0.29000	0.31000	15.21354	14.98340	False	-0.26105
121052	0.28750	0.30500	15.24281	15.04029	False	-0.37469
121052	0.28750	0.28250	15.24281	15.30172	False	-0.75746
121059	0.36000	0.36000	14.43384	14.43384	False	-1.00204
121059	0.36000	0.39000	14.43384	14.11728	False	0.09655
121073	0.34000	0.35000	14.64985	14.54128	False	-0.75924
121073	0.34000	0.35000	14.64985	14.54128	False	-0.75924
121078	0.66000	0.68000	11.35015	11.12920	False	-0.28759
121078	0.66000	0.66000	11.35015	11.35015	False	-1.01103
121082	0.64000	0.65000	11.56616	11.45872	False	-0.75487
121083	0.47000	0.53000	13.30108	12.69892	False	1.27413
121097	0.24250	0.25750	15.79313	15.60429	False	-0.43252
122667	0.44000	0.47000	13.60388	13.30108	False	0.04239
122667	0.43000	0.47000	13.70550	13.30108	False	0.45991
125576	0.15500	0.16500	17.06089	16.89646	False	-0.53659
125576	0.15500	0.16000	17.06089	16.97783	False	-0.87117
139769	0.52000	0.51000	12.79939	12.89972	False	-0.59425
139774	0.27000	0.29000	15.45125	15.21354	False	-0.23060
139774	0.27000	0.28000	15.45125	15.33137	False	-0.71505
139801	0.48000	0.48000	13.20061	13.20061	False	-1.00563
139815	0.43500	0.44500	13.65463	13.55322	False	-0.78575
139815	0.43500	0.41500	13.65463	13.85881	False	-0.16484
139817	0.16250	0.16000	16.93694	16.97783	False	-0.82661
139832	0.29000	0.30000	15.21354	15.09760	False	-0.73060

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
139832	0.29000	0.28000	15.21354	15.33137	False	-0.51530
139871	0.33750	0.34250	14.67718	14.62260	False	-0.98129
139871	0.33750	0.32500	14.67718	14.81505	False	-0.43448
139877	0.62000	0.63000	11.77808	11.67259	False	-0.76353
145176	0.53000	0.58000	12.69892	12.19243	False	0.88255
145176	0.53000	0.53000	12.69892	12.69892	False	-1.00710
145181	0.33000	0.27000	14.75965	15.45125	False	1.84249
145181	0.33000	0.26000	14.75965	15.57338	False	2.34463
145390	0.45000	0.46000	13.50265	13.40173	False	-0.78739
145409	0.31500	0.28500	14.92691	15.27221	False	0.41913
145645	0.66000	0.68000	11.35015	11.12920	False	-0.28759
145645	0.66000	0.67000	11.35015	11.24035	False	-0.74456
145647	0.40000	0.41000	14.01339	13.91018	False	-0.77943
145651	0.44000	0.35000	13.60388	14.54128	False	2.84977
145651	0.44000	0.36000	13.60388	14.43384	False	2.40799
146307	0.62000	0.63000	11.77808	11.67259	False	-0.76353
146468	0.65000	0.69000	11.45872	11.01660	False	0.62148
146468	0.65000	0.66000	11.45872	11.35015	False	-0.74993
169266	0.64000	0.62000	11.56616	11.77808	False	-0.13910
169266	0.64000	0.61000	11.56616	11.88272	False	0.29116
169370	0.54000	0.56000	12.59827	12.39612	False	-0.36853
169455	0.43000	0.44000	13.70550	13.60388	False	-0.78506
169455	0.43000	0.42000	13.70550	13.80757	False	-0.58446
194560	0.71000	0.68000	10.78646	11.12920	False	0.39655
194560	0.71000	0.69000	10.78646	11.01660	False	-0.06644
199732	0.47000	0.51000	13.30108	12.89972	False	0.44850
199783	0.40000	0.36500	14.01339	14.38050	False	0.50616
199783	0.38500	0.36500	14.16950	14.38050	False	-0.13525
204983	0.35000	0.37000	14.54128	14.32741	False	-0.32598
204990	0.32000	0.35000	14.87080	14.54128	False	0.14854
204990	0.32000	0.43000	14.87080	13.70550	True	3.58495
205780	0.37500	0.40000	14.27456	14.01339	False	-0.13072
205780	0.37500	0.37500	14.27456	14.27456	False	-1.00250
205782	0.49000	0.48000	13.10028	13.20061	False	-0.59337
206240	0.68000	0.68500	11.12920	11.07309	False	-0.96466
206240	0.68000	0.66000	11.12920	11.35015	False	-0.10325
206240	0.67500	0.68500	11.18495	11.07309	False	-0.73562
206240	0.67500	0.66000	11.18495	11.35015	False	-0.33229
206256	0.47000	0.40000	13.30108	14.01339	False	1.92338
206324	0.63000	0.61000	11.67259	11.88272	False	-0.14609
206335	0.51500	0.54000	12.84957	12.59827	False	-0.16714
224873	0.35000	0.36000	14.54128	14.43384	False	-0.76354
224873	0.34000	0.36000	14.64985	14.43384	False	-0.31746
224919	0.64000	0.63000	11.56616	11.67259	False	-0.57284
224919	0.64000	0.62000	11.56616	11.77808	False	-0.13910
224929	0.44000	0.52000	13.60388	12.79939	False	2.10515
224947	0.42000	0.44000	13.80757	13.60388	False	-0.36566
224947	0.41500	0.44000	13.85881	13.60388	False	-0.15517
225437	0.28000	0.32000	15.33137	14.87080	False	0.68605

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
225437	0.28000	0.29000	15.33137	15.21354	False	-0.72317
256490	0.27000	0.29500	15.45125	15.15534	False	0.00868
256490	0.27000	0.28000	15.45125	15.33137	False	-0.71505
269362	0.75000	0.75000	10.30204	10.30204	False	-1.01409
269362	0.74000	0.75000	10.42662	10.30204	False	-0.68111

Table J-12. 2012–13 NECAP Rescore Analysis
Mathematics Grade 8

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
121052	4	1.20098	1.16176	1.30456	1.30250	-0.03006	False
121032	4	2.39512	2.36585	1.34971	1.31408	-0.02168	False
139871	4	1.32683	1.41463	1.25230	1.24968	0.07011	False
139817	4	0.68627	0.70588	0.92391	0.94604	0.02122	False
121097	4	0.98049	0.92195	1.00286	0.92384	-0.05837	False
206335	4	2.07805	2.09756	1.51630	1.52800	0.01287	False

Figure J-7. 2012–13 NECAP: Delta Plot
Mathematics Grade 11

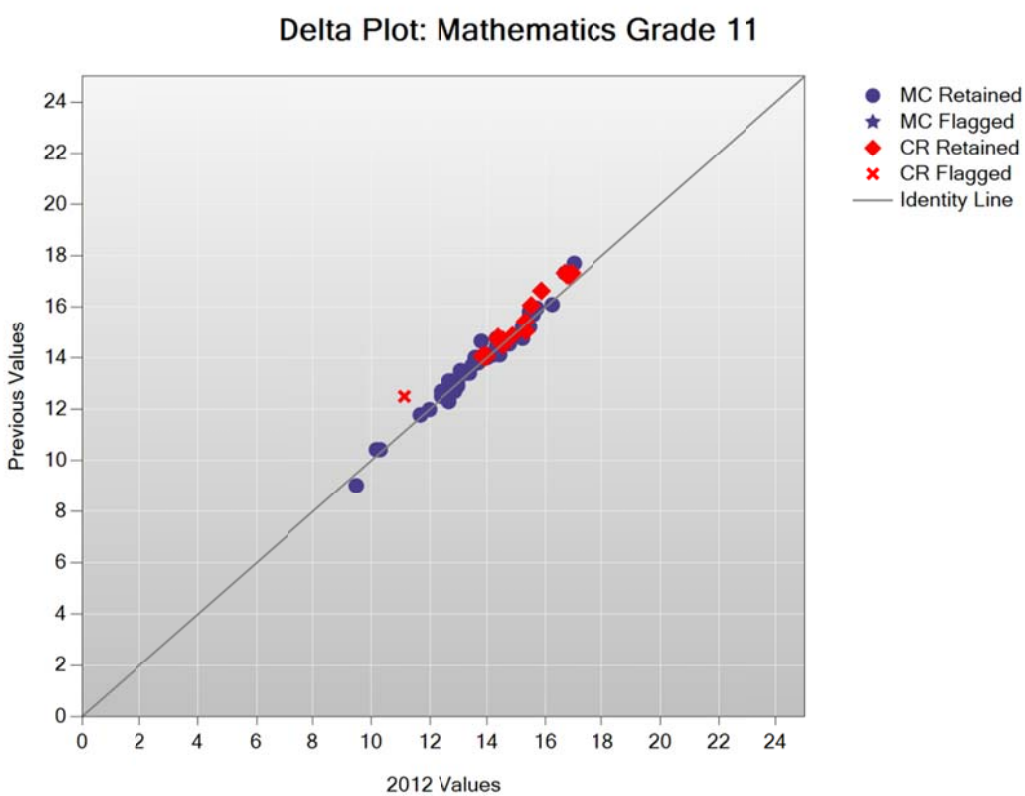


Table J-13. 2012–13 NECAP Delta Analysis
Mathematics Grade 11

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
119424	0.43000	0.45000	13.70550	13.50265	False	-0.67692
119434	0.51000	0.52000	12.89972	12.79939	False	-0.98098
119434	0.51000	0.52000	12.89972	12.79939	False	-0.98098
119434	0.51000	0.52000	12.89972	12.79939	False	-0.98098
119434	0.51000	0.52000	12.89972	12.79939	False	-0.98098
119472	0.33500	0.36500	14.70459	14.38050	False	-0.32994
119472	0.32500	0.36500	14.81505	14.38050	False	0.20488
119484	0.53000	0.56000	12.69892	12.39612	False	0.10956
119484	0.55000	0.56000	12.49735	12.39612	False	-0.86639
119502	0.32000	0.32000	14.87080	14.87080	False	-0.04860
119520	0.57000	0.53000	12.29450	12.69892	False	1.31566
119542	0.34000	0.42000	14.64985	13.80757	False	2.33570
119544	0.55000	0.68000	12.49735	11.12920	True	5.61424
119552	0.29000	0.29000	15.21354	15.21354	False	0.04512
119552	0.29000	0.27000	15.21354	15.45125	False	1.26108
119563	0.12000	0.16000	17.69995	16.97783	False	0.88704
119566	0.14500	0.17500	17.23249	16.73836	False	-0.15135
119566	0.14500	0.17000	17.23249	16.81666	False	-0.55189
119592	0.39000	0.38000	14.11728	14.22192	False	0.28066
119630	0.62000	0.63000	11.77808	11.67259	False	-0.64793
119630	0.62000	0.63000	11.77808	11.67259	False	-0.64793
141248	0.55000	0.55000	12.49735	12.49735	False	-0.69758
145213	0.14000	0.18000	17.32128	16.66146	False	0.67191
145213	0.14000	0.16500	17.32128	16.89646	False	-0.53015
145455	0.25000	0.27000	15.69796	15.45125	False	-0.99739
145455	0.25000	0.26000	15.69796	15.57338	False	-0.45967
145455	0.24000	0.27000	15.82521	15.45125	False	-0.38126
145455	0.24000	0.26000	15.82521	15.57338	False	-1.00599
145459	0.53000	0.52000	12.69892	12.79939	False	-0.12855
145459	0.53000	0.53000	12.69892	12.69892	False	-0.64246
169505	0.46000	0.47000	13.40173	13.30108	False	-0.96517
169505	0.46000	0.46000	13.40173	13.40173	False	-0.45029
169512	0.36000	0.37000	14.43384	14.32741	False	-0.71245
169512	0.33000	0.37000	14.75965	14.32741	False	0.20822
169580	0.74000	0.76000	10.42662	10.17479	False	0.47016
169580	0.74000	0.75000	10.42662	10.30204	False	-0.18076
169584	0.14000	0.18000	17.32128	16.66146	False	0.67191
169590	0.23000	0.25000	15.95539	15.69796	False	-1.01294
169599	0.33500	0.35000	14.70459	14.54128	False	-0.92942
169657	0.30500	0.28000	15.04029	15.33137	False	1.48666
169703	0.42000	0.43000	13.80757	13.70550	False	-0.86147
169704	0.84000	0.81000	9.02217	9.48841	False	0.73717
178080	0.22000	0.21000	16.08877	16.22568	False	0.98478
198612	0.39750	0.40250	14.03929	13.98753	False	-0.54077
198612	0.39250	0.40250	14.09124	13.98753	False	-0.79226
198612	0.39250	0.41500	14.09124	13.85881	False	-0.63109
198612	0.39750	0.41500	14.03929	13.85881	False	-0.88259

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
205859	0.27250	0.28250	15.42108	15.30172	False	-0.50866
206071	0.45000	0.49000	13.50265	13.10028	False	0.39914
206081	0.60000	0.60000	11.98661	11.98661	False	-0.83723
206092	0.35000	0.33000	14.54128	14.75965	False	0.97833
206106	0.18250	0.23750	16.62352	15.85747	False	1.40609
206318	0.49000	0.53000	13.10028	12.69892	False	0.50397
206318	0.49000	0.52000	13.10028	12.79939	False	-0.00994
206346	0.53000	0.53000	12.69892	12.69892	False	-0.64246
206346	0.53000	0.51000	12.69892	12.89972	False	0.38470
206381	0.35500	0.35000	14.48742	14.54128	False	0.12207
259784	0.39000	0.39000	14.11728	14.11728	False	-0.25464
259784	0.39000	0.40000	14.11728	14.01339	False	-0.78605
259828	0.51000	0.50000	12.89972	13.00000	False	-0.07462
259828	0.50000	0.50000	13.00000	13.00000	False	-0.56014
259849	0.22250	0.26500	16.05511	15.51202	False	0.42100
259849	0.22250	0.26500	16.05511	15.51202	False	0.42100
259873	0.39000	0.36000	14.11728	14.43384	False	1.36465
259873	0.39000	0.38000	14.11728	14.22192	False	0.28066
259914	0.40000	0.40000	14.01339	14.01339	False	-0.28304
259914	0.40000	0.40000	14.01339	14.01339	False	-0.28304
259944	0.35000	0.34000	14.54128	14.64985	False	0.41667
270853	0.40000	0.41000	14.01339	13.91018	False	-0.81098
270854	0.40000	0.44000	14.01339	13.60388	False	0.29601
270854	0.40000	0.44000	14.01339	13.60388	False	0.29601
270854	0.40000	0.43000	14.01339	13.70550	False	-0.22380
270854	0.40000	0.43000	14.01339	13.70550	False	-0.22380
270856	0.33000	0.29000	14.75965	15.21354	False	2.24276
270864	0.36000	0.36000	14.43384	14.43384	False	-0.16808
270864	0.36000	0.36000	14.43384	14.43384	False	-0.16808
270864	0.36000	0.36000	14.43384	14.43384	False	-0.16808
270864	0.36000	0.36000	14.43384	14.43384	False	-0.16808

Table J-14. 2012–13 NECAP Rescore Analysis
Mathematics Grade 11

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
259849	4	1.08780	1.07317	1.25876	1.27702	-0.01163	False
198612	4	1.64904	1.65385	1.08680	1.08613	0.00442	False
205859	4	1.16505	1.18932	1.50235	1.50110	0.01616	False
206106	4	0.81250	0.81731	1.36606	1.37549	0.00352	False
206381	4	1.49756	1.50732	1.25648	1.28338	0.00776	False
169657	4	1.25366	1.30244	1.07089	1.13389	0.04555	False

Figure J-8. 2012–13 NECAP: Delta Plot
Reading Grade 3

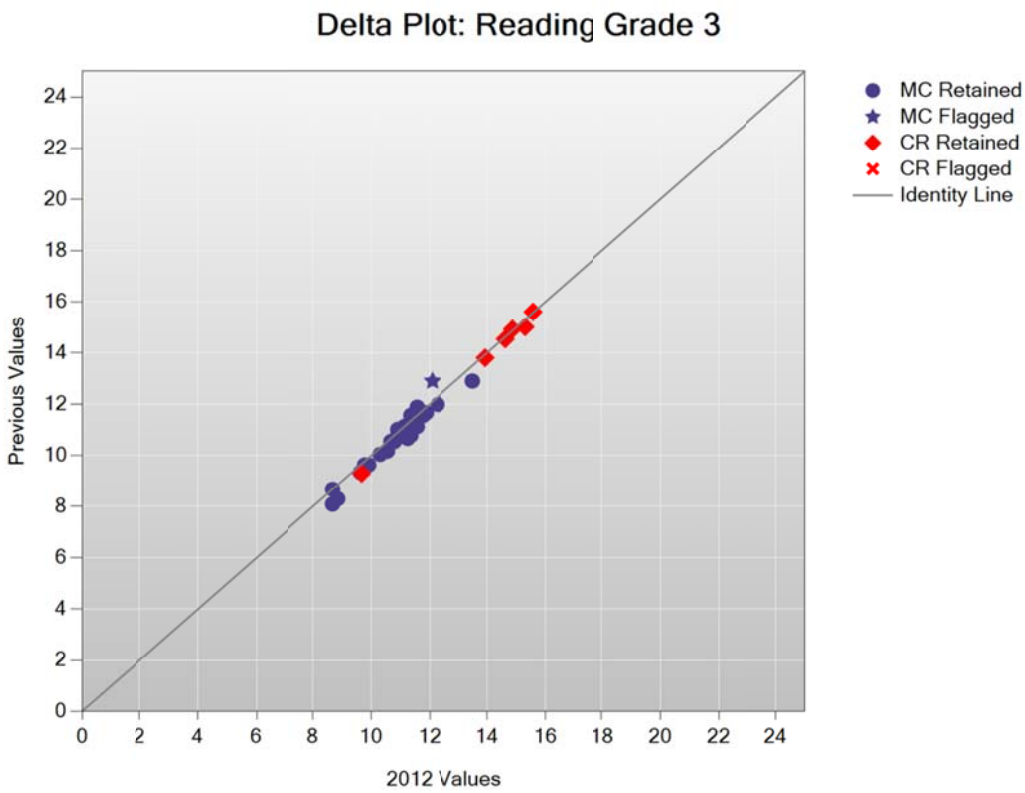


Table J-15. 2012–13 NECAP Delta Analysis
Reading Grade 3

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
117807	0.51000	0.59000	12.89972	12.08982	True	3.82770
171437	0.66000	0.66000	11.35015	11.35015	False	-0.14906
171448	0.80000	0.78000	9.63352	9.91123	False	-0.87993
171450	0.82000	0.80000	9.33854	9.63352	False	-0.86024
171452	0.64000	0.62000	11.56616	11.77808	False	-0.75421
171458	0.77000	0.75000	10.04461	10.30204	False	-0.88705
171461	0.73000	0.71000	10.54875	10.78646	False	-0.86800
171470	0.71000	0.68000	10.78646	11.12920	False	-0.24331
171472	0.60000	0.57000	11.98661	12.29450	False	-0.13288
171501	0.30750	0.28250	15.01180	15.30172	False	0.52184
171502	0.35000	0.34250	14.54128	14.62260	False	-0.71849
207603	0.64000	0.66000	11.56616	11.35015	False	0.96063
207605	0.80000	0.79000	9.63352	9.77432	False	-0.48091
207615	0.69000	0.70000	11.01660	10.90240	False	0.54881
207620	0.51000	0.45000	12.89972	13.50265	False	1.68281
207623	0.42000	0.40750	13.80757	13.93592	False	-0.64746
207737	0.88000	0.85000	8.30005	8.85427	False	0.27791
207740	0.64000	0.65000	11.56616	11.45872	False	0.37592

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
207758	0.68000	0.68000	11.12920	11.12920	False	-0.09418
207788	0.73000	0.72000	10.54875	10.66863	False	-0.59563
207795	0.31500	0.32000	14.92691	14.87080	False	-0.73536
207968	0.89000	0.86000	8.09389	8.67872	False	0.39161
207984	0.71000	0.70000	10.78646	10.90240	False	-0.63341
208009	0.72000	0.67000	10.66863	11.24035	False	0.96052
208110	0.64000	0.63000	11.56616	11.67259	False	-0.77585
208126	0.71000	0.69000	10.78646	11.01660	False	-0.84974
208132	0.63000	0.61000	11.67259	11.88272	False	-0.73734
208140	0.86000	0.86000	8.67872	8.67872	False	0.51453
208145	0.61000	0.64000	11.88272	11.56616	False	1.42345
208148	0.68000	0.64000	11.12920	11.56616	False	0.34923
208151	0.76000	0.73000	10.17479	10.54875	False	-0.22715
208163	0.71000	0.66000	10.78646	11.35015	False	0.94656
208184	0.82250	0.79750	9.30026	9.66910	False	-0.47196
208196	0.25750	0.26000	15.60429	15.57338	False	-1.03937

Table J-16. 2012–13 NECAP Rescore Analysis
Reading Grade 3

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
207795	4	1.27106	1.09646	1.01807	0.97861	-0.17150	False
171501	4	1.09756	0.92683	0.91667	0.83250	-0.18625	False
171502	4	1.43415	1.42439	0.80402	0.90603	-0.01213	False
207623	4	1.64878	1.57073	1.11097	1.05577	-0.07025	False
208184	4	3.27805	3.27317	1.00098	1.00961	-0.00487	False
208196	4	1.00493	0.97537	0.84566	0.99040	-0.03495	False

Figure J-9. 2012–13 NECAP: Delta Plot
Reading Grade 4

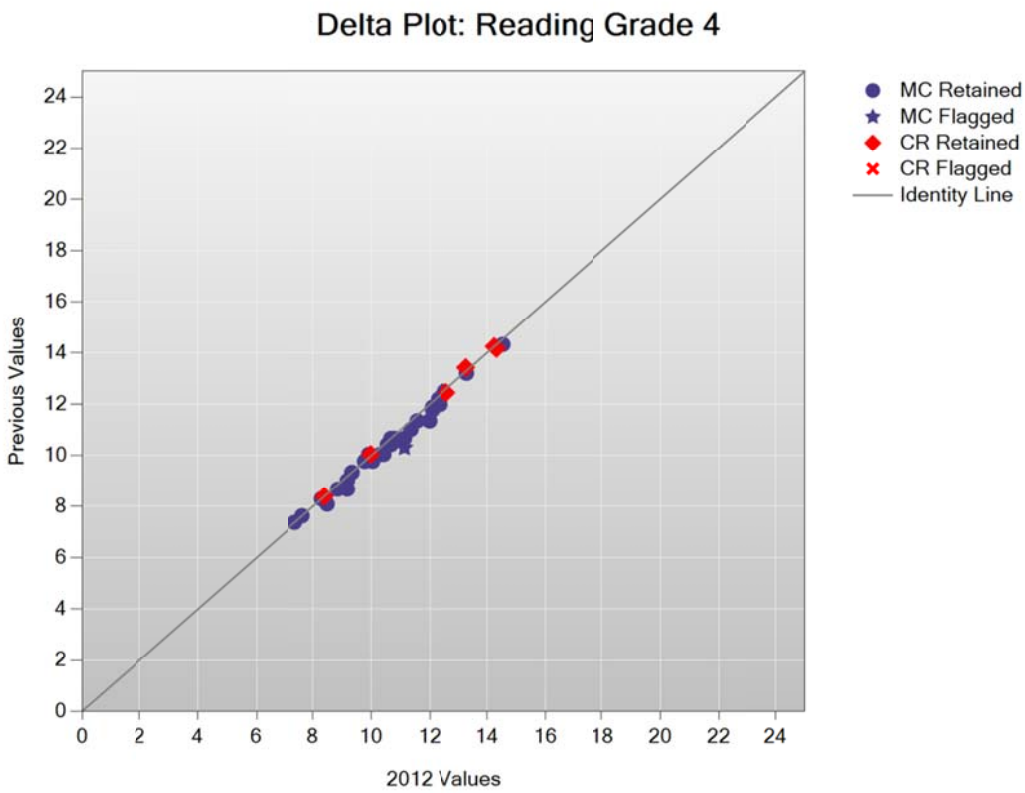


Table J-17. 2012–13 NECAP Delta Analysis
Reading Grade 4

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
171486	0.60000	0.56000	11.98661	12.39612	False	0.29574
171492	0.92000	0.92000	7.37971	7.37971	False	-0.16113
171505	0.77000	0.78000	10.04461	9.91123	False	1.00161
171510	0.89000	0.87000	8.09389	8.49444	False	0.50773
171519	0.37750	0.37750	14.24821	14.24821	False	0.32827
172189	0.86000	0.85000	8.67872	8.85427	False	-1.17503
208211	0.74000	0.73000	10.42662	10.54875	False	-0.83479
208214	0.37000	0.35000	14.32741	14.54128	False	-1.22595
208227	0.79000	0.79000	9.77432	9.77432	False	0.00949
208230	0.62000	0.59000	11.77808	12.08982	False	-0.40248
208230	0.61000	0.59000	11.88272	12.08982	False	-1.17319
208232	0.82000	0.82000	9.33854	9.33854	False	-0.02156
208238	0.79000	0.77000	9.77432	10.04461	False	-0.56200
208250	0.84000	0.83000	9.02217	9.18334	False	-1.21961
208250	0.86000	0.83000	8.67872	9.18334	False	1.22509
208263	0.72000	0.71000	10.66863	10.78646	False	-0.78617
208272	0.77000	0.77500	10.04461	9.97834	False	0.51212
208276	0.45750	0.47250	13.42694	13.27595	False	1.37100

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
208431	0.88000	0.88000	8.30005	8.30005	False	-0.09555
208438	0.58000	0.57000	12.19243	12.29450	False	-0.56272
208473	0.48000	0.47000	13.20061	13.30108	False	-0.47912
208483	0.77000	0.74000	10.04461	10.42662	False	0.23350
208498	0.38500	0.37000	14.16950	14.32741	False	-0.82910
208748	0.55000	0.55000	12.49735	12.49735	False	0.20352
208756	0.72000	0.72000	10.66863	10.66863	False	0.07322
208756	0.74000	0.72000	10.42662	10.66863	False	-0.81475
208760	0.77000	0.75000	10.04461	10.30204	False	-0.67512
208763	0.69000	0.66000	11.01660	11.35015	False	-0.18919
208776	0.57000	0.56000	12.29450	12.39612	False	-0.55211
208788	0.66000	0.60000	11.35015	11.98661	False	1.99639
208788	0.66000	0.60000	11.35015	11.98661	False	1.99639
208800	0.66000	0.64000	11.35015	11.56616	False	-1.07018
208802	0.72000	0.68000	10.66863	11.12920	False	0.76205
208802	0.75000	0.68000	10.30204	11.12920	True	3.46195
208807	0.87500	0.87500	8.39860	8.39860	False	-0.08853
208827	0.55500	0.54000	12.44678	12.59827	False	-0.90493
208863	0.79000	0.79000	9.77432	9.77432	False	0.00949
208864	0.82000	0.82000	9.33854	9.33854	False	-0.02156
208867	0.91000	0.91000	7.63698	7.63698	False	-0.14280

Table J-18. 2012–13 NECAP Rescore Analysis
Reading Grade 4

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
171519	4	1.47253	1.55067	0.97658	0.86965	0.08002	False
208807	4	3.47317	3.46341	0.87595	0.87560	-0.01114	False
208827	4	2.39512	2.35610	1.24427	1.10279	-0.03136	False
208498	4	1.57561	1.63902	0.88421	0.89848	0.07172	False
208272	4	3.08780	3.09756	1.01375	1.00317	0.00962	False
208276	4	1.90244	1.97073	0.92727	0.83812	0.07365	False

Figure J-10. 2012–13 NECAP: Delta Plot
Reading Grade 5

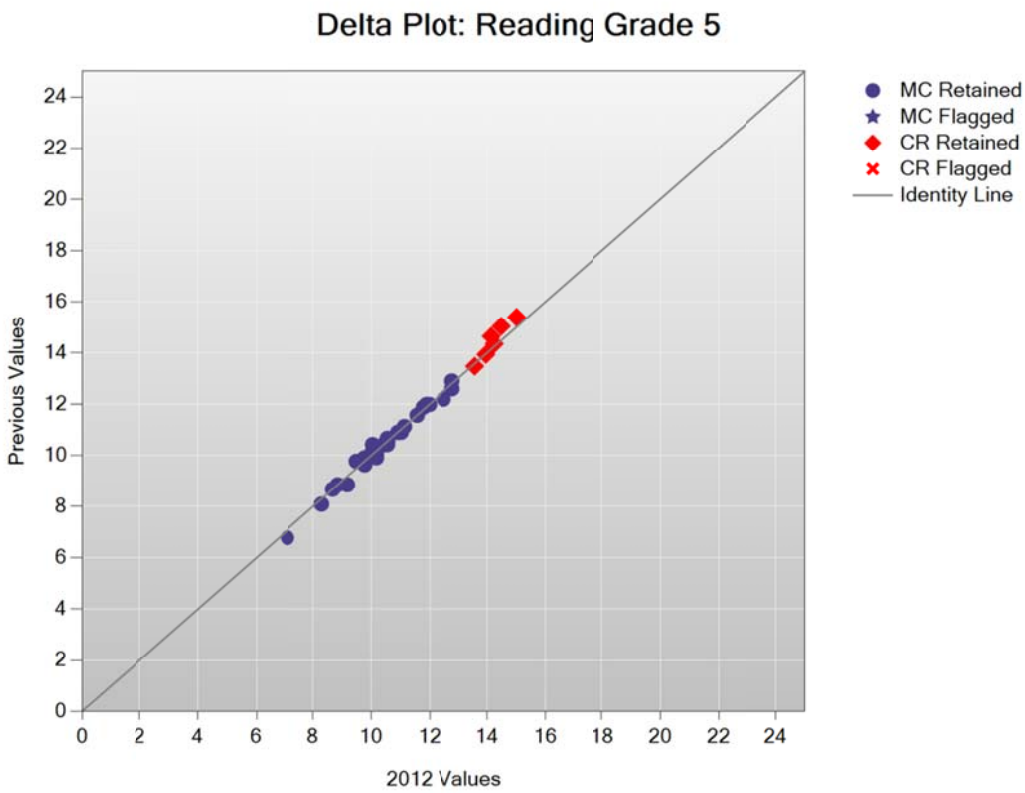


Table J-19. 2012–13 NECAP Delta Analysis
Reading Grade 5

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
148390	0.74000	0.73000	10.42662	10.54875	False	-0.26898
149070	0.94000	0.93000	6.78091	7.09684	False	-0.28844
149074	0.77000	0.76000	10.04461	10.17479	False	-0.37633
149079	0.74000	0.77000	10.42662	10.04461	False	2.29481
149082	0.45250	0.44250	13.47739	13.57854	False	0.96016
171306	0.79000	0.81000	9.77432	9.48841	False	1.77059
171308	0.58000	0.55000	12.19243	12.49735	False	2.11707
171313	0.74000	0.74000	10.42662	10.42662	False	-0.98671
171317	0.78000	0.79000	9.91123	9.77432	False	0.42750
171320	0.40750	0.40500	13.93592	13.96170	False	0.52460
171482	0.85000	0.85000	8.85427	8.85427	False	-0.26031
171493	0.76000	0.77000	10.17479	10.04461	False	0.24788
171564	0.70000	0.70000	10.90240	10.90240	False	-1.09830
171637	0.74000	0.74000	10.42662	10.42662	False	-0.98671
171639	0.80000	0.79000	9.63352	9.77432	False	-0.47500
171655	0.51000	0.52000	12.89972	12.79939	False	-1.03750
171656	0.86000	0.86000	8.67872	8.67872	False	-0.17921
171666	0.73000	0.73000	10.54875	10.54875	False	-1.04313

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
171668	0.36750	0.37750	14.35393	14.24821	False	-0.41187
171670	0.27500	0.30750	15.39104	15.01180	False	-0.02239
172065	0.78000	0.76000	9.91123	10.17479	False	0.70786
172080	0.61000	0.62000	11.88272	11.77808	False	-0.76046
209017	0.85000	0.83000	8.85427	9.18334	False	0.78231
209025	0.72000	0.73000	10.66863	10.54875	False	-0.06867
209025	0.73000	0.73000	10.54875	10.54875	False	-1.04313
209026	0.70000	0.69000	10.90240	11.01660	False	-0.11729
209039	0.60000	0.61000	11.98661	11.88272	False	-0.81498
209043	0.54000	0.52000	12.59827	12.79939	False	1.41283
209047	0.64000	0.64000	11.56616	11.56616	False	-0.79165
209050	0.60000	0.60000	11.98661	11.98661	False	-0.59741
209051	0.73000	0.73000	10.54875	10.54875	False	-1.04313
209057	0.30500	0.35500	15.04029	14.48742	False	1.63113
209060	0.34000	0.38750	14.64985	14.14336	False	1.41309
209702	0.68000	0.68000	11.12920	11.12920	False	-0.99352
209715	0.89000	0.88000	8.09389	8.30005	False	-0.62477

Table J-20. 2012–13 NECAP Rescore Analysis
Reading Grade 5

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
171320	4	1.62414	1.66340	1.05685	1.04211	0.03714	False
209057	4	1.27941	1.37255	0.95809	1.05233	0.09721	False
209060	4	1.37073	1.30732	0.75203	0.74490	-0.08432	False
149082	4	1.83902	1.83415	0.94189	0.94879	-0.00518	False
171668	4	1.51471	1.45588	0.89420	0.95174	-0.06578	False
171670	4	1.05854	1.30244	0.99141	0.97139	0.24602	False

Figure J-11. 2012–13 NECAP: Delta Plot
Reading Grade 6

Table J-21. 2012–13 NECAP Delta Analysis
Reading Grade 6

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
171785	0.68000	0.68000	11.12920	11.12920	False	-0.70464
171801	0.89000	0.88000	8.09389	8.30005	False	0.19585
171807	0.44000	0.41000	13.60388	13.91018	False	-0.80812
171818	0.65000	0.60000	11.45872	11.98661	False	1.14090
171823	0.42250	0.35750	13.78201	14.46060	False	1.36620
171853	0.71000	0.70000	10.78646	10.90240	False	-1.12691
171857	0.49000	0.49000	13.10028	13.10028	False	-0.13128
171867	0.87000	0.87000	8.49444	8.49444	False	-1.15346
171868	0.94000	0.94000	6.78091	6.78091	False	-0.65502
171872	0.76000	0.76000	10.17479	10.17479	False	-0.98226
171873	0.95000	0.95000	6.42059	6.42059	False	-0.55021
171882	0.77000	0.76000	10.04461	10.17479	False	-0.82597
171885	0.92000	0.92000	7.37971	7.37971	False	-0.82921
171889	0.45000	0.46000	13.50265	13.40173	False	0.58917
171893	0.49750	0.47250	13.02507	13.27595	False	-0.97116
172035	0.52000	0.52000	12.79939	12.79939	False	-0.21881
172042	0.63000	0.64000	11.67259	11.56616	False	0.08979
172045	0.89000	0.91000	8.09389	7.63698	False	1.14459

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
172056	0.52000	0.51000	12.79939	12.89972	False	-0.81880
172069	0.69000	0.71000	11.01660	10.78646	False	0.63875
172074	0.83000	0.84000	9.18334	9.02217	False	-0.30691
172081	0.87000	0.89000	8.49444	8.09389	False	0.92409
172094	0.68000	0.62000	11.12920	11.77808	False	1.96016
172099	0.39750	0.38750	14.03929	14.14336	False	-0.48044
172103	0.40750	0.42250	13.93592	13.78201	False	1.03215
172206	0.51000	0.50000	12.89972	13.00000	False	-0.78923
172232	0.80000	0.83000	9.63352	9.18334	False	1.55219
208949	0.69000	0.68000	11.01660	11.12920	False	-1.21377
208957	0.70000	0.66000	10.90240	11.35015	False	0.82349
209416	0.76000	0.74000	10.17479	10.42662	False	-0.13640
209425	0.88000	0.87000	8.30005	8.49444	False	0.06542
209428	0.87000	0.87000	8.49444	8.49444	False	-1.15346
209432	0.57000	0.49000	12.29450	13.10028	False	2.55940
209437	0.52250	0.52250	12.77428	12.77428	False	-0.22611

Table J-22. 2012–13 NECAP Rescore Analysis
Reading Grade 6

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
171823	4	1.65446	1.40049	0.93898	0.83669	-0.27047	False
209437	4	2.07389	2.09852	0.78170	0.88842	0.03151	False
171893	4	2.15122	1.94146	1.00381	0.93055	-0.20896	False
171889	4	1.77561	1.91220	0.90495	0.81025	0.15093	False
172103	4	1.69268	1.71220	0.93667	0.84449	0.02083	False
172099	4	1.51220	1.48293	0.80663	0.81856	-0.03628	False

Figure J-12. 2012–13 NECAP: Delta Plot
Reading Grade 7

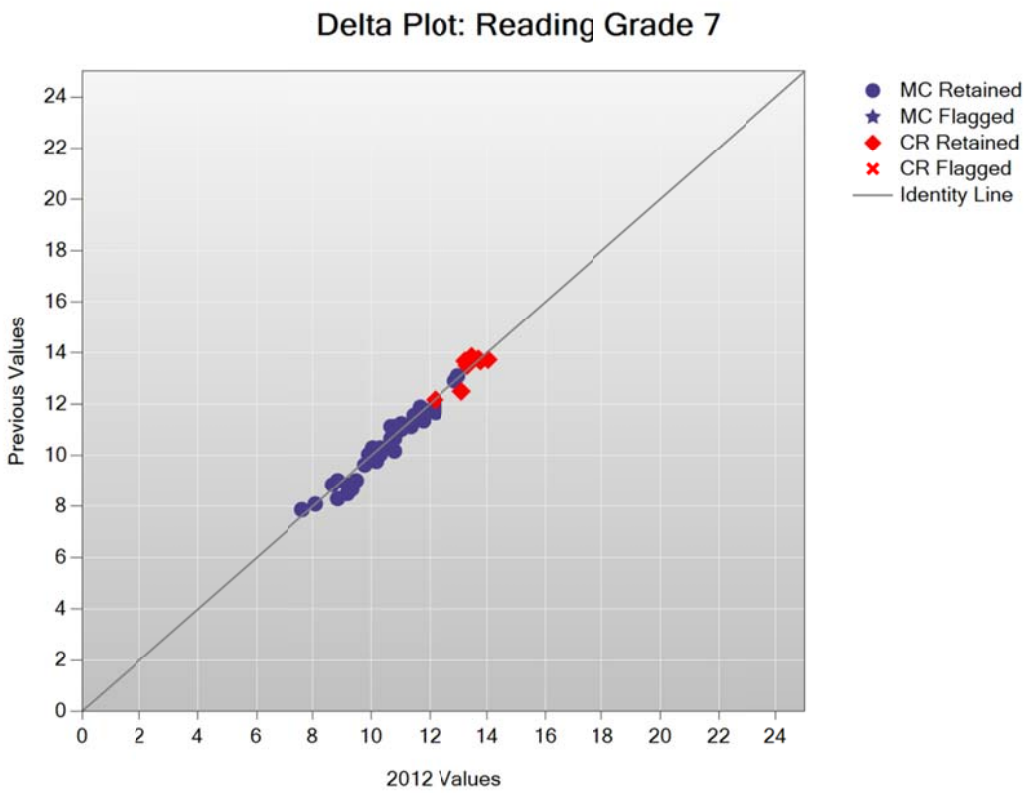


Table J-23. 2012–13 NECAP Delta Analysis
Reading Grade 7

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
147125	0.80000	0.79000	9.63352	9.77432	False	-1.58404
147130	0.64000	0.65000	11.56616	11.45872	False	-0.41424
147132	0.60000	0.58000	11.98661	12.19243	False	-0.62495
147135	0.75000	0.75000	10.30204	10.30204	False	-0.82078
147136	0.68000	0.72000	11.12920	10.66863	False	1.99317
147140	0.67000	0.69000	11.24035	11.01660	False	0.42108
147141	0.68000	0.69000	11.12920	11.01660	False	-0.27860
147144	0.61000	0.63000	11.88272	11.67259	False	0.18233
147149	0.55000	0.48750	12.49735	13.12535	False	2.25055
147153	0.42750	0.39750	13.73097	14.03929	False	0.45132
147526	0.90000	0.91000	7.87379	7.63698	False	1.29190
147539	0.90000	0.91000	7.87379	7.63698	False	1.29190
147546	0.62000	0.62000	11.77808	11.77808	False	-1.16518
147549	0.89000	0.89000	8.09389	8.09389	False	-0.30555
147555	0.58250	0.58000	12.16683	12.19243	False	-1.42302
147677	0.76000	0.76000	10.17479	10.17479	False	-0.79109
147690	0.75000	0.77000	10.30204	10.04461	False	0.85991
173294	0.69000	0.69000	11.01660	11.01660	False	-0.98751

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
173304	0.76000	0.71000	10.17479	10.78646	False	1.60203
173309	0.51000	0.51000	12.89972	12.89972	False	-1.42690
173310	0.77000	0.78000	10.04461	9.91123	False	0.11013
173312	0.72000	0.72000	10.66863	10.66863	False	-0.90631
173315	0.62000	0.58000	11.77808	12.19243	False	0.68786
173315	0.61000	0.58000	11.88272	12.19243	False	0.02907
173316	0.66000	0.62000	11.35015	11.77808	False	0.67668
173316	0.64000	0.62000	11.56616	11.77808	False	-0.68324
173319	0.68000	0.66000	11.12920	11.35015	False	-0.72624
173320	0.43250	0.47500	13.68005	13.25083	False	1.19333
173323	0.45000	0.46750	13.50265	13.32622	False	-0.41576
207506	0.85000	0.83000	8.85427	9.18334	False	-0.55110
207530	0.72000	0.71000	10.66863	10.78646	False	-1.50692
207539	0.84000	0.85000	9.02217	8.85427	False	0.57404
207558	0.49000	0.50000	13.10028	13.00000	False	-0.81902
207565	0.87000	0.83000	8.49444	9.18334	False	1.71419
207566	0.86000	0.82000	8.67872	9.33854	False	1.56729
207566	0.85000	0.82000	8.85427	9.33854	False	0.46217
207583	0.71000	0.71000	10.78646	10.78646	False	-0.93381
207583	0.72000	0.71000	10.66863	10.78646	False	-1.50692
207585	0.75000	0.73000	10.30204	10.54875	False	-0.75103
207588	0.63000	0.58000	11.67259	12.19243	False	1.35197
207591	0.68000	0.69000	11.12920	11.01660	False	-0.27860
207598	0.85000	0.82000	8.85427	9.33854	False	0.46217
207604	0.68000	0.71000	11.12920	10.78646	False	1.22390
207604	0.70000	0.71000	10.90240	10.78646	False	-0.20394
207608	0.43250	0.42250	13.68005	13.78201	False	-0.90788
207622	0.44000	0.46750	13.60388	13.32622	False	0.22153
207669	0.62000	0.59000	11.77808	12.08982	False	0.01798
207671	0.84000	0.81000	9.02217	9.48841	False	0.38365
207675	0.88000	0.85000	8.30005	8.85427	False	0.78947
207678	0.85000	0.86000	8.85427	8.67872	False	0.66311
207702	0.41500	0.45250	13.85881	13.47739	False	0.83949
207793	0.75000	0.72000	10.30204	10.66863	False	0.03167
207817	0.77000	0.75000	10.04461	10.30204	False	-0.74110
207819	0.79000	0.76000	9.77432	10.17479	False	0.12974
207820	0.67000	0.66000	11.24035	11.35015	False	-1.42593
207827	0.42500	0.43000	13.75647	13.70550	False	-1.29398

Table J-24. 2012–13 NECAP Rescore Analysis
Reading Grade 7

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
207702	4	1.72650	1.80952	0.78083	0.89435	0.10633	False
207608	4	1.71220	1.87805	0.98343	0.99314	0.16865	False
207622	4	1.77451	1.97549	1.20455	1.27818	0.16685	False
207827	4	1.72683	1.79512	1.05224	1.05847	0.06490	False
173320	4	1.72549	1.86765	0.87115	1.00898	0.16318	False

continued

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
173323	4	1.77561	1.95610	0.97760	1.00452	0.18462	False
147555	4	2.28431	2.26961	0.97455	0.96101	-0.01509	False
147149	4	2.18627	1.99510	0.96801	1.02244	-0.19749	False
147153	4	1.71845	1.59223	0.98511	0.97039	-0.12812	False

Figure J-13. 2012–13 NECAP: Delta Plot
Reading Grade 8

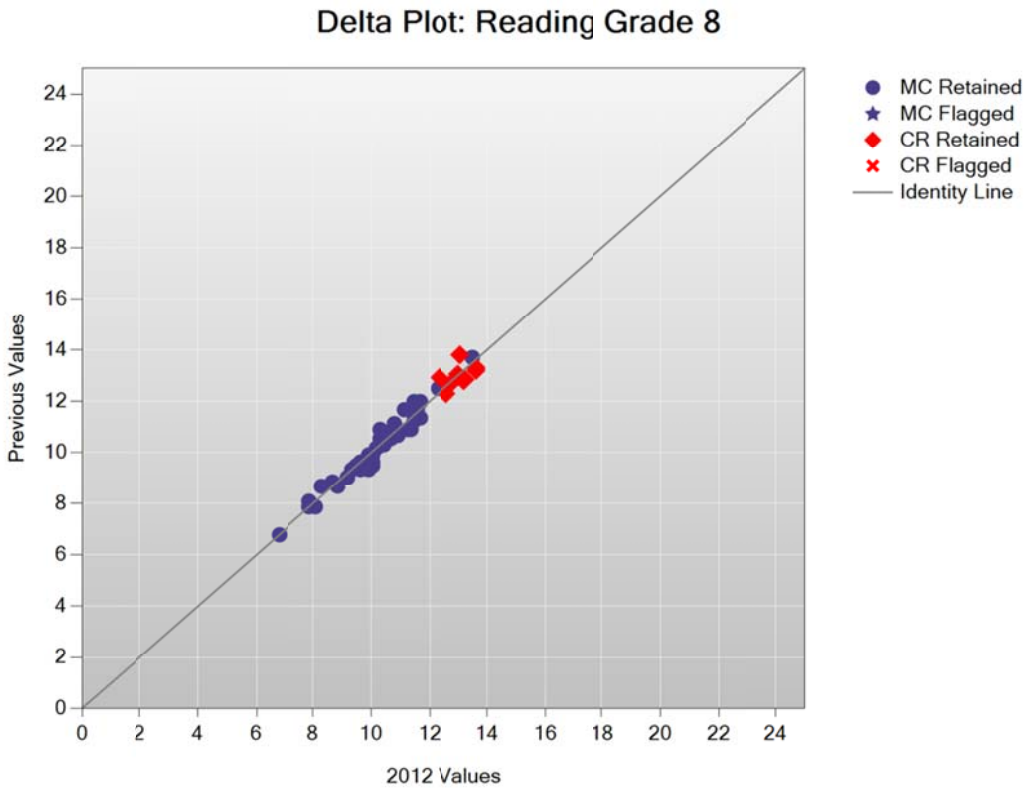


Table J-25. 2012–13 NECAP Delta Analysis
Reading Grade 8

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
147516	0.70000	0.66000	10.90240	11.35015	False	1.03395
147521	0.90000	0.90000	7.87379	7.87379	False	-0.92611
147523	0.51000	0.50000	12.89972	13.00000	False	-0.69314
147531	0.73000	0.72000	10.54875	10.66863	False	-0.75945
147537	0.82000	0.82000	9.33854	9.33854	False	-1.03329
147540	0.80000	0.80000	9.63352	9.63352	False	-1.05488
147545	0.76000	0.76000	10.17479	10.17479	False	-1.09449
147547	0.90000	0.89000	7.87379	8.09389	False	-0.41497
147557	0.57000	0.54000	12.29450	12.59827	False	0.35957
147562	0.51500	0.47500	12.84957	13.25083	False	0.92581

continued

IREF	P		Delta		Discard	Std
	Old	New	Old	New		
147600	0.78000	0.78000	9.91123	9.91123	False	-1.07520
172442	0.89000	0.90000	8.09389	7.87379	False	0.24432
172446	0.60000	0.63000	11.98661	11.67259	False	0.46586
172450	0.86000	0.85000	8.67872	8.85427	False	-0.59624
172451	0.94000	0.94000	6.78091	6.78091	False	-0.84614
172468	0.48000	0.43750	13.20061	13.62924	False	1.09904
172653	0.63000	0.68000	11.67259	11.12920	False	1.72531
172658	0.73000	0.73000	10.54875	10.54875	False	-1.12185
172666	0.73000	0.75000	10.54875	10.30204	False	0.20816
172675	0.55000	0.57000	12.49735	12.29450	False	-0.17086
172678	0.80000	0.79000	9.63352	9.77432	False	-0.71368
172697	0.63000	0.64000	11.67259	11.56616	False	-0.63036
172699	0.50000	0.50000	13.00000	13.00000	False	-1.22639
172700	0.54000	0.54000	12.59827	12.59827	False	-1.25579
172704	0.49500	0.50000	13.05013	13.00000	False	-1.03461
172709	0.52000	0.48000	12.79939	13.20061	False	0.92197
172719	0.60000	0.65000	11.98661	11.45872	False	1.61883
172734	0.70000	0.75000	10.90240	10.30204	False	2.08882
207571	0.74000	0.74000	10.42662	10.42662	False	-1.11291
207576	0.71000	0.72000	10.78646	10.66863	False	-0.50403
207580	0.84000	0.83000	9.02217	9.18334	False	-0.64859
207586	0.82000	0.80000	9.33854	9.63352	False	0.09591
207600	0.47000	0.43750	13.30108	13.62924	False	0.56478
207859	0.81000	0.81000	9.48841	9.48841	False	-1.04426
207860	0.67000	0.65000	11.24035	11.45872	False	-0.17791
207862	0.78000	0.77000	9.91123	10.04461	False	-0.73333
207867	0.43000	0.45000	13.70550	13.50265	False	-0.25926
207875	0.70000	0.67000	10.90240	11.24035	False	0.44201
207897	0.85000	0.86000	8.85427	8.67872	False	-0.05149
207903	0.66000	0.63000	11.35015	11.67259	False	0.39116
207912	0.80000	0.77000	9.63352	10.04461	False	0.74351
207912	0.81000	0.77000	9.48841	10.04461	False	1.51513
207914	0.53750	0.53000	12.62345	12.69892	False	-0.84708
207921	0.42000	0.49250	13.80757	13.07520	False	2.58793
207956	0.86000	0.88000	8.67872	8.30005	False	1.05641
207959	0.75000	0.74000	10.30204	10.42662	False	-0.75221
207964	0.82000	0.78000	9.33854	9.91123	False	1.59306
207969	0.68000	0.71000	11.12920	10.78646	False	0.68342
207969	0.69000	0.71000	11.01660	10.78646	False	0.08460
207974	0.50750	0.56500	12.92480	12.34537	False	1.82802
208271	0.84000	0.83000	9.02217	9.18334	False	-0.64859
208284	0.80000	0.79000	9.63352	9.77432	False	-0.71368
208289	0.72000	0.70000	10.66863	10.90240	False	-0.13676

Table J-26. 2012–13 NECAP Rescore Analysis
Reading Grade 8

IREF	*Max*	*Mean*		*St Dev*		*Effect Size*	*Discard*
		Old	*New*	*Old*	*New*		
207600	4	1.93162	1.85836	0.94090	0.99606	-0.07786	False
147557	4	2.26601	2.19704	1.04081	1.17121	-0.06626	False
147562	4	2.18227	2.02956	1.02317	1.05537	-0.14925	False
172468	4	1.89216	1.61275	0.97986	1.02102	-0.28515	False
172704	4	2.21287	2.15347	0.97515	1.00122	-0.06092	False
172709	4	2.13235	1.95098	1.07493	1.05670	-0.16873	False

Figure J-14. 2012–13 NECAP: Delta Plot
Reading Grade 11

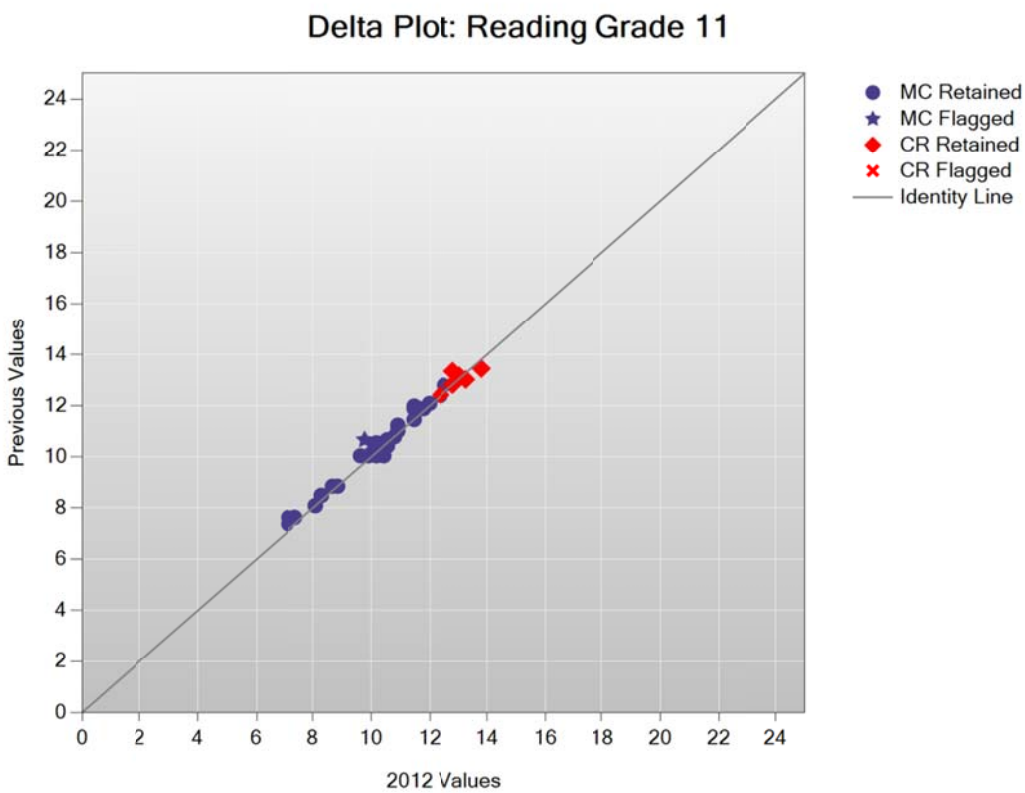


Table J-27. 2012–13 NECAP Delta Analysis
Reading Grade 11

IREF	*P*		*Delta*		*Discard*	*Std*
	Old	*New*	*Old*	*New*		
118874	0.87000	0.88000	8.49444	8.30005	False	-0.92999
118876	0.72000	0.73000	10.66863	10.54875	False	-0.84691
118878	0.69000	0.70000	11.01660	10.90240	False	-0.87000
118880	0.73000	0.76000	10.54875	10.17479	False	0.18900
118883	0.45500	0.42000	13.45215	13.80757	False	1.47547
147735	0.91000	0.92000	7.63698	7.37971	False	-0.96230

continued

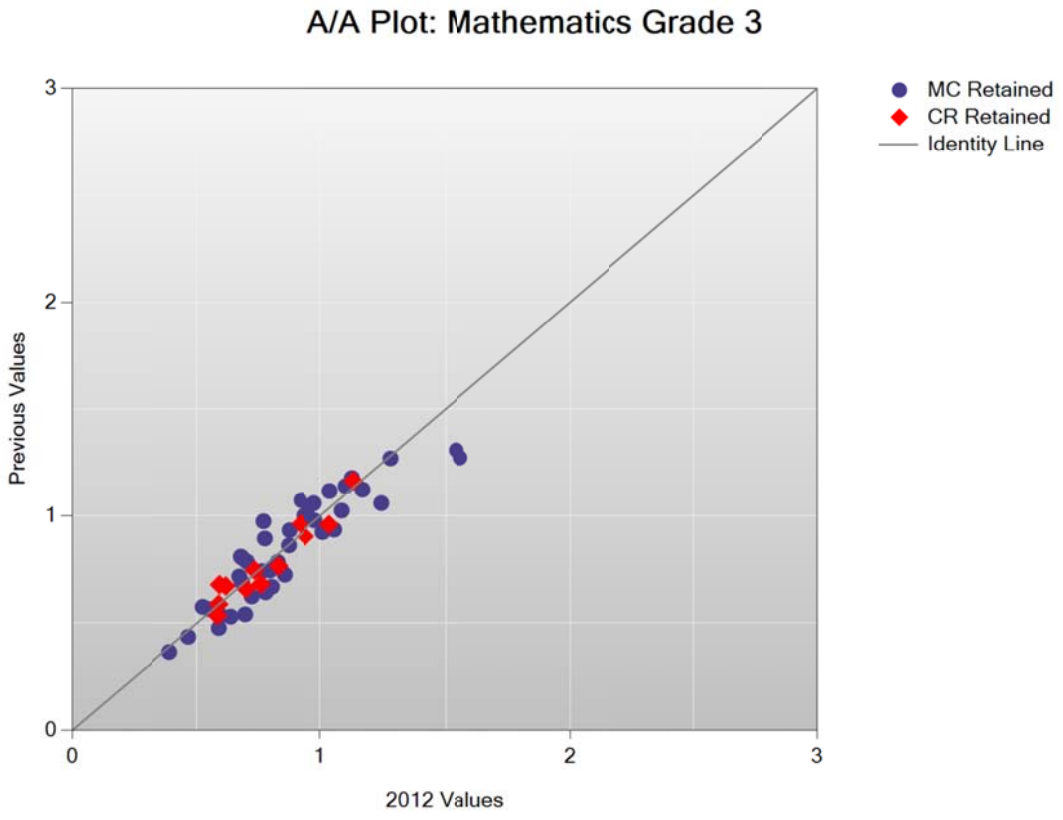
IREF	P		Delta		Discard	Std
	Old	New	Old	New		
147736	0.91000	0.92000	7.63698	7.37971	False	-0.96230
147747	0.74000	0.73000	10.42662	10.54875	False	0.60381
147749	0.59000	0.60000	12.08982	11.98661	False	-0.97936
147753	0.77000	0.80000	10.04461	9.63352	False	0.32413
147764	0.65000	0.65000	11.45872	11.45872	False	-0.27537
147766	0.77000	0.78000	10.04461	9.91123	False	-0.82478
147767	0.76000	0.77000	10.17479	10.04461	False	-0.82711
147774	0.55750	0.56000	12.42146	12.39612	False	-0.57880
147775	0.49750	0.47250	13.02507	13.27595	False	0.93476
208312	0.72000	0.79000	10.66863	9.77432	True	3.24348
208324	0.52000	0.55000	12.79939	12.49735	False	0.13327
208339	0.61000	0.63000	11.88272	11.67259	False	-0.55089
208343	0.61000	0.62000	11.88272	11.77808	False	-0.95428
208350	0.77000	0.74000	10.04461	10.42662	False	2.18134
208358	0.56000	0.56000	12.39612	12.39612	False	-0.42689
208380	0.67000	0.70000	11.24035	10.90240	False	0.09076
208381	0.71000	0.71000	10.78646	10.78646	False	-0.16670
208391	0.48250	0.49500	13.17552	13.05013	False	-0.83625
208423	0.85000	0.85000	8.85427	8.85427	False	0.14562
208424	0.87000	0.88000	8.49444	8.30005	False	-0.92999
208425	0.85000	0.86000	8.85427	8.67872	False	-0.87827
208426	0.73000	0.76000	10.54875	10.17479	False	0.18900
208426	0.77000	0.76000	10.04461	10.17479	False	0.71250
208430	0.72000	0.73000	10.66863	10.54875	False	-0.84691
208432	0.91000	0.93000	7.63698	7.09684	False	0.68764
208432	0.92000	0.93000	7.37971	7.09684	False	-0.85450
208434	0.61000	0.65000	11.88272	11.45872	False	0.69654
208434	0.60000	0.65000	11.98661	11.45872	False	1.31927
208437	0.89000	0.89000	8.09389	8.09389	False	0.26853
208437	0.89000	0.89000	8.09389	8.09389	False	0.26853
208474	0.51750	0.51750	12.82448	12.82448	False	-0.49613
208476	0.46500	0.51750	13.35138	12.82448	False	1.53408

Table J-28. 2012–13 NECAP Rescore Analysis
Reading Grade 11

IREF	Max	Mean		St Dev		Effect Size	Discard
		Old	New	Old	New		
208474	4	2.21661	2.05295	0.95502	0.92728	-0.17137	False
208476	4	1.90291	2.01456	0.77650	0.76037	0.14379	False
118883	4	1.89268	1.61951	0.96745	0.94364	-0.28236	False
147774	4	2.32020	2.15764	1.10193	1.09467	-0.14752	False
147775	4	2.05419	1.85714	0.98423	1.00518	-0.20020	False
208391	4	1.91584	1.85644	1.07597	1.13268	-0.05521	False

APPENDIX K—*α*-PLOTS AND *b*-PLOTS

Figure K-1. 2012–13 NECAP: Mathematics Grade 3 Plots
Top: *α*-Plot Bottom: *b*-Plot



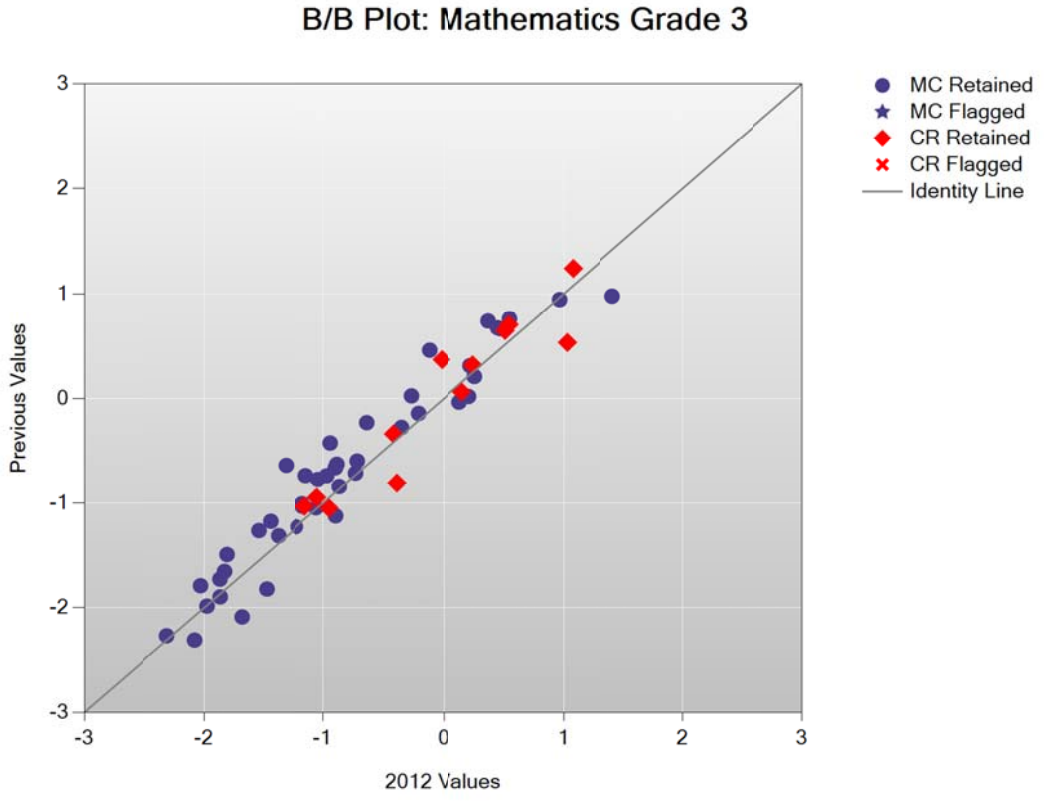
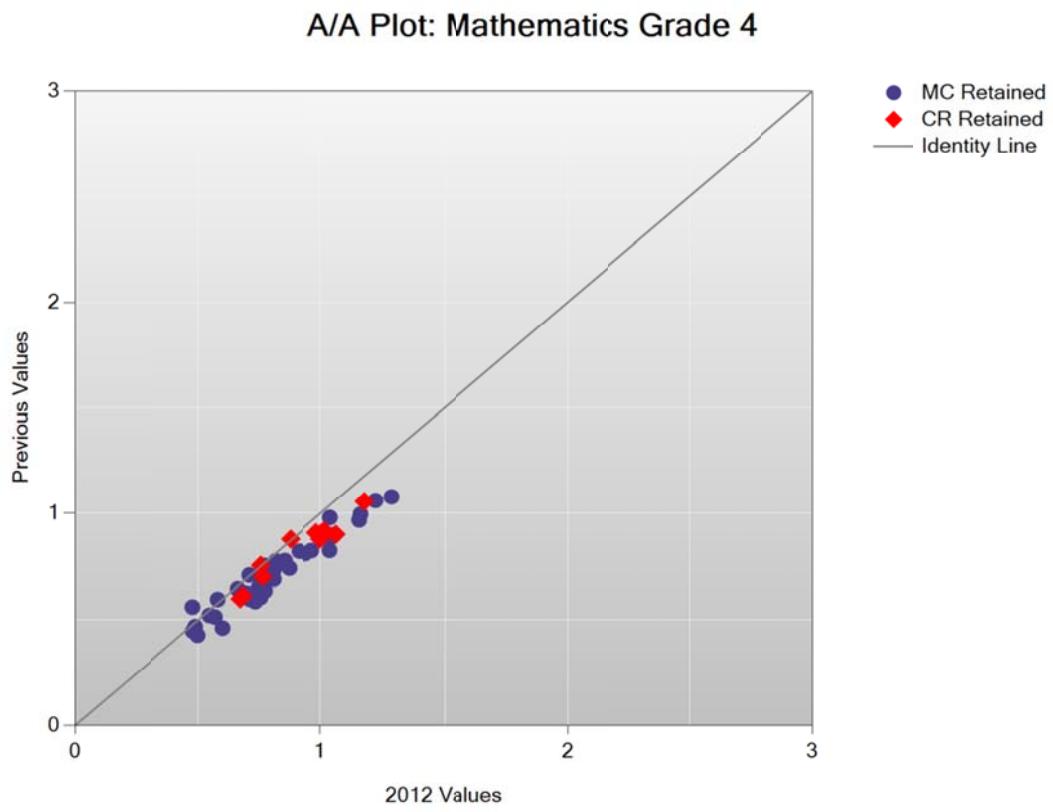


Figure K-2. 2012–13 NECAP: Mathematics Grade 4 Plots
Top: *α*-Plot Bottom: *b*-Plot



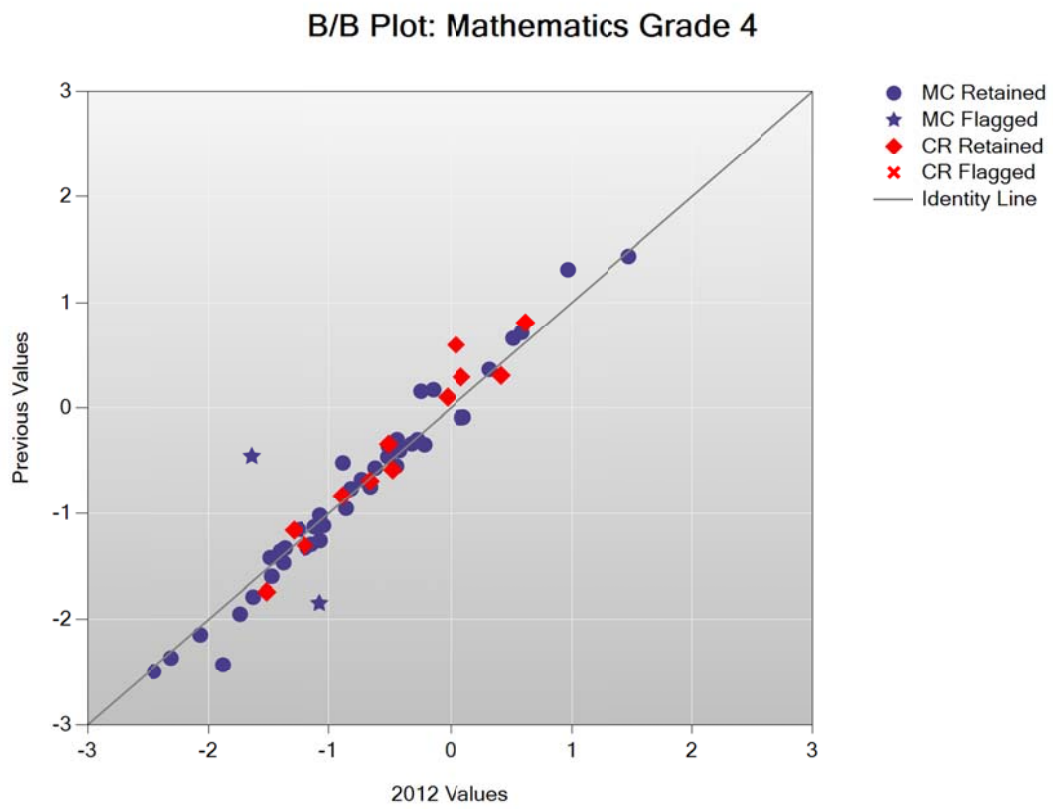
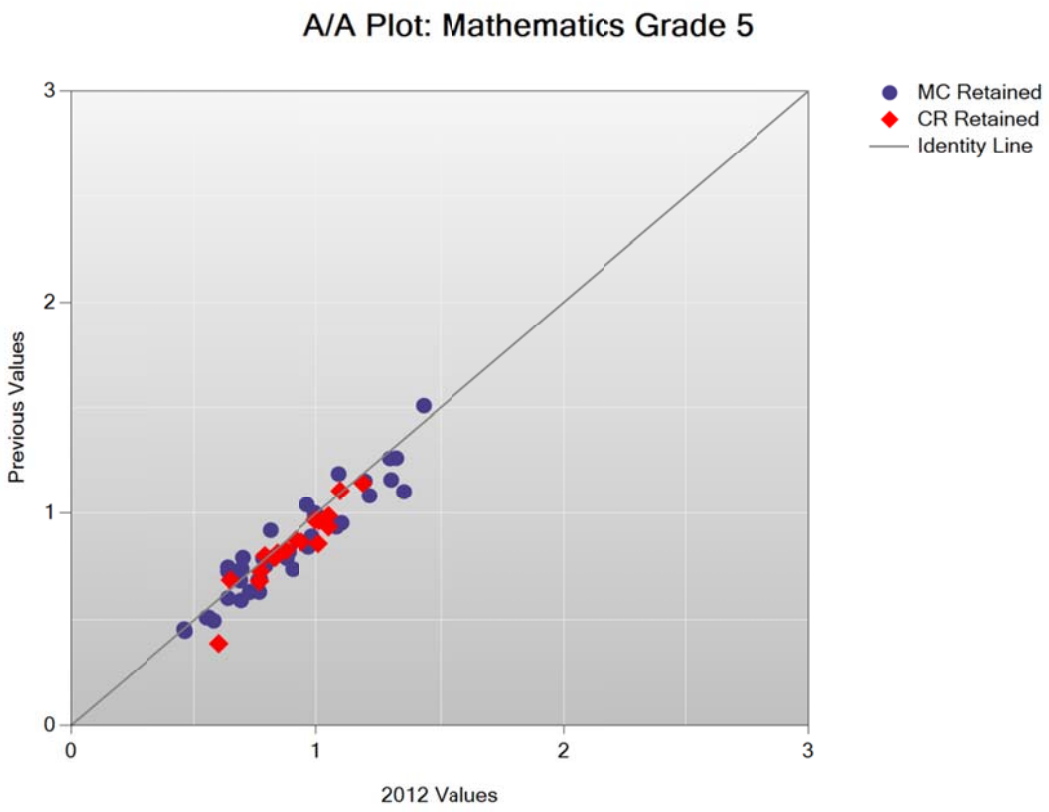


Figure K-3. 2012–13 NECAP: Mathematics Grade 5 Plots
Top: *α*-Plot Bottom: *b*-Plot



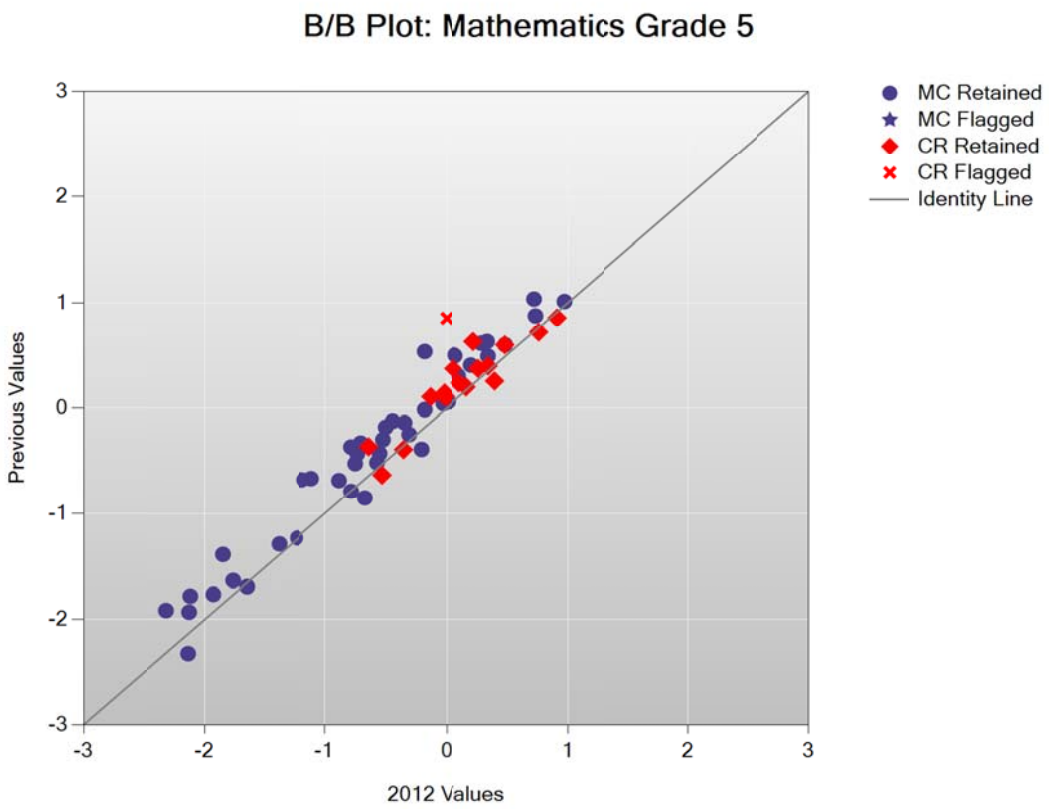
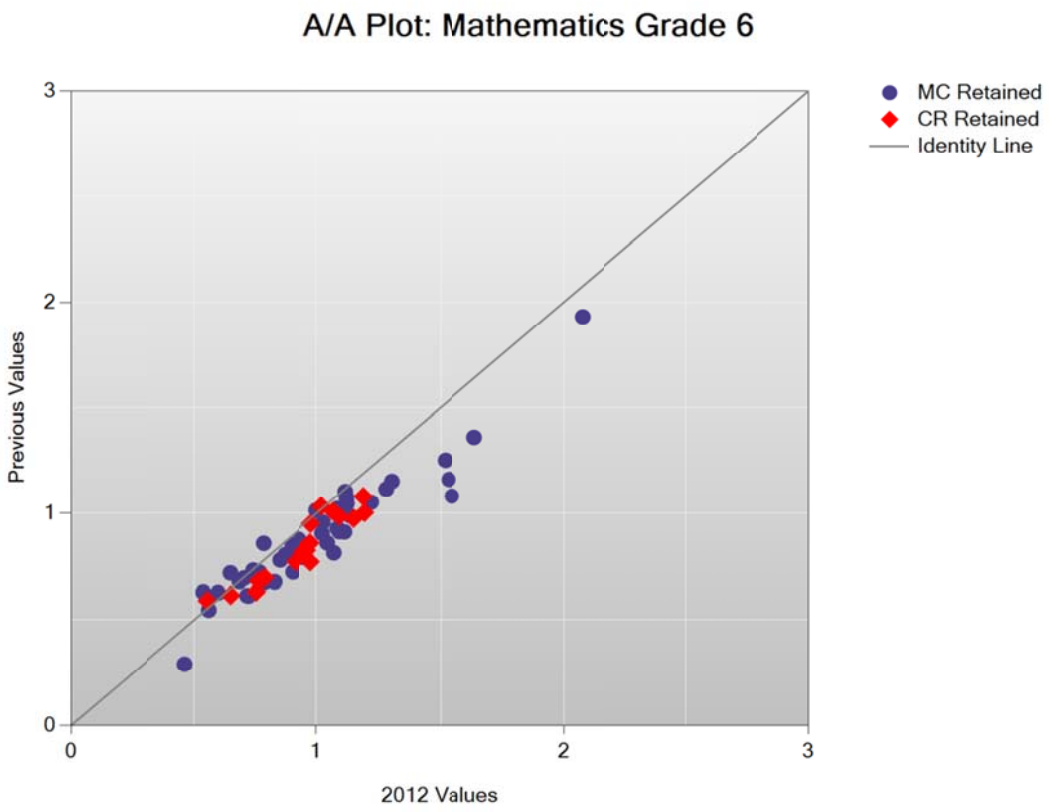


Figure K-4. 2012–13 NECAP: Mathematics Grade 6 Plots
Top: *α*-Plot Bottom: *b*-Plot



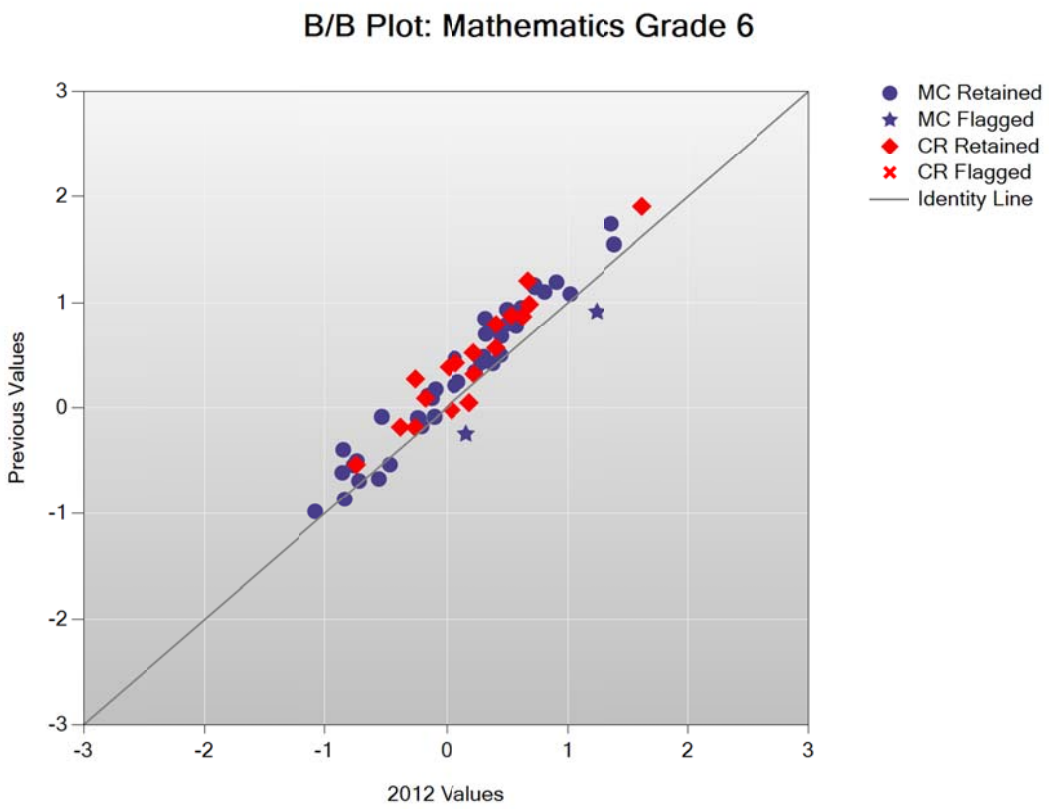
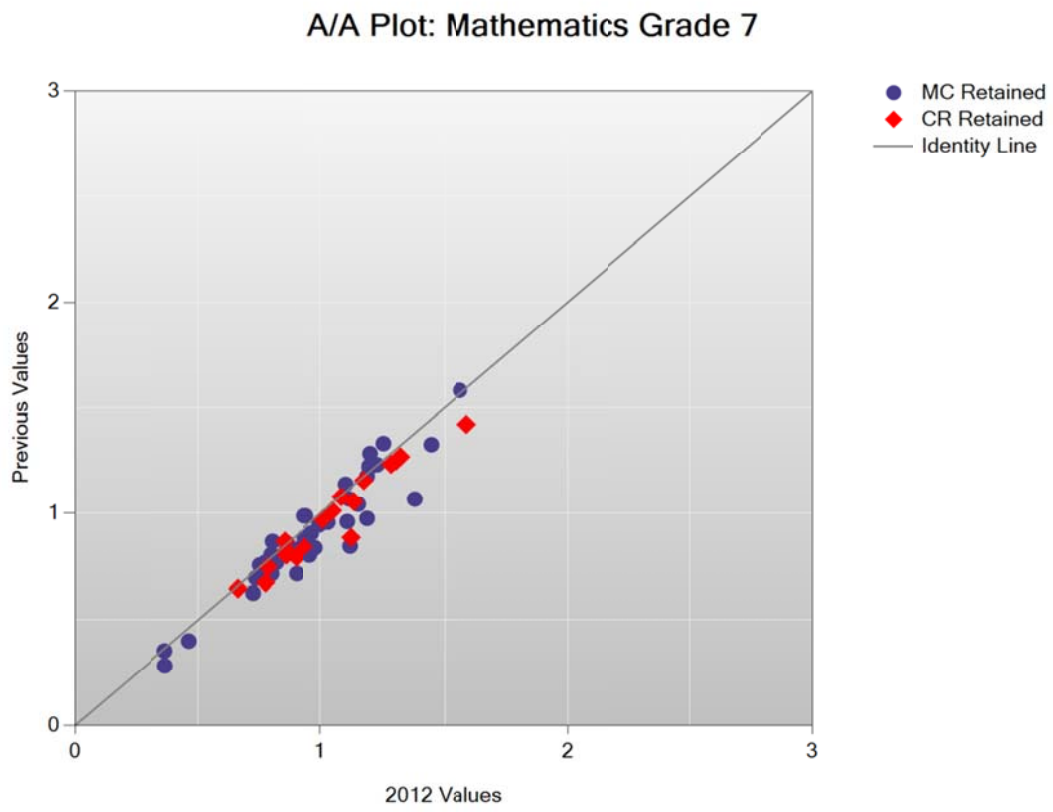


Figure K-5. 2012–13 NECAP: Mathematics Grade 7 Plots
Top: *α*-Plot Bottom: *b*-Plot



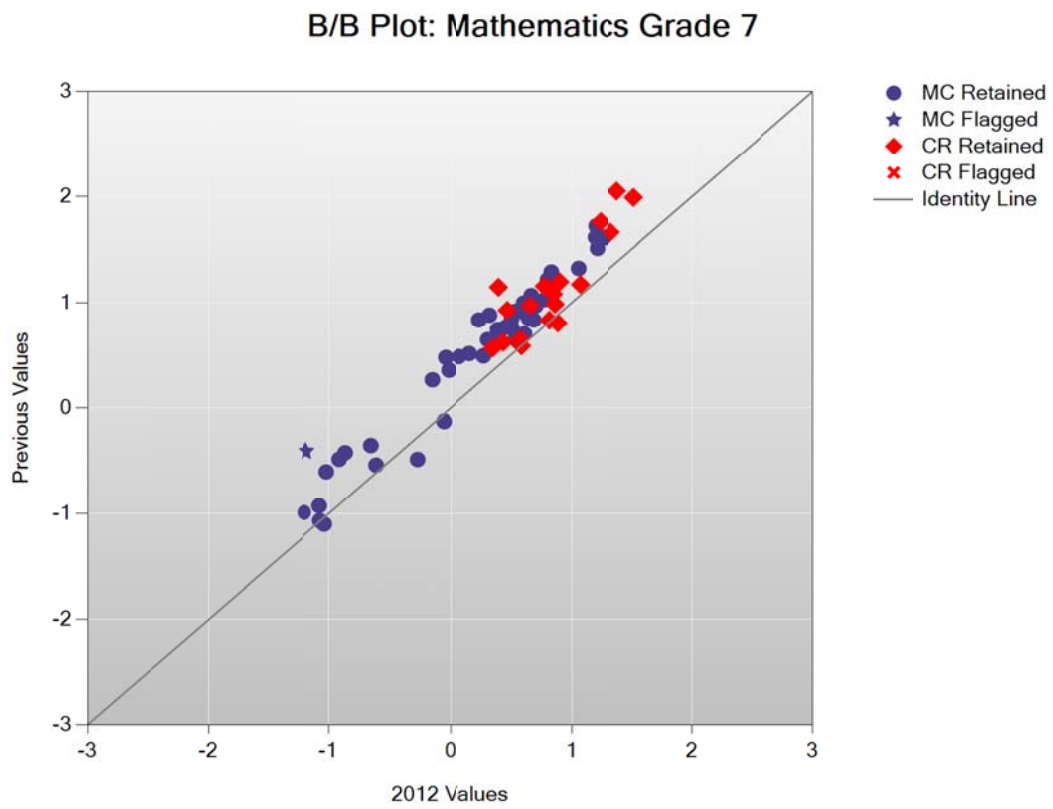
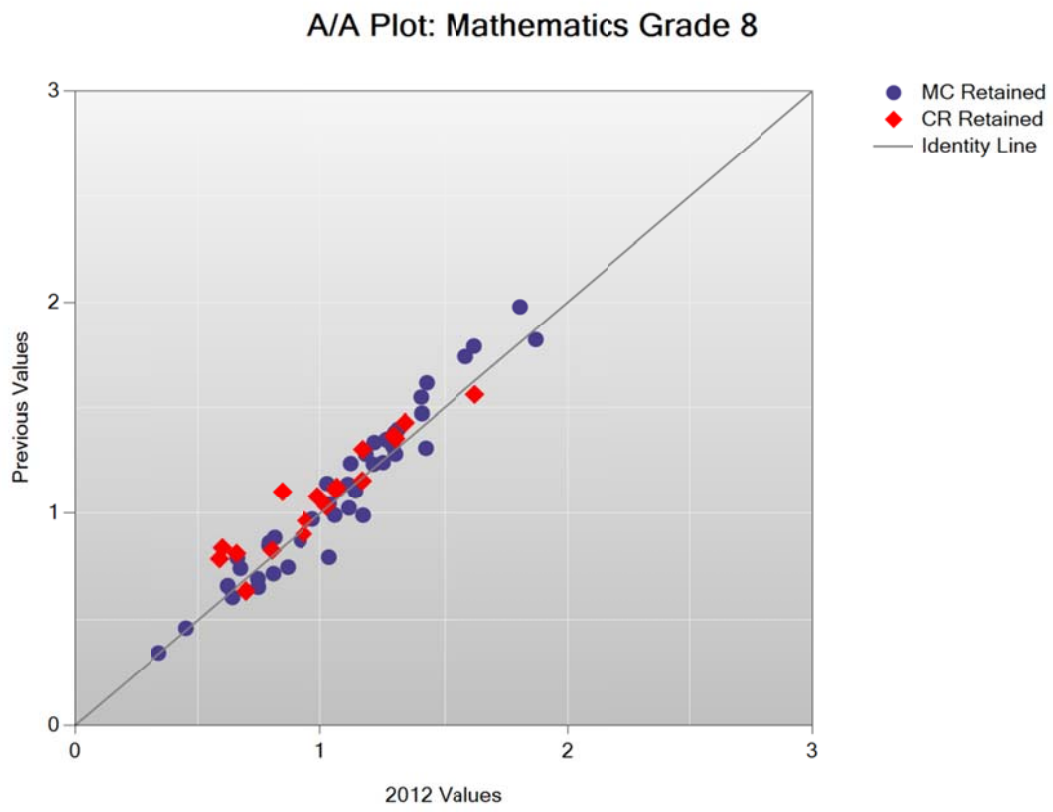


Figure K-6. 2012–13 NECAP: Mathematics Grade 8 Plots
Top: *α*-Plot Bottom: *b*-Plot



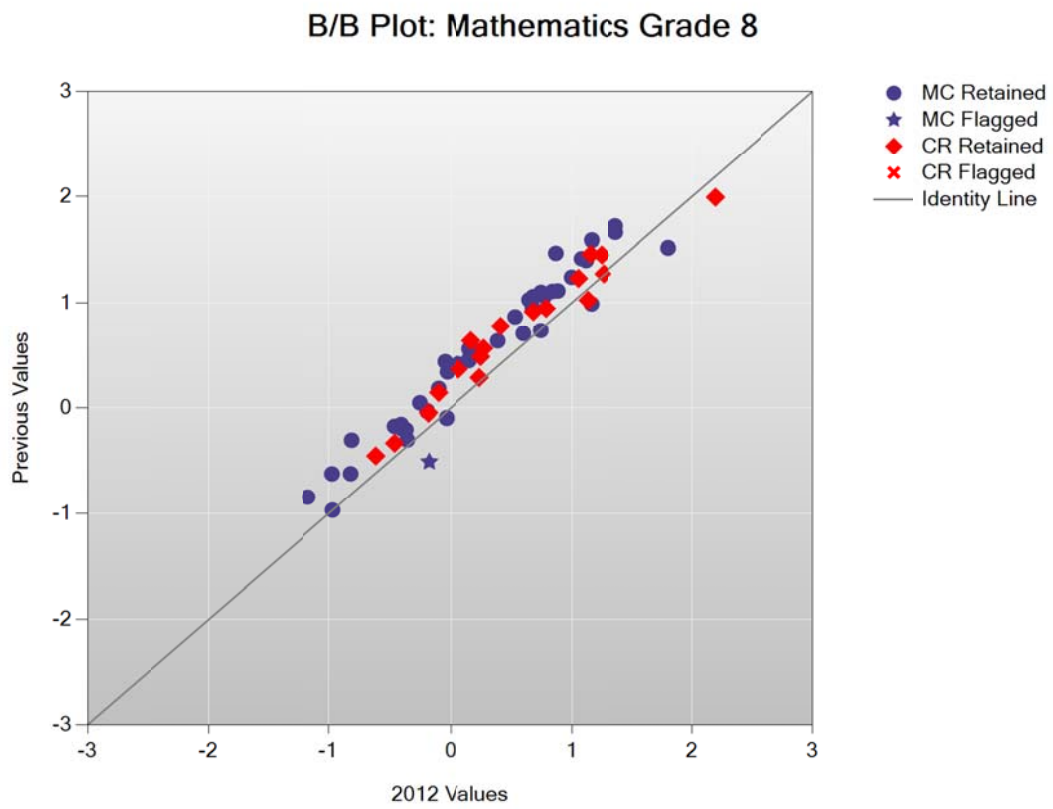
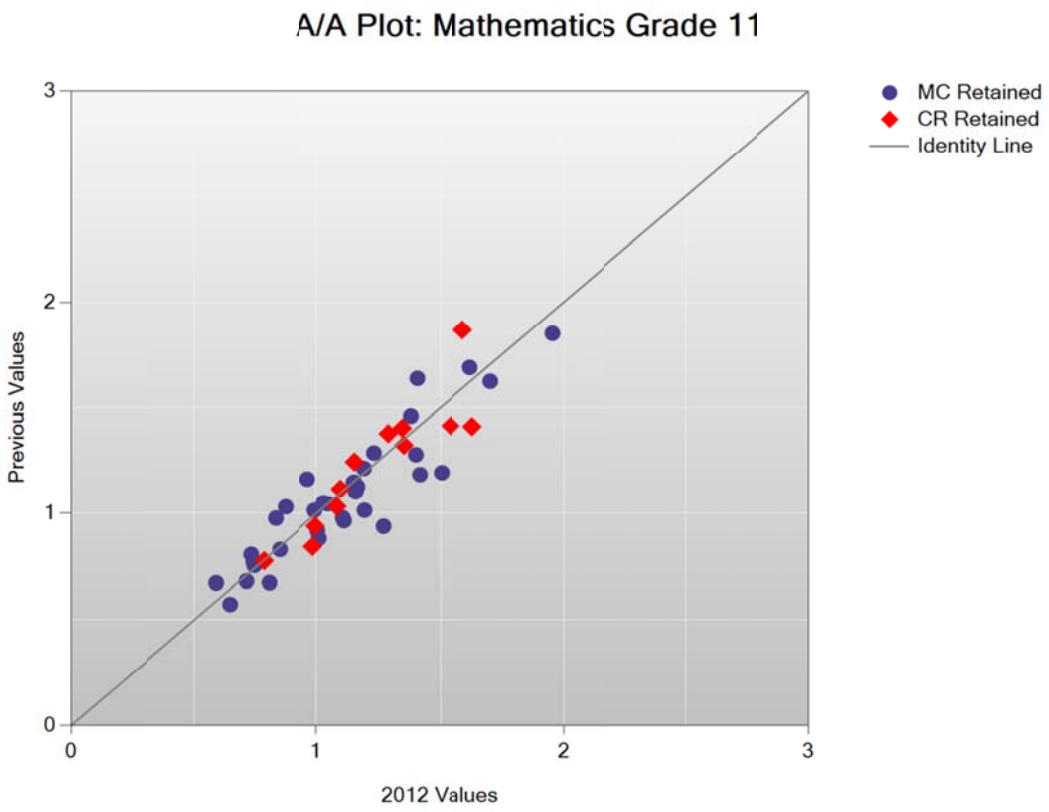


Figure K-7. 2012–13 NECAP: Mathematics Grade 11 Plots
Top: *α*-Plot Bottom: *b*-Plot



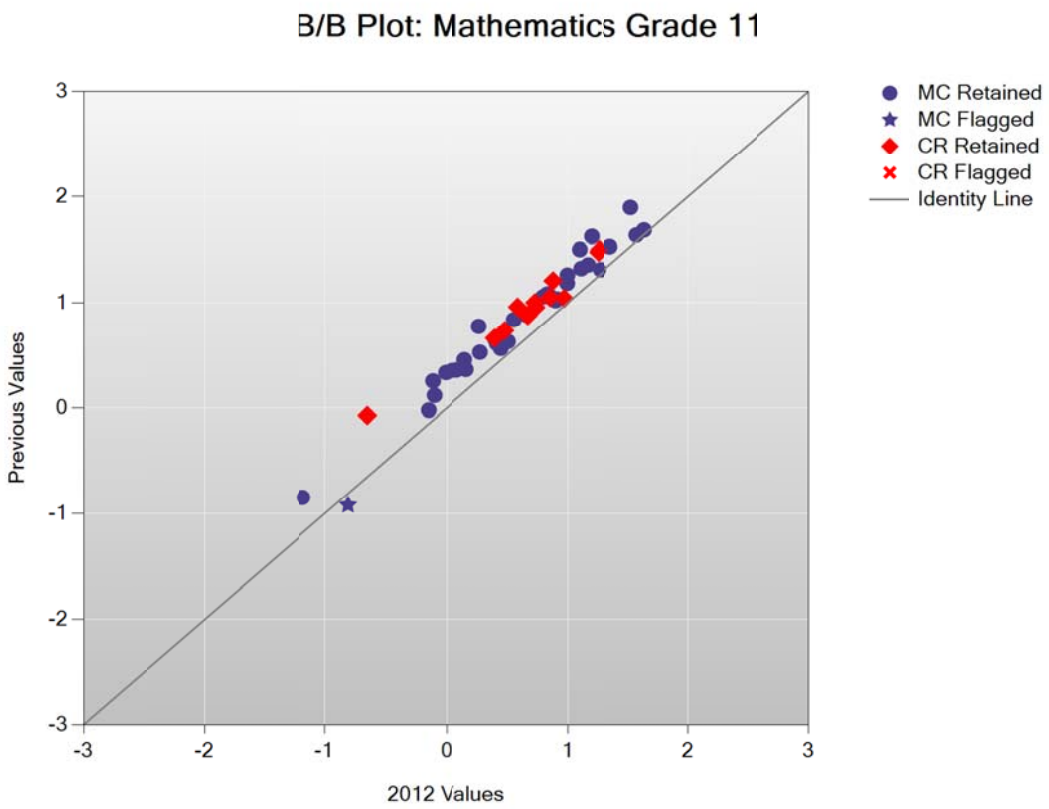
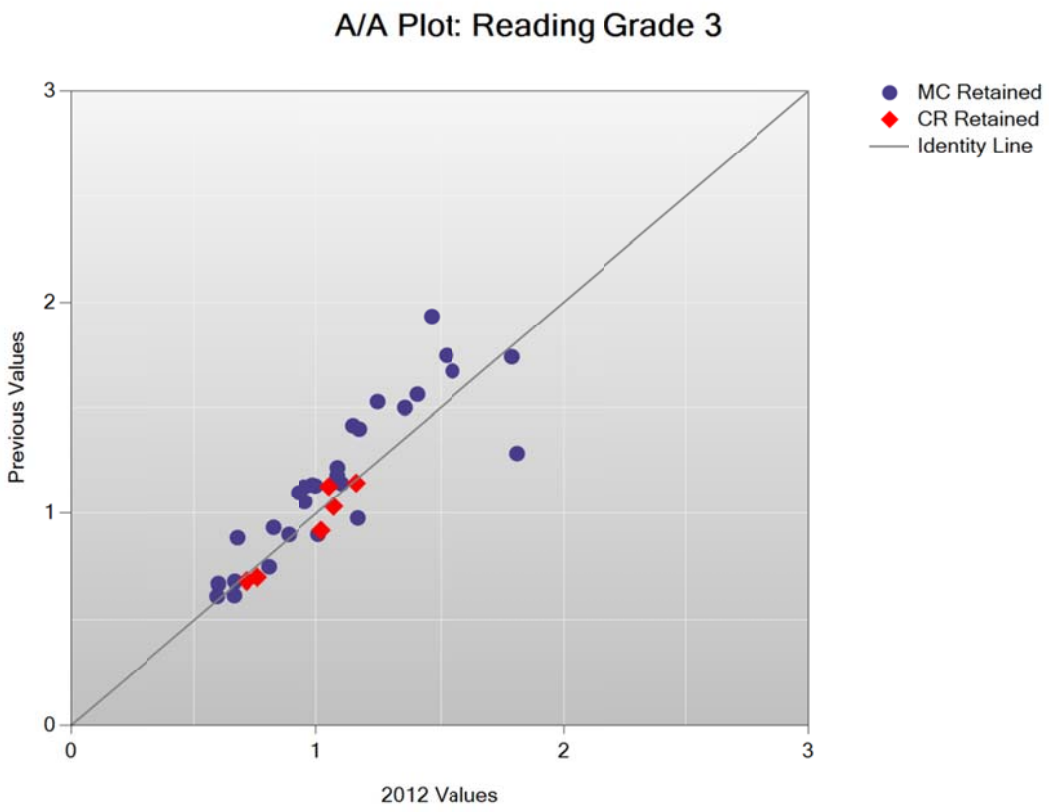


Figure K-8. 2012–13 NECAP: Reading Grade 3 Plots
Top: *α*-Plot Bottom: *b*-Plot



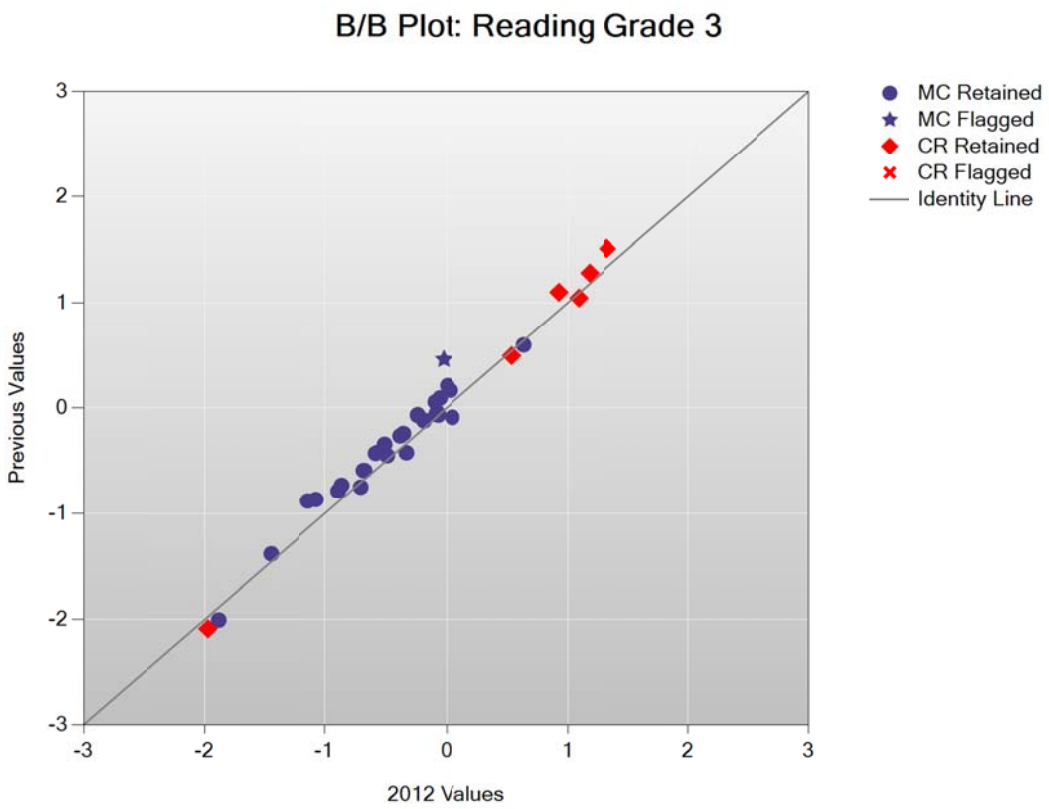
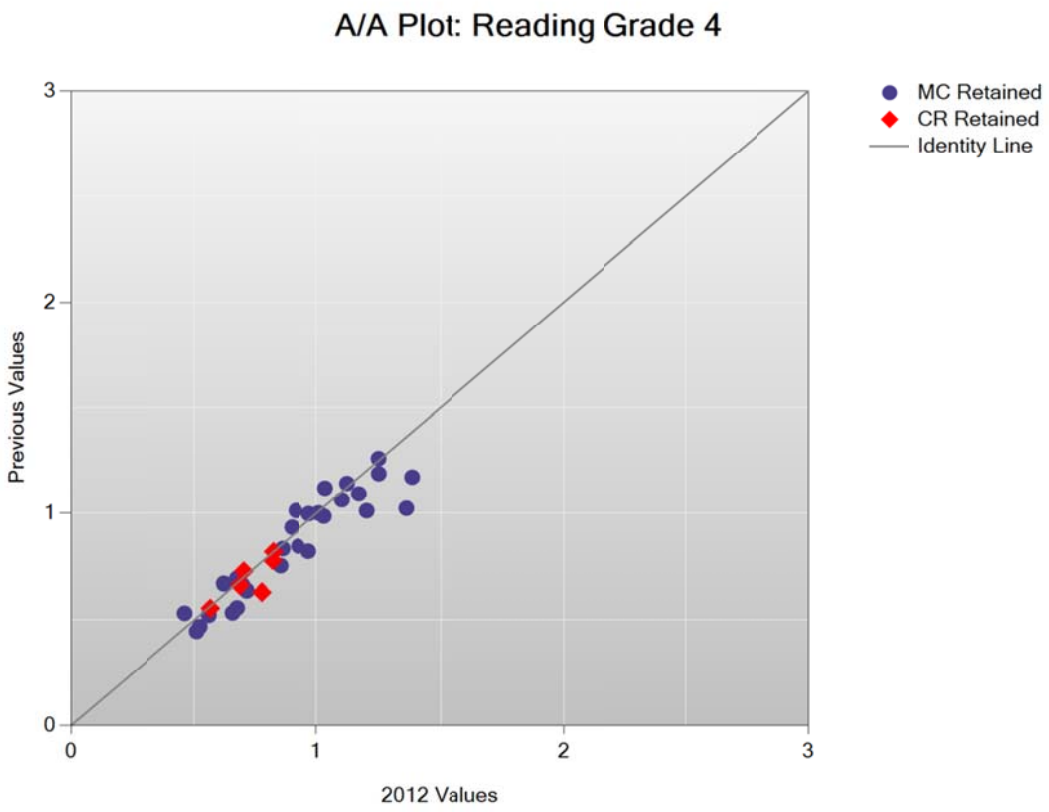


Figure K-9. 2012–13 NECAP: Reading Grade 4 Plots
Top: *α*-Plot Bottom: *b*-Plot



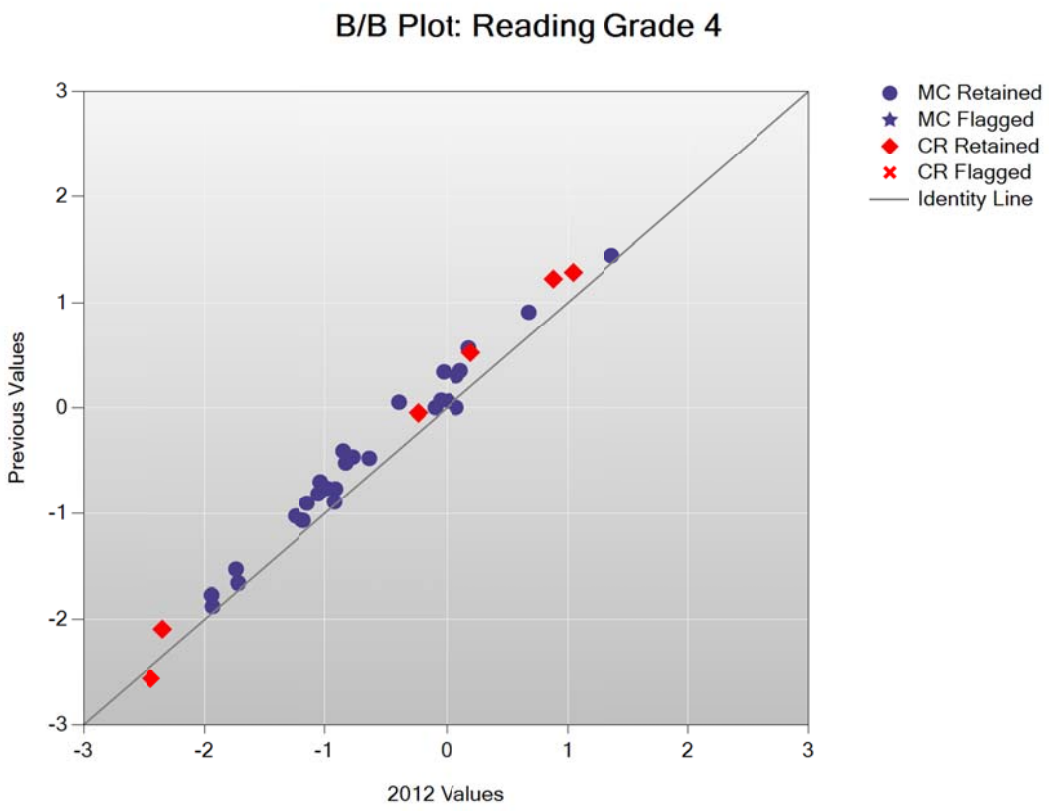
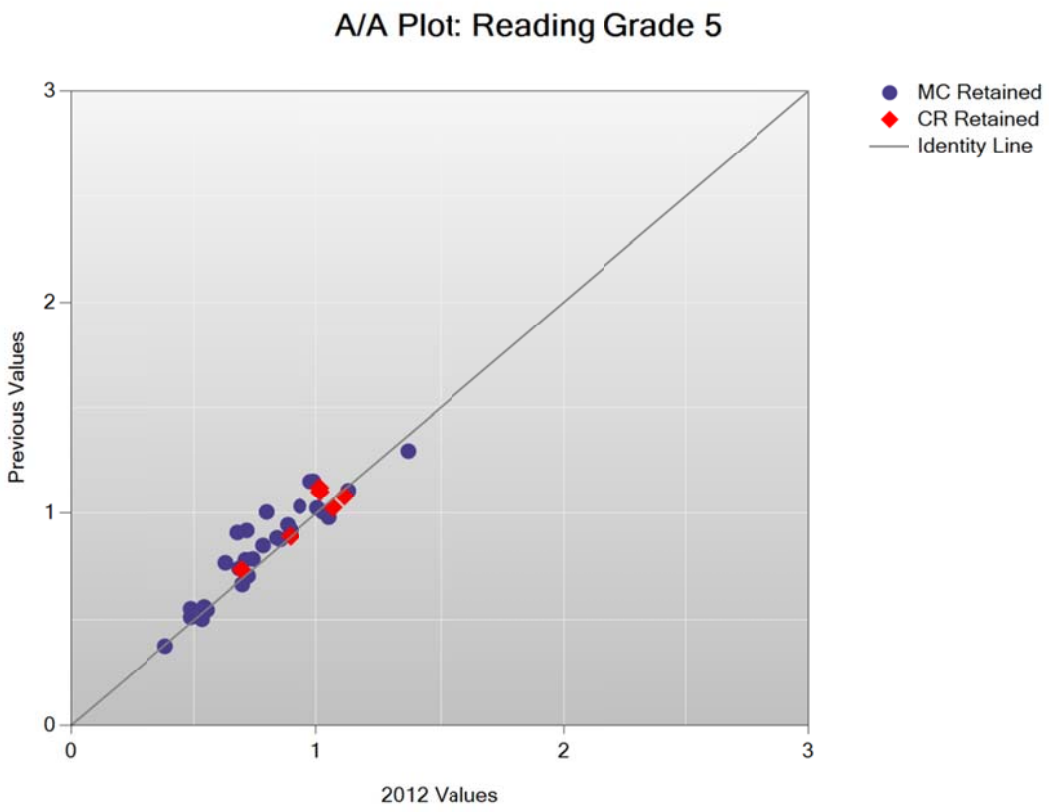


Figure K-10. 2012–13 NECAP: Reading Grade 5 Plots
Top: *α*-Plot Bottom: *b*-Plot



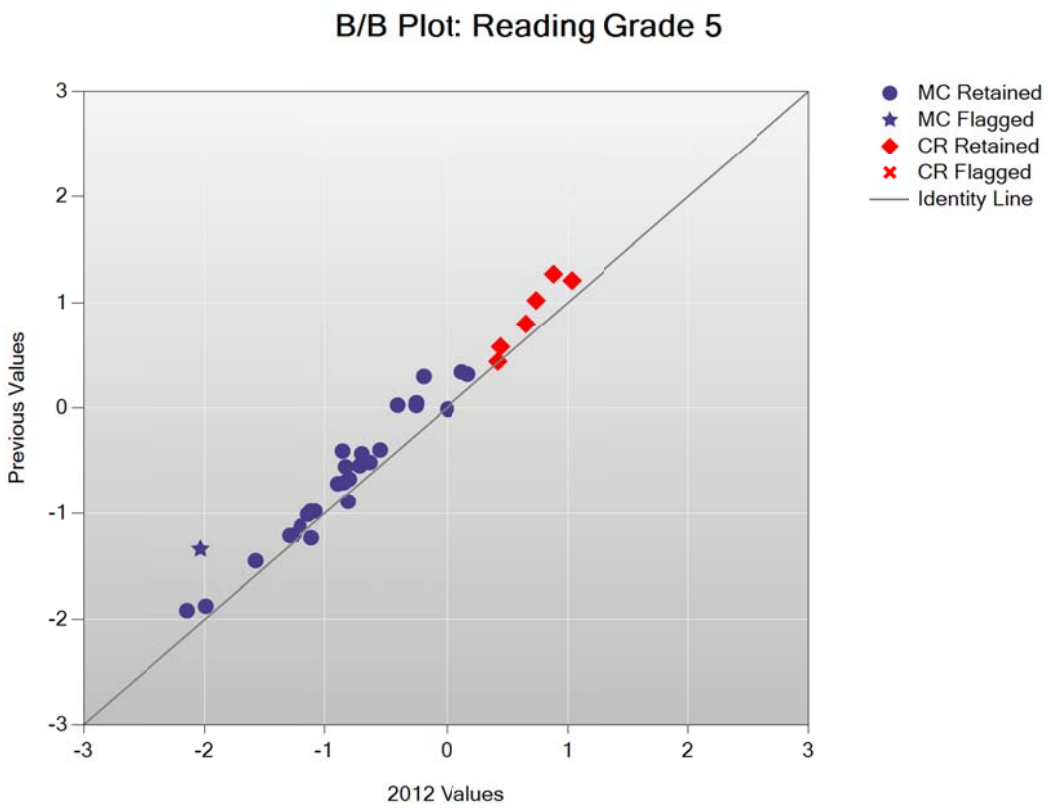
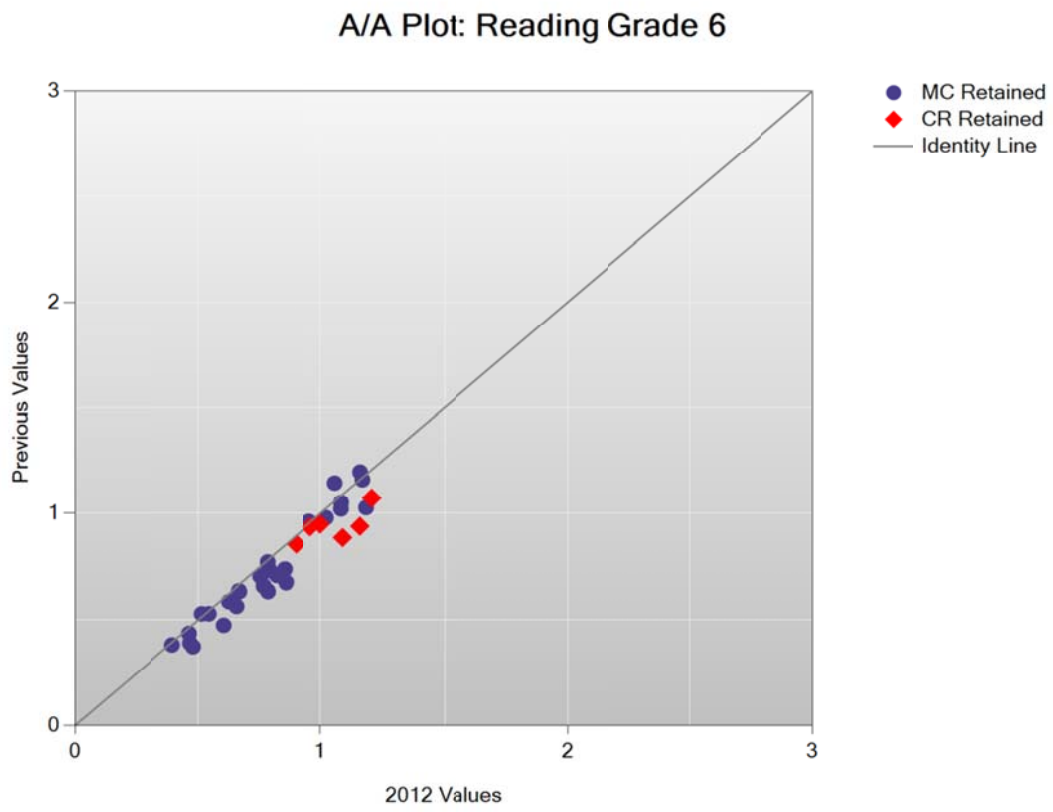


Figure K-11. 2012–13 NECAP: Reading Grade 6 Plots
Top: *α*-Plot Bottom: *b*-Plot



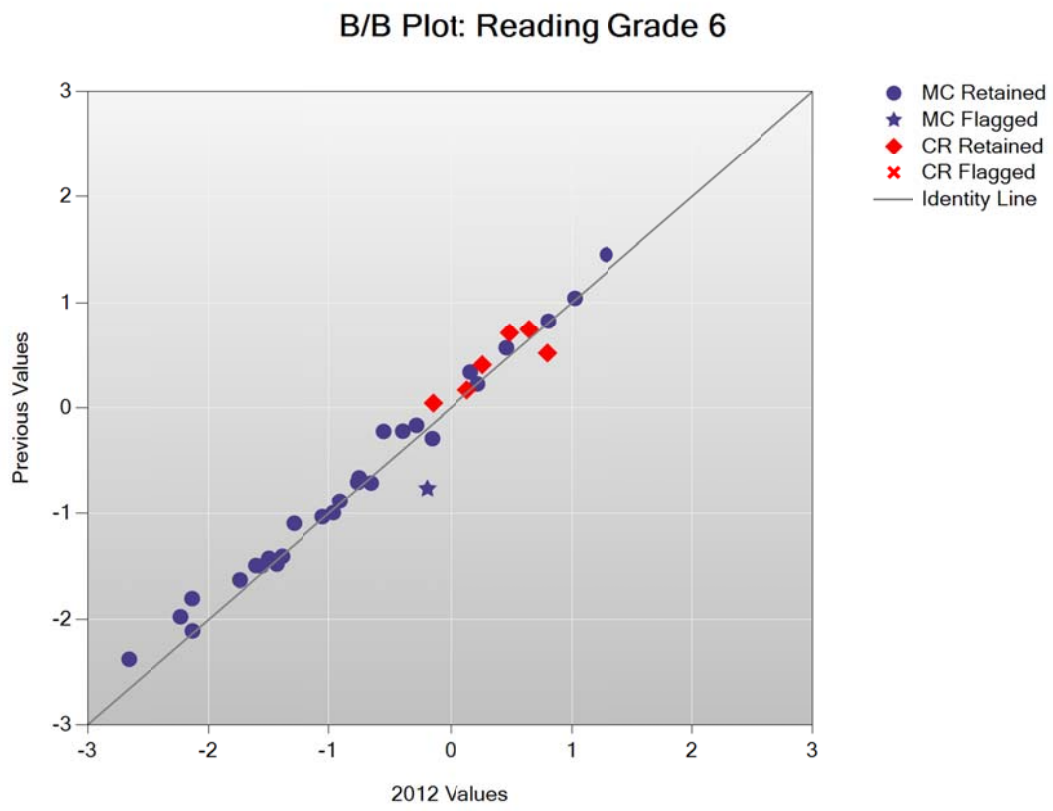
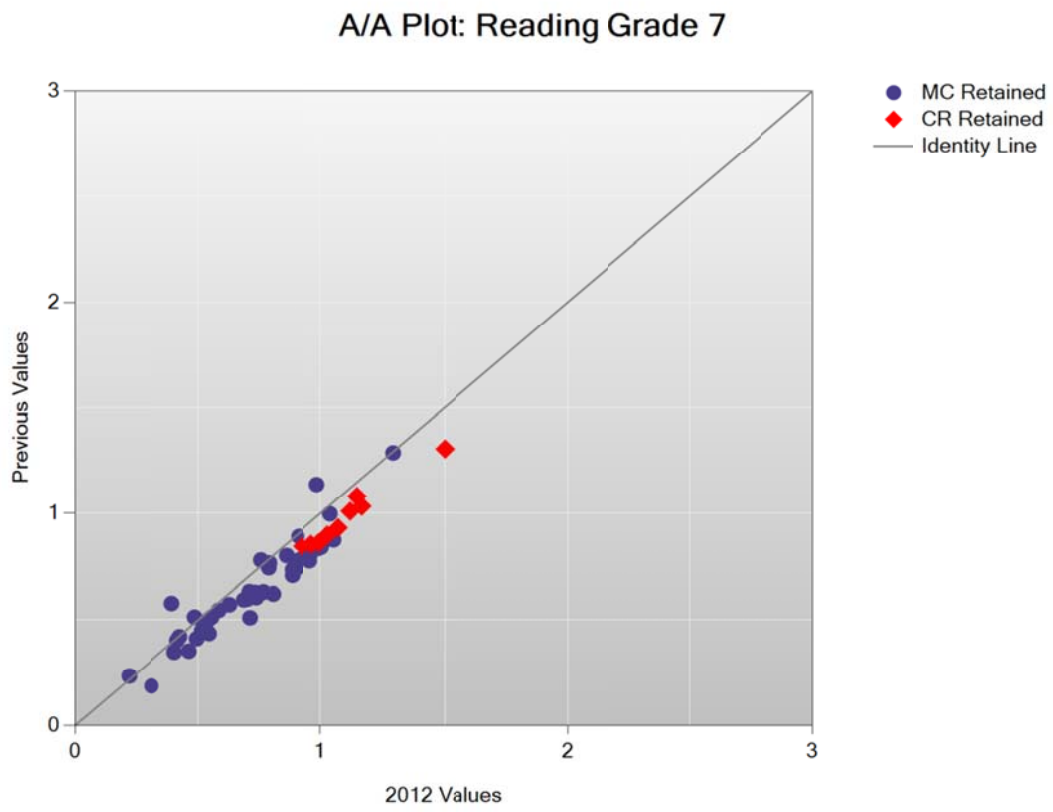


Figure K-12. 2012–13 NECAP: Reading Grade 7 Plots
Top: *α*-Plot Bottom: *b*-Plot



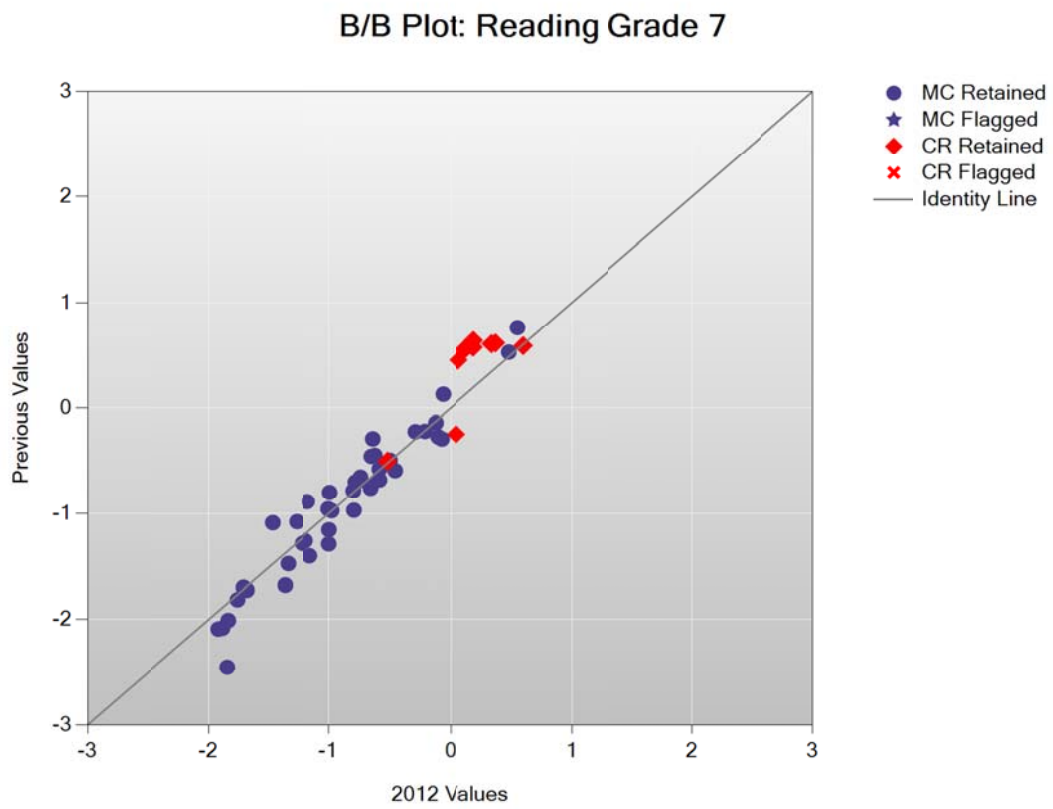
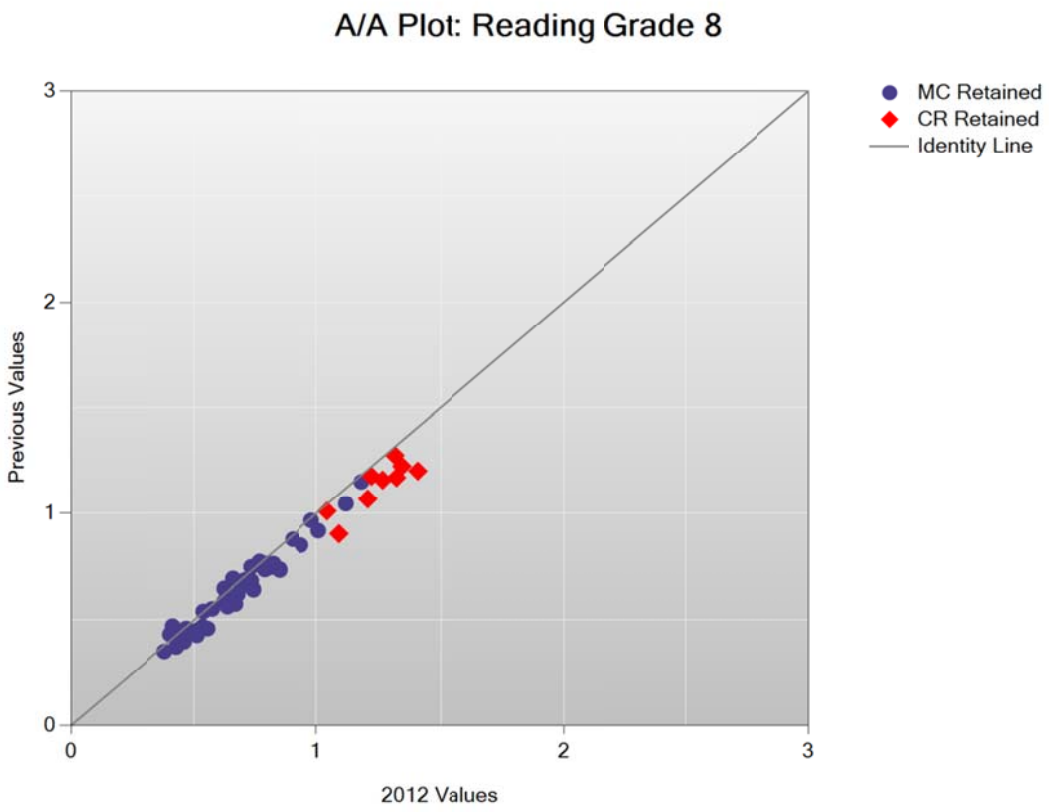


Figure K-13. 2012–13 NECAP: Reading Grade 8 Plots
Top: *α*-Plot Bottom: *b*-Plot



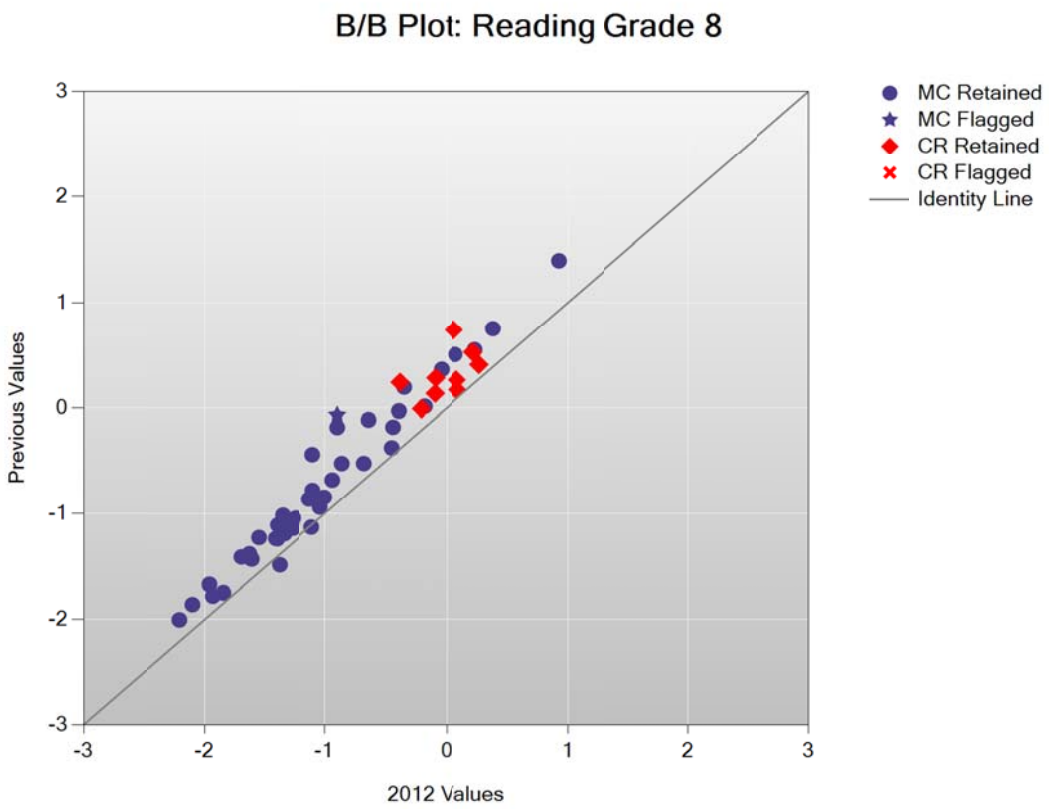
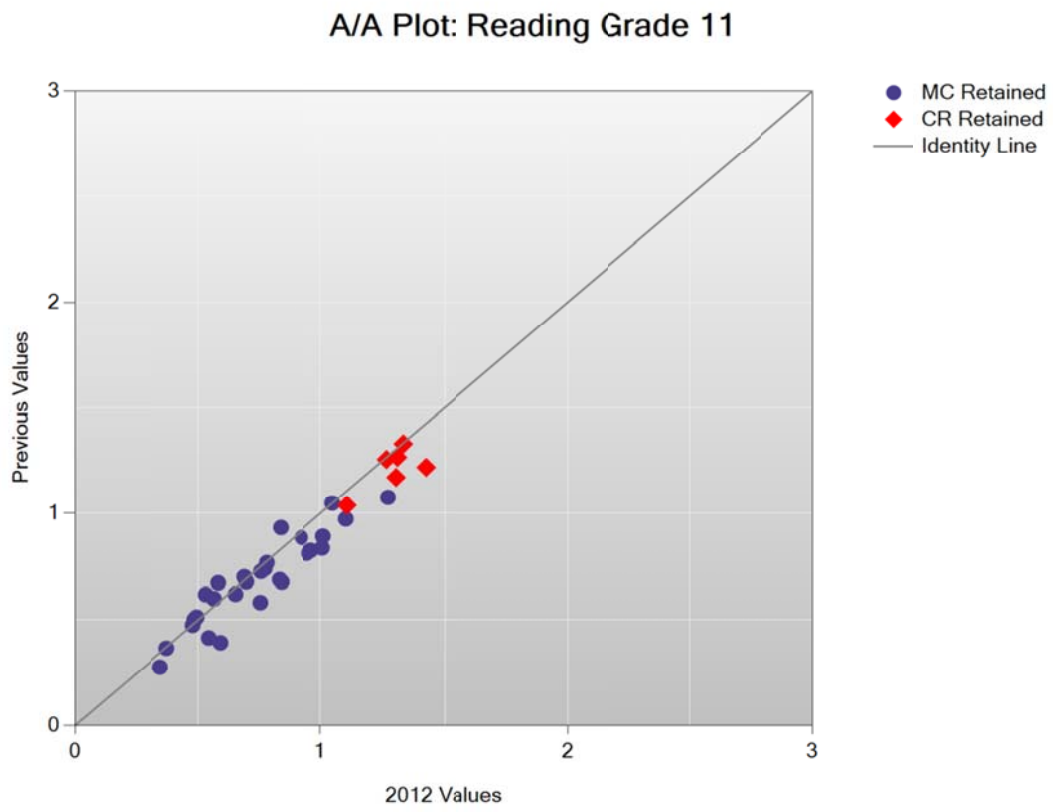
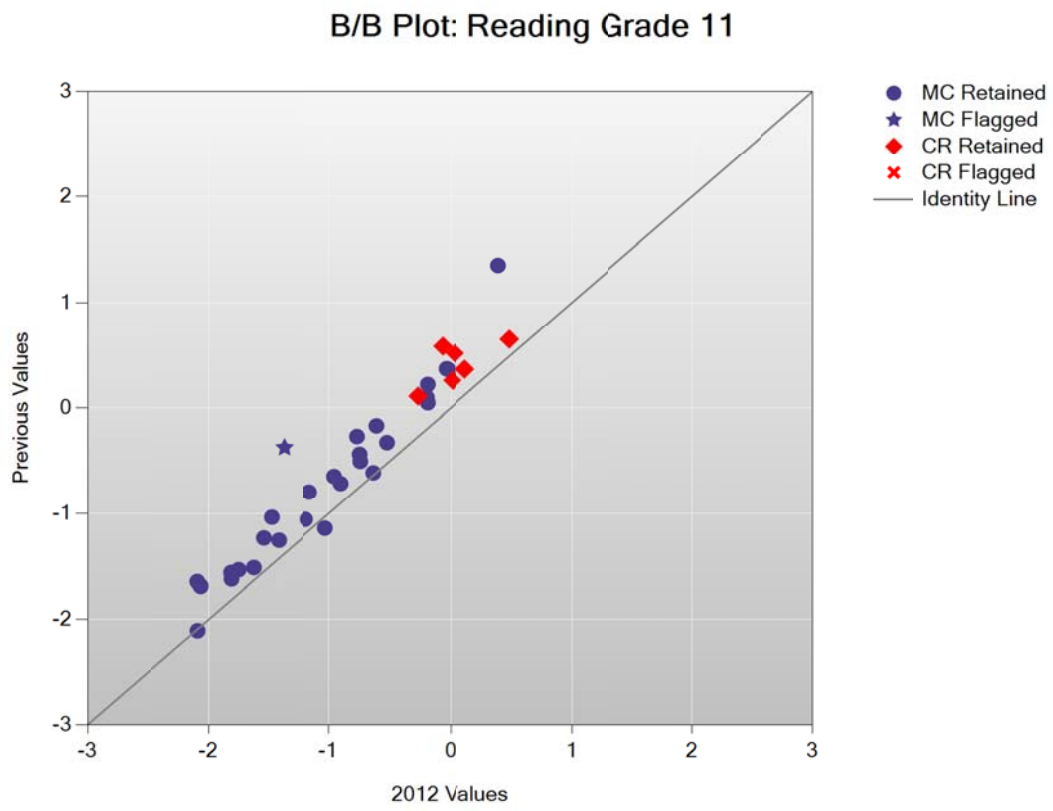


Figure K-14. 2012–13 NECAP: Reading Grade 11 Plots
Top: *α*-Plot Bottom: *b*-Plot





APPENDIX L—PERFORMANCE LEVEL SCORE DISTRIBUTIONS

Table L-1. 2012–13 NECAP: Performance Level Distributions by Grade—Mathematics

Grade	*Performance Level*	*Percent in Level*		
		2012-13	*2011-12*	*2010-11*
3	4	19.04	21.40	21.48
	3	46.30	45.81	45.52
	2	20.18	18.94	20.08
	1	14.48	13.85	12.92
4	4	23.83	23.61	20.62
	3	45.33	45.82	45.29
	2	17.59	17.11	20.72
	1	13.25	13.46	13.37
5	4	20.85	20.24	18.94
	3	45.60	47.57	46.74
	2	16.45	16.03	16.94
	1	17.10	16.15	17.38
6	4	23.55	24.63	23.77
	3	43.59	42.48	42.00
	2	15.57	15.80	17.15
	1	17.30	17.09	17.08
7	4	22.35	23.24	18.89
	3	40.03	39.23	41.44
	2	18.86	17.93	19.67
	1	18.76	19.59	20.01
8	4	18.09	19.51	18.33
	3	44.81	43.08	42.15
	2	17.99	19.46	20.34
	1	19.10	17.95	19.18
11	4	2.61	2.54	2.67
	3	33.98	31.62	32.30
	2	25.82	26.08	28.30
	1	37.59	39.77	36.73

Table L-2. 2012–13 NECAP: Performance Level Distributions by Grade—Reading

Grade	*Performance Level*	*Percent in Level*		
		2012-13	*2011-12*	*2010-11*
3	4	15.08	20.87	14.45
	3	56.87	54.16	59.19
	2	17.60	15.77	16.83
	1	10.44	9.20	9.53
4	4	19.63	22.55	21.03
	3	52.32	50.78	50.05
	2	17.95	16.99	19.16
	1	10.10	9.67	9.76
5	4	18.35	19.05	19.67
	3	55.18	52.15	54.21

continued

Grade	Performance Level	Percent in Level		
		2012-13	*2011-12*	*2010-11*
5	2	18.71	19.91	18.46
	1	7.76	8.89	7.66
6	4	15.91	21.77	16.61
	3	58.58	53.54	57.28
	2	17.67	17.27	18.31
	1	7.85	7.41	7.80
7	4	18.07	16.49	13.80
	3	54.36	56.72	54.43
	2	18.53	18.06	22.16
	1	9.04	8.73	9.61
8	4	23.52	25.07	21.99
	3	55.41	53.87	53.34
	2	15.42	15.11	19.31
	1	5.65	5.95	5.36
11	4	30.85	30.30	26.50
	3	46.07	45.56	47.43
	2	14.84	14.87	17.04
	1	8.24	9.27	9.04

Table L-3. 2012–13 NECAP: Performance Level Distributions by Grade—Writing

Grade	Performance Level	Percent in Level		
		2012-13	*2011-12*	*2010-11*
5	4	13.40	8.90	12.49
	3	39.61	40.41	39.62
	2	35.53	39.06	39.81
	1	11.46	11.63	8.08
8	4	15.16	9.93	10.91
	3	48.50	47.26	48.95
	2	26.88	32.19	31.83
	1	9.46	10.62	8.31
11	4	3.23	4.26	1.06
	3	41.86	43.25	46.96
	2	45.72	44.04	45.58
	1	9.19	8.45	6.40

APPENDIX M—RAW TO SCALED SCORE LOOK-UP TABLES

Table M-1. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 3

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	300	10.0	1	300	10.0	1
1	300	10.0	1	300	10.0	1
2	300	10.0	1	300	10.0	1
3	300	10.0	1	300	10.0	1
4	300	10.0	1	300	10.0	1
5	300	10.0	1	300	10.0	1
6	300	10.0	1	300	10.0	1
7	302	9.5	1	302	8.8	1
8	307	7.4	1	307	7.0	1
9	311	6.2	1	310	5.9	1
10	313	5.4	1	312	5.2	1
11	316	4.8	1	314	4.7	1
12	318	4.3	1	316	4.4	1
13	319	4.0	1	318	4.1	1
14	321	3.7	1	319	3.8	1
15	322	3.5	1	321	3.6	1
16	323	3.4	1	322	3.5	1
17	324	3.2	1	323	3.3	1
18	325	3.1	1	324	3.2	1
19	326	3.0	1	325	3.1	1
20	327	2.9	1	326	3.0	1
21	328	2.9	1	327	2.9	1
22	329	2.8	1	328	2.8	1
23	330	2.8	1	329	2.8	1
24	331	2.7	1	329	2.7	1
25	331	2.7	1	330	2.7	1
26	332	2.7	2	331	2.7	1
27	333	2.7	2	331	2.6	1
28	334	2.6	2	332	2.6	2
29	335	2.6	2	333	2.6	2
30	335	2.6	2	334	2.5	2
31	336	2.6	2	335	2.5	2
32	337	2.6	2	335	2.5	2
33	338	2.6	2	336	2.5	2
34	338	2.6	2	337	2.5	2
35	339	2.6	2	338	2.5	2
36	339	2.6	2	338	2.5	2
37	341	2.6	3	339	2.5	2
38	341	2.6	3	339	2.5	2
39	342	2.6	3	341	2.6	3
40	343	2.6	3	341	2.6	3
41	344	2.6	3	342	2.6	3
42	344	2.6	3	343	2.6	3
43	345	2.6	3	344	2.6	3
44	346	2.6	3	344	2.7	3

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
45	347	2.6	3	345	2.7	3
46	348	2.6	3	346	2.7	3
47	349	2.7	3	347	2.8	3
48	349	2.7	3	348	2.8	3
49	350	2.7	3	349	2.8	3
50	351	2.8	3	350	2.9	3
51	352	2.8	3	351	2.9	3
52	352	2.9	3	352	3.0	3
53	354	3.0	4	352	3.0	3
54	355	3.1	4	354	3.1	4
55	357	3.2	4	355	3.2	4
56	358	3.4	4	357	3.4	4
57	359	3.5	4	358	3.6	4
58	361	3.8	4	360	3.8	4
59	363	4.0	4	362	4.1	4
60	365	4.4	4	364	4.6	4
61	367	4.9	4	367	5.2	4
62	370	5.6	4	370	6.3	4
63	374	6.7	4	375	8.0	4
64	380	9.4	4	380	10.0	4
65	380	10.0	4	380	10.0	4

Table M-2. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 4

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	400	10.0	1	400	10.0	1
1	400	10.0	1	400	10.0	1
2	400	10.0	1	400	10.0	1
3	400	10.0	1	400	10.0	1
4	400	10.0	1	400	10.0	1
5	400	10.0	1	400	10.0	1
6	400	10.0	1	401	8.7	1
7	401	8.5	1	406	7.0	1
8	406	6.8	1	409	6.0	1
9	409	5.9	1	411	5.4	1
10	411	5.3	1	414	4.9	1
11	413	4.9	1	415	4.6	1
12	415	4.5	1	417	4.3	1
13	417	4.3	1	419	4.1	1
14	418	4.1	1	420	3.9	1
15	420	3.9	1	421	3.7	1
16	421	3.8	1	422	3.6	1
17	422	3.7	1	423	3.4	1
18	424	3.6	1	425	3.3	1
19	425	3.5	1	426	3.3	1

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
20	426	3.4	1	427	3.2	1
21	427	3.4	1	427	3.1	1
22	428	3.3	1	428	3.0	1
23	429	3.3	1	429	3.0	1
24	430	3.2	1	430	2.9	1
25	430	3.2	1	430	2.9	1
26	432	3.1	2	432	2.9	2
27	432	3.1	2	433	2.8	2
28	433	3.1	2	433	2.8	2
29	434	3.0	2	434	2.8	2
30	435	3.0	2	435	2.8	2
31	436	3.0	2	436	2.8	2
32	437	3.0	2	436	2.8	2
33	438	3.0	2	437	2.8	2
34	438	3.0	2	438	2.7	2
35	439	3.0	2	439	2.7	2
36	440	2.9	3	439	2.7	2
37	441	2.9	3	440	2.7	3
38	442	2.9	3	441	2.8	3
39	443	2.9	3	442	2.8	3
40	443	2.9	3	442	2.8	3
41	444	3.0	3	443	2.8	3
42	445	3.0	3	444	2.8	3
43	446	3.0	3	445	2.8	3
44	447	3.0	3	446	2.8	3
45	448	3.0	3	446	2.9	3
46	449	3.1	3	447	2.9	3
47	450	3.1	3	448	2.9	3
48	451	3.1	3	449	3.0	3
49	452	3.2	3	450	3.0	3
50	453	3.2	3	451	3.0	3
51	454	3.3	3	452	3.1	3
52	455	3.4	4	453	3.1	3
53	456	3.5	4	454	3.2	3
54	457	3.6	4	455	3.3	4
55	459	3.7	4	456	3.4	4
56	460	3.9	4	457	3.5	4
57	462	4.1	4	459	3.6	4
58	463	4.3	4	460	3.8	4
59	465	4.6	4	462	4.0	4
60	468	5.1	4	464	4.3	4
61	471	5.7	4	466	4.8	4
62	474	6.6	4	469	5.6	4
63	479	8.4	4	473	7.0	4
64	480	10.0	4	480	10.0	4
65	480	10.0	4	480	10.0	4

Table M-3. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 5

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	500	10.0	1	500	10.0	1
1	500	10.0	1	500	10.0	1
2	500	10.0	1	500	10.0	1
3	500	10.0	1	500	10.0	1
4	500	10.0	1	500	10.0	1
5	500	10.0	1	500	10.0	1
6	500	10.0	1	500	10.0	1
7	508	9.4	1	503	10.0	1
8	513	6.8	1	509	7.7	1
9	517	5.6	1	513	6.4	1
10	519	4.9	1	516	5.7	1
11	521	4.5	1	518	5.2	1
12	523	4.1	1	520	4.8	1
13	524	3.9	1	522	4.5	1
14	526	3.7	1	524	4.3	1
15	527	3.6	1	525	4.1	1
16	528	3.5	1	527	3.9	1
17	529	3.3	1	528	3.8	1
18	530	3.2	1	529	3.6	1
19	531	3.2	1	530	3.5	1
20	532	3.1	1	531	3.4	1
21	532	3.0	1	532	3.3	1
22	534	2.9	2	533	3.3	2
23	535	2.9	2	534	3.2	2
24	535	2.8	2	535	3.1	2
25	536	2.8	2	536	3.1	2
26	537	2.7	2	537	3.0	2
27	538	2.7	2	538	2.9	2
28	538	2.6	2	539	2.9	2
29	539	2.6	2	539	2.8	2
30	539	2.6	2	540	2.8	3
31	540	2.5	3	541	2.8	3
32	541	2.5	3	542	2.7	3
33	542	2.5	3	542	2.7	3
34	542	2.5	3	543	2.7	3
35	543	2.5	3	544	2.6	3
36	544	2.5	3	545	2.6	3
37	544	2.4	3	545	2.6	3
38	545	2.4	3	546	2.6	3
39	545	2.4	3	547	2.5	3
40	546	2.4	3	547	2.5	3
41	547	2.4	3	548	2.5	3
42	547	2.4	3	549	2.5	3
43	548	2.4	3	549	2.5	3
44	549	2.4	3	550	2.5	3
45	549	2.4	3	551	2.5	3

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
46	550	2.5	3	552	2.5	3
47	551	2.5	3	552	2.5	3
48	551	2.5	3	553	2.5	3
49	552	2.5	3	553	2.6	3
50	553	2.6	3	555	2.6	4
51	553	2.6	3	555	2.6	4
52	554	2.6	4	556	2.7	4
53	555	2.7	4	557	2.7	4
54	556	2.8	4	558	2.8	4
55	557	2.8	4	559	2.8	4
56	558	2.9	4	560	2.9	4
57	559	3.0	4	561	3.1	4
58	560	3.2	4	562	3.2	4
59	561	3.3	4	564	3.4	4
60	563	3.6	4	565	3.7	4
61	564	3.9	4	567	4.0	4
62	566	4.4	4	569	4.5	4
63	569	5.2	4	572	5.3	4
64	573	6.5	4	576	6.7	4
65	579	9.8	4	580	9.9	4
66	580	10.0	4	580	10.0	4

Table M-4. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 6

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	600	10.0	1	600	10.0	1
1	600	10.0	1	600	10.0	1
2	600	10.0	1	600	10.0	1
3	600	10.0	1	600	10.0	1
4	600	10.0	1	600	10.0	1
5	600	10.0	1	600	10.0	1
6	600	10.0	1	600	10.0	1
7	606	10.0	1	607	9.9	1
8	612	7.8	1	613	7.6	1
9	616	6.4	1	617	6.4	1
10	619	5.5	1	620	5.6	1
11	622	4.9	1	622	5.0	1
12	623	4.5	1	624	4.6	1
13	625	4.2	1	626	4.3	1
14	627	4.0	1	628	4.0	1
15	628	3.8	1	629	3.8	1
16	629	3.7	1	630	3.6	1
17	630	3.5	1	631	3.5	1
18	631	3.4	1	632	3.4	1
19	632	3.3	1	634	3.3	2

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
20	633	3.2	2	635	3.2	2
21	634	3.1	2	635	3.1	2
22	635	3.1	2	636	3.0	2
23	636	3.0	2	637	3.0	2
24	636	2.9	2	638	2.9	2
25	637	2.9	2	639	2.9	2
26	638	2.8	2	639	2.8	2
27	639	2.8	2	640	2.8	3
28	639	2.7	2	641	2.8	3
29	640	2.7	3	642	2.7	3
30	641	2.7	3	643	2.7	3
31	641	2.6	3	643	2.7	3
32	642	2.6	3	644	2.7	3
33	643	2.6	3	645	2.6	3
34	643	2.6	3	645	2.6	3
35	644	2.5	3	646	2.6	3
36	645	2.5	3	647	2.6	3
37	645	2.5	3	647	2.5	3
38	646	2.5	3	648	2.5	3
39	647	2.5	3	649	2.5	3
40	647	2.5	3	649	2.5	3
41	648	2.5	3	650	2.5	3
42	649	2.5	3	651	2.5	3
43	649	2.5	3	651	2.5	3
44	650	2.5	3	652	2.5	3
45	651	2.6	3	652	2.5	3
46	651	2.6	3	652	2.5	3
47	652	2.6	3	654	2.5	4
48	652	2.6	3	655	2.5	4
49	654	2.7	4	655	2.5	4
50	654	2.7	4	656	2.6	4
51	655	2.8	4	657	2.6	4
52	656	2.8	4	658	2.7	4
53	657	2.9	4	659	2.7	4
54	658	3.0	4	660	2.8	4
55	659	3.1	4	661	2.9	4
56	660	3.2	4	662	3.0	4
57	661	3.3	4	663	3.1	4
58	662	3.5	4	664	3.3	4
59	664	3.7	4	665	3.5	4
60	665	3.9	4	667	3.8	4
61	667	4.2	4	669	4.1	4
62	669	4.7	4	671	4.6	4
63	672	5.5	4	673	5.3	4
64	676	6.8	4	677	6.6	4
65	680	10.0	4	680	9.5	4
66	680	10.0	4	680	10.0	4

Table M-5. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 7

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	700	10.0	1	700	10.0	700
1	700	10.0	1	700	10.0	700
2	700	10.0	1	700	10.0	700
3	700	10.0	1	700	10.0	700
4	700	10.0	1	700	10.0	700
5	700	10.0	1	700	10.0	700
6	702	10.0	1	700	10.0	700
7	713	9.4	1	711	8.2	711
8	718	6.8	1	716	6.3	716
9	721	5.5	1	719	5.3	719
10	724	4.8	1	722	4.8	722
11	725	4.3	1	724	4.4	724
12	727	3.9	1	726	4.1	726
13	728	3.6	1	727	3.8	727
14	730	3.4	1	729	3.6	729
15	731	3.2	1	730	3.4	730
16	732	3.1	1	731	3.3	731
17	733	3.0	1	732	3.2	732
18	734	2.9	2	733	3.0	733
19	735	2.8	2	734	2.9	734
20	735	2.7	2	735	2.8	735
21	736	2.6	2	736	2.8	736
22	737	2.6	2	737	2.7	737
23	738	2.5	2	738	2.6	738
24	738	2.5	2	739	2.6	739
25	739	2.4	2	739	2.5	739
26	739	2.4	2	739	2.5	739
27	740	2.4	3	741	2.4	741
28	741	2.3	3	741	2.4	741
29	742	2.3	3	742	2.4	742
30	742	2.3	3	743	2.3	743
31	743	2.3	3	743	2.3	743
32	743	2.2	3	744	2.3	744
33	744	2.2	3	744	2.3	744
34	745	2.2	3	745	2.3	745
35	745	2.2	3	746	2.3	746
36	746	2.2	3	746	2.3	746
37	746	2.2	3	747	2.3	747
38	747	2.2	3	748	2.3	748
39	748	2.2	3	748	2.3	748
40	748	2.2	3	749	2.3	749
41	749	2.3	3	749	2.3	749
42	749	2.3	3	750	2.3	750
43	750	2.3	3	751	2.4	751
44	751	2.3	3	751	2.4	751
45	751	2.4	3	752	2.4	752

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
46	752	2.4	4	753	2.4	753
47	753	2.4	4	753	2.4	4
48	753	2.5	4	754	2.5	4
49	754	2.5	4	755	2.5	4
50	755	2.6	4	756	2.5	4
51	756	2.7	4	757	2.6	4
52	757	2.7	4	757	2.6	4
53	757	2.8	4	758	2.7	4
54	758	2.9	4	759	2.7	4
55	759	3.0	4	760	2.8	4
56	760	3.2	4	761	2.9	4
57	762	3.3	4	762	3.0	4
58	763	3.5	4	763	3.2	4
59	764	3.7	4	765	3.4	4
60	766	4.0	4	766	3.6	4
61	768	4.4	4	768	4.0	4
62	770	5.0	4	770	4.5	4
63	773	5.9	4	773	5.2	4
64	778	7.5	4	777	6.4	4
65	780	9.3	4	780	7.7	4
66	780	9.3	4	780	7.7	4

Table M-6. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 8

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	800	10.0	1	800	10.0	1
1	800	10.0	1	800	10.0	1
2	800	10.0	1	800	10.0	1
3	800	10.0	1	800	10.0	1
4	800	10.0	1	800	10.0	1
5	800	10.0	1	800	10.0	1
6	800	10.0	1	803	10.0	1
7	811	9.7	1	814	8.3	1
8	817	6.7	1	819	6.0	1
9	820	5.3	1	821	4.9	1
10	822	4.5	1	824	4.3	1
11	824	3.9	1	825	3.8	1
12	826	3.6	1	827	3.5	1
13	827	3.3	1	828	3.3	1
14	828	3.1	1	829	3.1	1
15	829	2.9	1	830	2.9	1
16	830	2.8	1	831	2.8	1
17	831	2.7	1	832	2.7	1
18	832	2.6	1	833	2.6	1
19	833	2.5	1	834	2.5	2

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
20	833	2.4	1	835	2.5	2
21	834	2.4	2	835	2.4	2
22	835	2.3	2	836	2.4	2
23	836	2.3	2	837	2.3	2
24	836	2.2	2	837	2.3	2
25	837	2.2	2	838	2.3	2
26	838	2.2	2	839	2.2	2
27	838	2.1	2	839	2.2	2
28	839	2.1	2	839	2.2	2
29	839	2.1	2	841	2.2	3
30	839	2.1	2	841	2.2	3
31	840	2.1	3	842	2.1	3
32	841	2.1	3	842	2.1	3
33	842	2.0	3	843	2.1	3
34	842	2.0	3	843	2.1	3
35	843	2.0	3	844	2.1	3
36	843	2.0	3	845	2.1	3
37	844	2.0	3	845	2.1	3
38	844	2.0	3	846	2.0	3
39	845	2.0	3	846	2.0	3
40	845	2.0	3	847	2.0	3
41	846	2.0	3	847	2.0	3
42	847	2.0	3	848	2.0	3
43	847	2.0	3	849	2.0	3
44	848	2.0	3	849	2.0	3
45	848	2.0	3	850	2.0	3
46	849	2.0	3	850	2.0	3
47	849	2.1	3	851	2.0	3
48	850	2.1	3	851	2.0	3
49	851	2.1	3	852	2.0	4
50	851	2.1	3	853	2.0	4
51	852	2.2	4	853	2.1	4
52	853	2.2	4	854	2.1	4
53	853	2.2	4	855	2.1	4
54	854	2.3	4	855	2.2	4
55	855	2.3	4	856	2.2	4
56	856	2.4	4	857	2.3	4
57	857	2.5	4	858	2.4	4
58	858	2.6	4	859	2.5	4
59	859	2.7	4	860	2.7	4
60	860	2.9	4	861	2.9	4
61	861	3.2	4	863	3.1	4
62	863	3.5	4	864	3.4	4
63	865	4.1	4	867	3.9	4
64	868	5.1	4	870	4.5	4
65	874	7.7	4	874	6.1	4
66	880	10.0	4	880	10.0	4

Table M-7. 2012–13 NECAP: Raw to Scaled Score Correspondence
Mathematics Grade 11

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	1100	10.0	1	1100	10.0	1
1	1100	10.0	1	1100	10.0	1
2	1100	10.0	1	1100	10.0	1
3	1100	10.0	1	1100	10.0	1
4	1100	10.0	1	1108	10.0	1
5	1114	8.0	1	1116	6.3	1
6	1119	5.3	1	1120	4.6	1
7	1122	4.2	1	1122	3.8	1
8	1124	3.6	1	1124	3.3	1
9	1125	3.2	1	1126	3.0	1
10	1127	2.9	1	1127	2.7	1
11	1128	2.7	1	1128	2.6	1
12	1129	2.5	1	1129	2.4	1
13	1130	2.4	1	1130	2.3	1
14	1131	2.2	1	1131	2.2	1
15	1132	2.2	1	1132	2.2	1
16	1132	2.1	1	1133	2.1	1
17	1133	2.0	1	1133	2.0	1
18	1133	1.9	1	1134	2.0	2
19	1134	1.9	2	1135	1.9	2
20	1135	1.8	2	1135	1.9	2
21	1135	1.8	2	1136	1.9	2
22	1136	1.7	2	1137	1.8	2
23	1136	1.7	2	1137	1.8	2
24	1137	1.7	2	1138	1.8	2
25	1137	1.7	2	1138	1.7	2
26	1138	1.6	2	1139	1.7	2
27	1138	1.6	2	1139	1.7	2
28	1139	1.6	2	1139	1.7	2
29	1139	1.6	2	1140	1.7	3
30	1139	1.6	2	1141	1.6	3
31	1140	1.6	3	1141	1.6	3
32	1141	1.6	3	1142	1.6	3
33	1141	1.6	3	1142	1.6	3
34	1141	1.6	3	1143	1.6	3
35	1142	1.6	3	1143	1.6	3
36	1142	1.6	3	1144	1.6	3
37	1143	1.6	3	1144	1.6	3
38	1143	1.6	3	1145	1.6	3
39	1144	1.6	3	1145	1.6	3
40	1144	1.6	3	1145	1.5	3
41	1145	1.6	3	1146	1.5	3
42	1145	1.6	3	1146	1.6	3
43	1146	1.6	3	1147	1.6	3
44	1146	1.6	3	1147	1.6	3
45	1146	1.6	3	1148	1.6	3

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
46	1147	1.6	3	1148	1.6	3
47	1147	1.6	3	1149	1.6	3
48	1148	1.6	3	1149	1.7	3
49	1149	1.7	3	1150	1.7	3
50	1149	1.7	3	1151	1.7	3
51	1150	1.7	3	1151	1.8	3
52	1150	1.8	3	1151	1.9	3
53	1151	1.8	3	1151	1.9	3
54	1151	1.9	3	1153	2.0	4
55	1151	2.0	3	1154	2.2	4
56	1153	2.1	4	1155	2.3	4
57	1154	2.2	4	1156	2.5	4
58	1155	2.4	4	1157	2.7	4
59	1156	2.6	4	1158	3.1	4
60	1158	2.9	4	1160	3.5	4
61	1159	3.2	4	1162	4.2	4
62	1161	3.9	4	1165	5.3	4
63	1165	5.4	4	1169	7.3	4
64	1180	8.1	4	1180	7.3	4

Table M-8. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 3

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	300	10.0	1	300	10.0	1
1	300	10.0	1	300	10.0	1
2	300	10.0	1	300	10.0	1
3	300	10.0	1	300	10.0	1
4	300	10.0	1	300	10.0	1
5	300	10.0	1	300	10.0	1
6	300	10.0	1	300	10.0	1
7	300	9.7	1	300	10.0	1
8	304	8.6	1	301	10.0	1
9	308	7.7	1	306	9.1	1
10	311	7.1	1	311	7.7	1
11	314	6.5	1	314	6.6	1
12	316	6.0	1	317	5.9	1
13	319	5.6	1	319	5.3	1
14	321	5.2	1	321	4.8	1
15	322	4.9	1	323	4.4	1
16	324	4.7	1	325	4.0	1
17	326	4.5	1	326	3.7	1
18	327	4.3	1	327	3.5	1
19	328	4.1	1	329	3.3	1
20	330	4.0	1	330	3.2	1
21	331	3.9	2	331	3.1	2

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
22	332	3.8	2	332	3.0	2
23	334	3.7	2	333	3.0	2
24	335	3.7	2	334	2.9	2
25	336	3.6	2	335	2.9	2
26	337	3.6	2	336	2.9	2
27	338	3.6	2	337	2.9	2
28	339	3.5	2	338	2.9	2
29	341	3.5	3	339	3.0	2
30	342	3.5	3	340	3.0	3
31	343	3.5	3	341	3.1	3
32	344	3.5	3	342	3.1	3
33	346	3.5	3	343	3.2	3
34	347	3.5	3	345	3.3	3
35	348	3.6	3	346	3.4	3
36	349	3.6	3	347	3.5	3
37	351	3.7	3	349	3.7	3
38	352	3.8	3	350	3.8	3
39	354	3.9	3	352	4.0	3
40	356	4.1	3	354	4.2	3
41	356	4.3	3	356	4.5	3
42	359	4.6	4	358	4.8	4
43	362	4.9	4	360	5.1	4
44	364	5.3	4	363	5.5	4
45	367	5.7	4	366	5.8	4
46	370	6.1	4	369	6.2	4
47	374	6.6	4	373	6.4	4
48	378	7.1	4	377	6.7	4
49	380	7.9	4	380	7.2	4
50	380	9.2	4	380	8.4	4
51	380	9.7	4	380	9.3	4
52	380	9.7	4	380	9.3	4

Table M-9. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 4

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	400	10.0	1	400	10.0	1
1	400	10.0	1	400	10.0	1
2	400	10.0	1	400	10.0	1
3	400	10.0	1	400	10.0	1
4	400	10.0	1	400	10.0	1
5	400	10.0	1	400	10.0	1
6	400	9.7	1	400	9.8	1
7	400	8.8	1	404	8.4	1
8	404	7.9	1	408	7.4	1
9	408	7.1	1	411	6.7	1

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
10	411	6.5	1	414	6.2	1
11	413	6.0	1	416	5.7	1
12	416	5.6	1	418	5.3	1
13	418	5.2	1	420	5.0	1
14	420	4.9	1	422	4.7	1
15	422	4.7	1	423	4.4	1
16	423	4.5	1	425	4.2	1
17	425	4.3	1	426	4.0	1
18	426	4.1	1	427	3.8	1
19	427	3.9	1	429	3.7	1
20	429	3.8	1	430	3.5	1
21	430	3.6	1	431	3.4	2
22	431	3.5	2	432	3.4	2
23	432	3.4	2	433	3.3	2
24	433	3.4	2	434	3.2	2
25	434	3.3	2	435	3.2	2
26	436	3.2	2	436	3.1	2
27	437	3.2	2	437	3.1	2
28	438	3.2	2	438	3.1	2
29	439	3.1	2	439	3.1	2
30	439	3.1	2	440	3.1	3
31	441	3.1	3	441	3.1	3
32	442	3.1	3	442	3.1	3
33	443	3.1	3	443	3.2	3
34	444	3.1	3	444	3.2	3
35	445	3.2	3	446	3.3	3
36	446	3.3	3	447	3.4	3
37	448	3.3	3	448	3.5	3
38	449	3.5	3	449	3.6	3
39	450	3.6	3	451	3.8	3
40	452	3.8	3	452	4.0	3
41	453	4.0	3	454	4.2	3
42	455	4.2	3	456	4.5	4
43	457	4.5	4	458	4.7	4
44	459	4.8	4	460	5.1	4
45	462	5.1	4	463	5.4	4
46	465	5.5	4	466	5.9	4
47	468	6.1	4	469	6.3	4
48	472	6.9	4	473	6.7	4
49	478	8.1	4	477	7.3	4
50	480	9.9	4	480	8.3	4
51	480	10.0	4	480	9.5	4
52	480	10.0	4	480	9.5	4

Table M-10. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 5

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	500	10.0	1	500	10.0	1
1	500	10.0	1	500	10.0	1
2	500	10.0	1	500	10.0	1
3	501	10.0	1	500	10.0	1
4	508	7.4	1	500	10.0	1
5	512	6.1	1	504	10.0	1
6	515	5.3	1	510	7.4	1
7	517	4.8	1	514	6.0	1
8	519	4.4	1	516	5.2	1
9	521	4.1	1	519	4.6	1
10	523	3.9	1	520	4.2	1
11	524	3.7	1	522	3.9	1
12	526	3.5	1	524	3.7	1
13	527	3.4	1	525	3.5	1
14	528	3.3	1	526	3.3	1
15	529	3.2	1	527	3.2	1
16	530	3.1	2	528	3.1	1
17	531	3.1	2	529	3.1	1
18	532	3.1	2	531	3.0	2
19	534	3.0	2	532	3.0	2
20	535	3.0	2	533	2.9	2
21	536	3.0	2	534	2.9	2
22	537	3.0	2	535	2.9	2
23	538	3.0	2	536	2.9	2
24	539	3.1	2	537	3.0	2
25	539	3.1	2	538	3.0	2
26	541	3.1	3	539	3.0	2
27	542	3.2	3	539	3.1	2
28	543	3.2	3	541	3.1	3
29	544	3.3	3	542	3.2	3
30	546	3.3	3	543	3.2	3
31	547	3.4	3	544	3.3	3
32	548	3.5	3	546	3.4	3
33	550	3.6	3	547	3.5	3
34	551	3.7	3	548	3.6	3
35	553	3.8	3	550	3.8	3
36	554	3.9	3	551	3.9	3
37	555	4.0	3	553	4.1	3
38	558	4.2	4	555	4.2	3
39	559	4.3	4	557	4.4	4
40	561	4.4	4	559	4.5	4
41	563	4.5	4	561	4.7	4
42	566	4.6	4	563	4.8	4
43	568	4.7	4	565	5.0	4
44	570	4.8	4	568	5.1	4
45	573	4.9	4	570	5.3	4

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
46	575	5.0	4	573	5.4	4
47	578	5.1	4	576	5.5	4
48	580	5.4	4	579	5.7	4
49	580	5.8	4	580	6.1	4
50	580	6.8	4	580	7.0	4
51	580	6.9	4	580	7.8	4
52	580	6.9	4	580	7.8	4

Table M-11. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 6

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	600	10.0	1	600	10.0	1
1	600	10.0	1	600	10.0	1
2	600	10.0	1	600	10.0	1
3	600	9.8	1	600	9.7	1
4	601	8.2	1	600	8.3	1
5	605	6.6	1	604	6.8	1
6	609	5.7	1	607	5.8	1
7	611	5.1	1	610	5.2	1
8	614	4.6	1	612	4.7	1
9	615	4.3	1	614	4.4	1
10	617	4.1	1	616	4.1	1
11	619	3.9	1	618	3.9	1
12	620	3.8	1	619	3.8	1
13	622	3.7	1	621	3.7	1
14	623	3.6	1	622	3.6	1
15	624	3.5	1	623	3.5	1
16	626	3.5	1	624	3.5	1
17	627	3.4	1	626	3.4	1
18	628	3.4	1	627	3.4	1
19	629	3.4	2	628	3.4	1
20	630	3.4	2	629	3.4	2
21	632	3.4	2	630	3.4	2
22	633	3.4	2	632	3.4	2
23	634	3.4	2	633	3.4	2
24	635	3.4	2	634	3.4	2
25	636	3.5	2	635	3.5	2
26	638	3.5	2	637	3.5	2
27	639	3.6	2	638	3.5	2
28	640	3.6	3	639	3.6	2
29	642	3.7	3	640	3.6	3
30	643	3.8	3	642	3.7	3
31	644	3.8	3	643	3.8	3
32	646	3.9	3	645	3.9	3
33	648	4.0	3	646	4.0	3

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
34	649	4.1	3	648	4.1	3
35	651	4.2	3	650	4.2	3
36	653	4.3	3	651	4.4	3
37	655	4.4	3	653	4.5	3
38	657	4.5	3	655	4.5	3
39	658	4.5	3	657	4.6	3
40	661	4.6	4	659	4.6	4
41	663	4.7	4	661	4.6	4
42	665	4.8	4	664	4.7	4
43	668	4.9	4	666	4.8	4
44	670	4.9	4	668	4.8	4
45	673	5.0	4	670	4.9	4
46	675	5.1	4	673	5.0	4
47	678	5.1	4	676	5.0	4
48	680	5.3	4	678	5.2	4
49	680	5.7	4	680	5.7	4
50	680	6.6	4	680	6.6	4
51	680	7.4	4	680	8.8	4
52	680	7.4	4	680	8.8	4

Table M-12. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 7

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	700	10.0	1	700	10.0	1
1	700	10.0	1	700	10.0	1
2	700	10.0	1	700	10.0	1
3	700	9.4	1	700	10.0	1
4	700	8.3	1	700	9.2	1
5	701	7.3	1	705	7.1	1
6	705	6.4	1	708	6.0	1
7	708	5.8	1	711	5.2	1
8	711	5.3	1	713	4.8	1
9	713	5.0	1	715	4.4	1
10	715	4.8	1	717	4.2	1
11	717	4.6	1	719	4.0	1
12	718	4.4	1	720	3.8	1
13	720	4.2	1	722	3.7	1
14	722	4.1	1	723	3.7	1
15	723	4.0	1	724	3.6	1
16	724	4.0	1	726	3.6	1
17	726	3.9	1	727	3.5	1
18	727	3.8	1	728	3.5	1
19	728	3.8	1	729	3.5	2
20	730	3.8	2	730	3.5	2
21	731	3.8	2	732	3.5	2

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
22	732	3.7	2	733	3.5	2
23	733	3.7	2	734	3.5	2
24	735	3.7	2	735	3.5	2
25	736	3.7	2	736	3.5	2
26	737	3.8	2	738	3.5	2
27	738	3.8	2	739	3.6	2
28	739	3.8	2	740	3.6	3
29	741	3.9	3	741	3.6	3
30	742	3.9	3	743	3.6	3
31	743	4.0	3	744	3.7	3
32	745	4.0	3	745	3.7	3
33	746	4.1	3	747	3.8	3
34	748	4.1	3	748	3.8	3
35	749	4.2	3	750	3.9	3
36	751	4.3	3	751	4.0	3
37	752	4.3	3	753	4.0	3
38	754	4.4	3	754	4.1	3
39	756	4.4	3	756	4.1	3
40	757	4.5	3	758	4.2	3
41	759	4.6	3	759	4.2	3
42	761	4.6	4	761	4.3	4
43	763	4.7	4	763	4.4	4
44	765	4.8	4	765	4.4	4
45	767	4.9	4	767	4.5	4
46	770	5.0	4	770	4.6	4
47	772	5.2	4	772	4.7	4
48	775	5.5	4	775	5.0	4
49	779	6.1	4	778	5.5	4
50	780	7.3	4	780	6.6	4
51	780	10.0	4	780	10.0	4
52	780	10.0	4	780	10.0	4

Table M-13. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 8

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	800	10.0	1	800	10.0	1
1	800	10.0	1	800	10.0	1
2	800	10.0	1	800	10.0	1
3	800	10.0	1	800	10.0	1
4	800	10.0	1	800	10.0	1
5	803	9.1	1	804	8.5	1
6	808	7.3	1	808	6.9	1
7	811	6.3	1	812	5.9	1
8	814	5.6	1	814	5.3	1
9	817	5.1	1	817	4.8	1

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
10	819	4.7	1	819	4.5	1
11	821	4.5	1	821	4.3	1
12	822	4.3	1	822	4.1	1
13	824	4.1	1	824	3.9	1
14	826	4.0	1	825	3.8	1
15	827	3.9	1	827	3.7	1
16	828	3.8	2	827	3.7	1
17	830	3.8	2	829	3.6	2
18	831	3.8	2	830	3.6	2
19	832	3.7	2	832	3.6	2
20	834	3.7	2	833	3.5	2
21	835	3.7	2	834	3.5	2
22	836	3.7	2	835	3.5	2
23	838	3.7	2	837	3.5	2
24	839	3.7	2	838	3.5	2
25	840	3.7	3	839	3.5	2
26	841	3.7	3	840	3.5	3
27	843	3.7	3	841	3.5	3
28	844	3.8	3	842	3.6	3
29	845	3.8	3	844	3.6	3
30	847	3.8	3	845	3.6	3
31	848	3.8	3	846	3.7	3
32	849	3.9	3	848	3.7	3
33	851	3.9	3	849	3.8	3
34	852	4.0	3	850	3.8	3
35	854	4.0	3	852	3.8	3
36	855	4.1	3	853	3.9	3
37	857	4.1	3	855	3.9	3
38	858	4.2	3	856	3.9	3
39	860	4.2	4	858	4.0	3
40	862	4.3	4	860	4.1	4
41	864	4.3	4	861	4.1	4
42	866	4.4	4	863	4.2	4
43	868	4.5	4	865	4.2	4
44	870	4.5	4	867	4.3	4
45	872	4.6	4	869	4.4	4
46	875	4.7	4	871	4.5	4
47	877	4.9	4	874	4.6	4
48	880	5.1	4	876	4.9	4
49	880	5.7	4	879	5.4	4
50	880	6.8	4	880	6.3	4
51	880	8.5	4	880	9.2	4
52	880	8.5	4	880	10.0	4

Table M-14. 2012–13 NECAP: Raw to Scaled Score Correspondence
Reading Grade 11

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
0	1100	10.0	1	1100	10.0	1
1	1100	10.0	1	1100	10.0	1
2	1100	10.0	1	1100	10.0	1
3	1100	10.0	1	1100	10.0	1
4	1100	10.0	1	1100	10.0	1
5	1106	7.7	1	1107	7.4	1
6	1110	6.2	1	1111	5.8	1
7	1113	5.3	1	1114	4.9	1
8	1115	4.7	1	1116	4.3	1
9	1117	4.3	1	1118	4.0	1
10	1119	4.0	1	1120	3.7	1
11	1121	3.8	1	1121	3.5	1
12	1122	3.6	1	1123	3.4	1
13	1124	3.5	1	1124	3.3	1
14	1125	3.4	1	1125	3.2	1
15	1126	3.3	1	1126	3.2	1
16	1127	3.3	1	1128	3.1	1
17	1128	3.2	1	1129	3.1	1
18	1129	3.2	1	1129	3.1	1
19	1130	3.1	2	1131	3.1	2
20	1131	3.1	2	1132	3.1	2
21	1132	3.1	2	1133	3.0	2
22	1133	3.1	2	1134	3.0	2
23	1134	3.0	2	1135	3.0	2
24	1135	3.0	2	1136	3.1	2
25	1136	3.0	2	1137	3.1	2
26	1137	3.0	2	1138	3.1	2
27	1138	3.1	2	1139	3.1	2
28	1139	3.1	2	1140	3.1	3
29	1141	3.1	3	1141	3.2	3
30	1142	3.2	3	1142	3.2	3
31	1143	3.2	3	1143	3.2	3
32	1144	3.3	3	1145	3.2	3
33	1145	3.3	3	1146	3.3	3
34	1147	3.4	3	1147	3.3	3
35	1148	3.4	3	1148	3.3	3
36	1149	3.5	3	1150	3.3	3
37	1151	3.5	3	1151	3.3	3
38	1152	3.5	3	1152	3.4	3
39	1153	3.6	3	1153	3.4	3
40	1155	3.6	4	1155	3.5	4
41	1157	3.7	4	1156	3.5	4
42	1158	3.8	4	1158	3.6	4
43	1160	3.8	4	1160	3.6	4
44	1162	3.9	4	1161	3.6	4
45	1164	4.0	4	1163	3.7	4

continued

Raw Score	2012–13			2011–12		
	Scaled Score	Standard Error	Performance Level	Scaled Score	Standard Error	Performance Level
46	1166	4.0	4	1165	3.8	4
47	1168	4.1	4	1167	3.9	4
48	1171	4.3	4	1169	4.1	4
49	1174	4.7	4	1172	4.6	4
50	1177	5.5	4	1175	5.5	4
51	1180	8.0	4	1180	8.2	4
52	1180	10.0	4	1180	10.0	4

APPENDIX N—SCALED SCORE DISTRIBUTIONS

Figure N-1. 2012–13 NECAP: Cumulative Score Distribution
Mathematics Grade 3

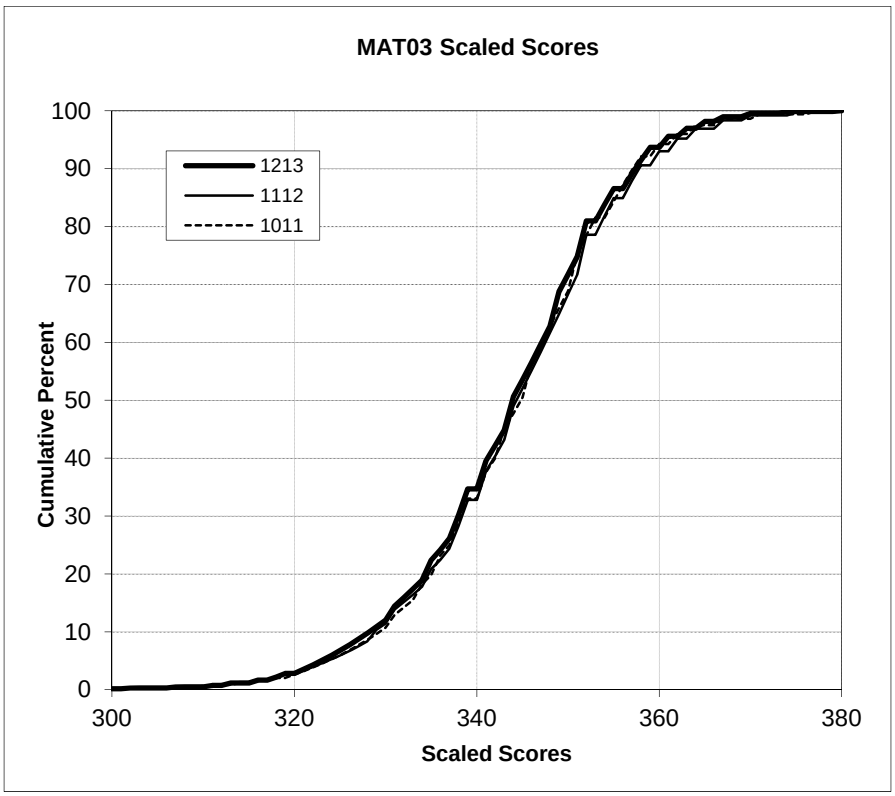


Figure N-2. 2012–13 NECAP: Cumulative Score Distribution
Mathematics Grade 4

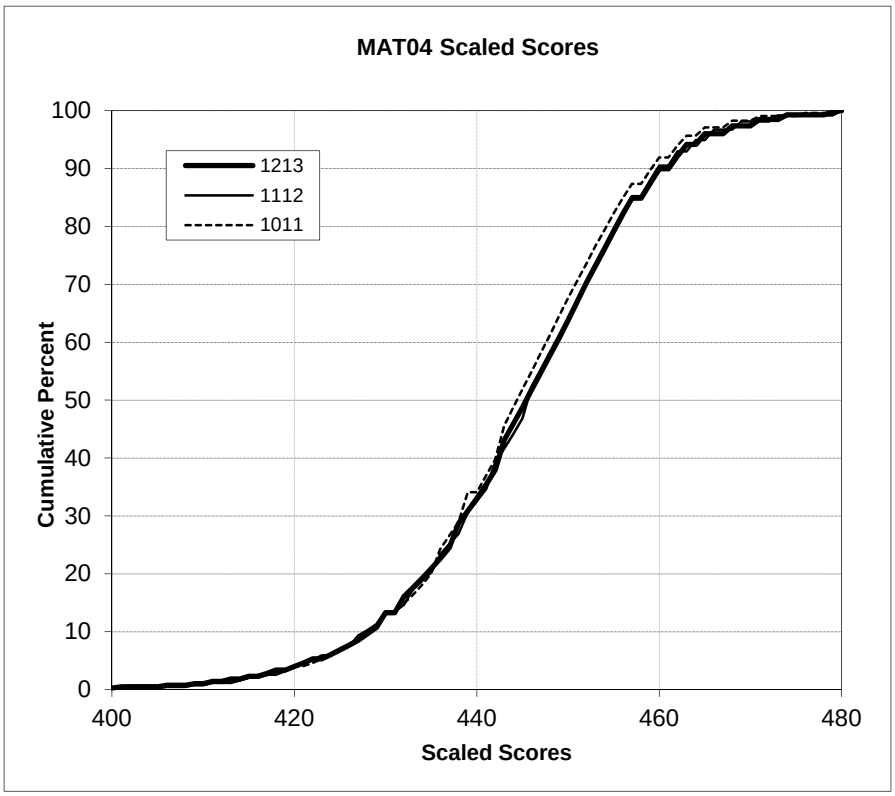


Figure N-3. 2012–13 NECAP: Cumulative Score Distribution Mathematics Grade 5

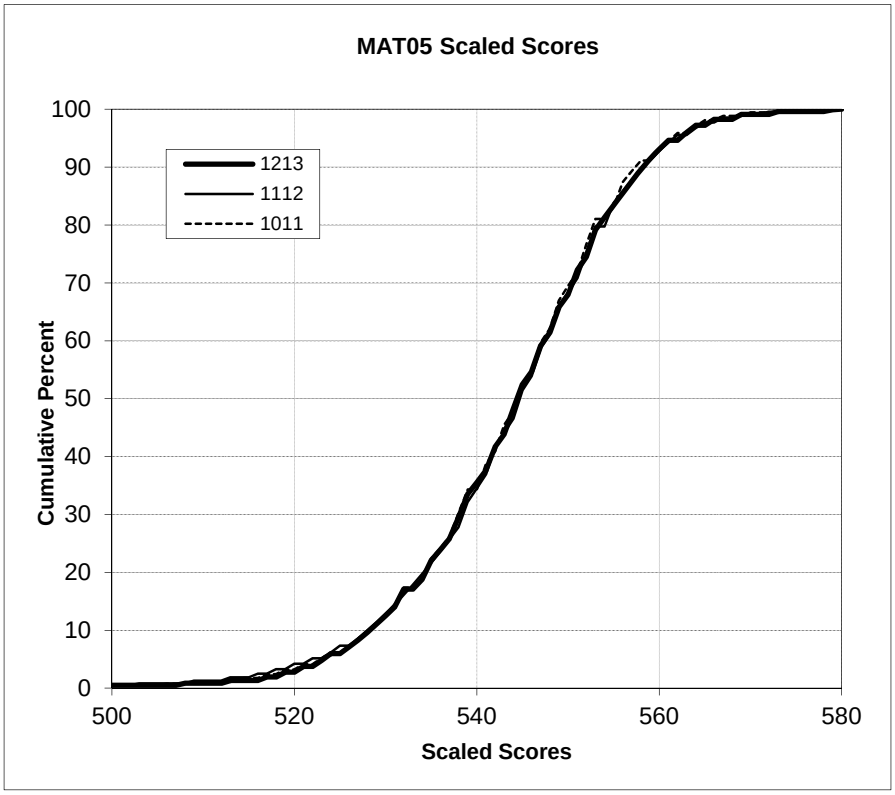


Figure N-4. 2012–13 NECAP: Cumulative Score Distribution Mathematics Grade 6

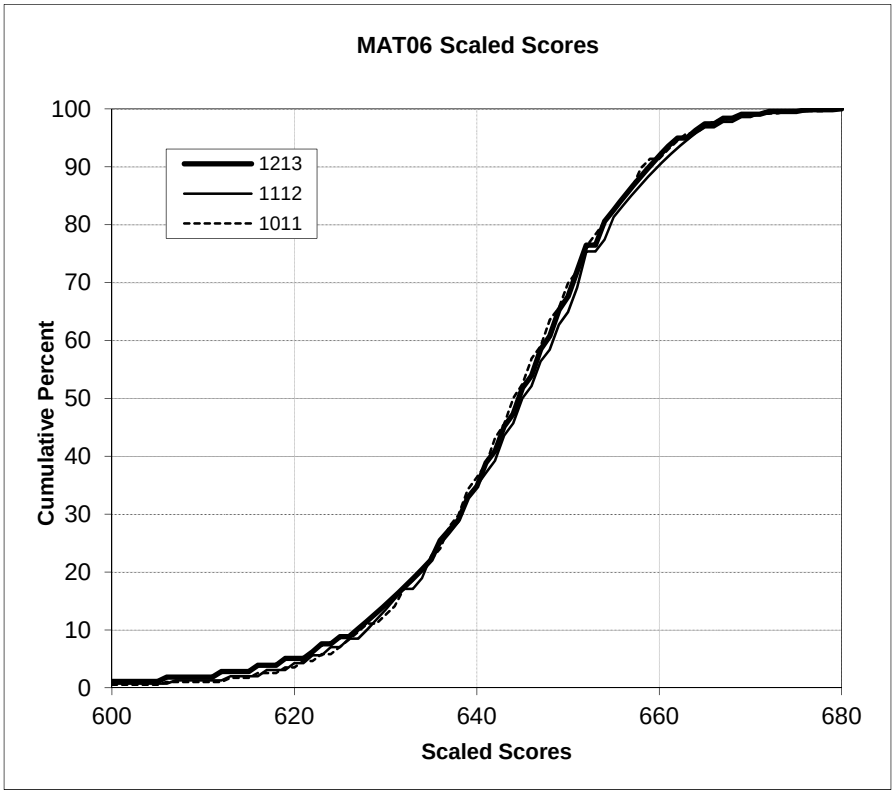


Figure N-5. 2012–13 NECAP: Cumulative Score Distribution Mathematics Grade 7

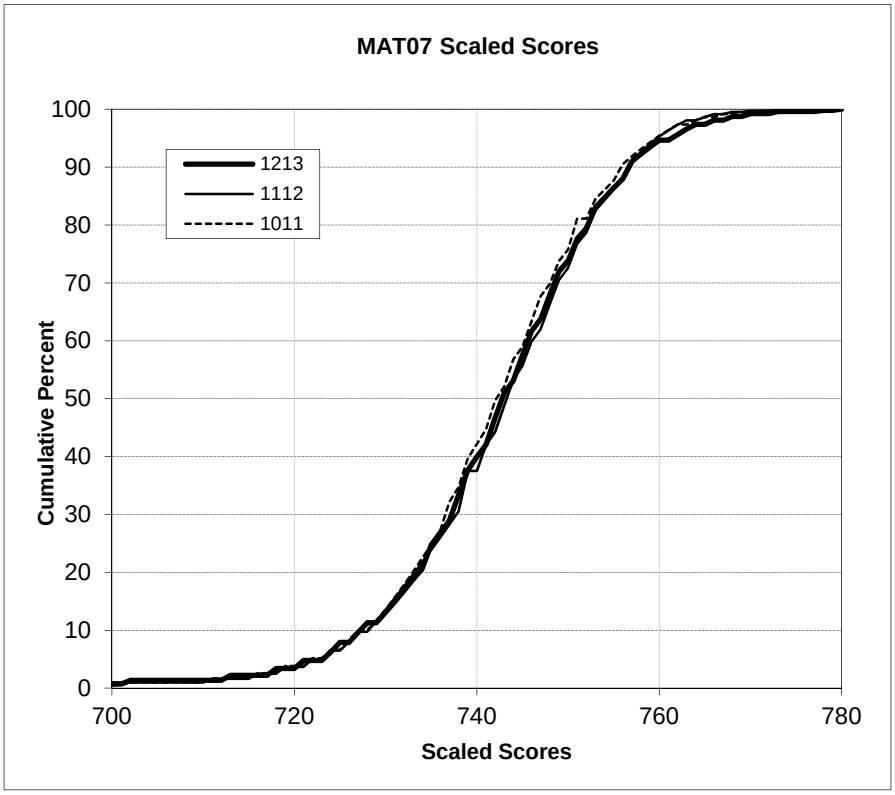


Figure N-6. 2012–13 NECAP: Cumulative Score Distribution Mathematics Grade 8

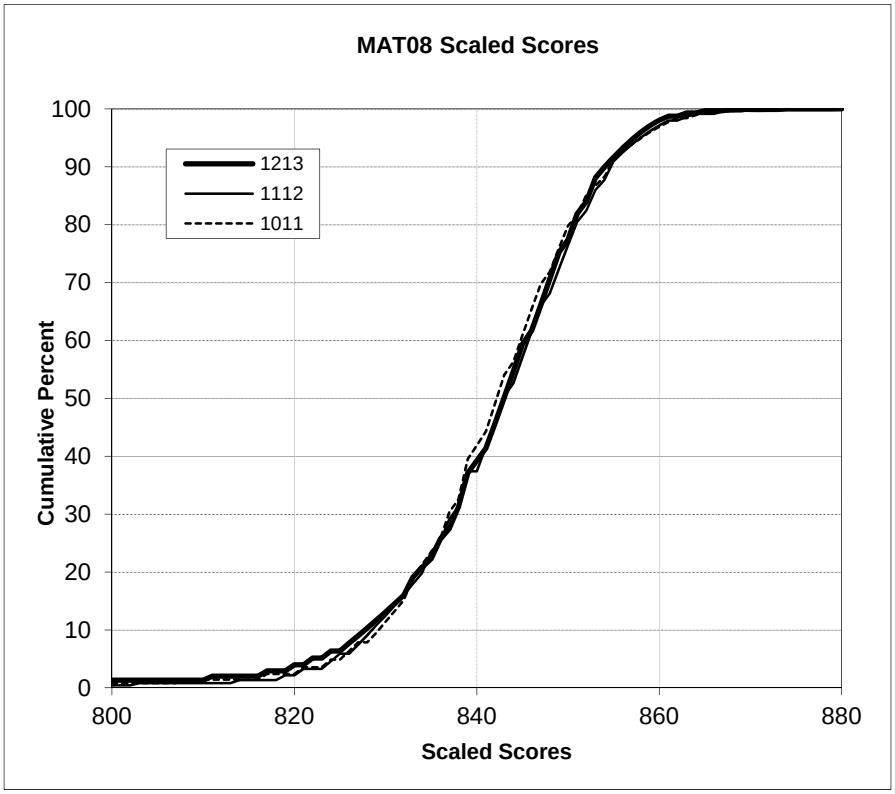


Figure N-7. 2012–13 NECAP: Cumulative Score Distribution
Mathematics Grade 11

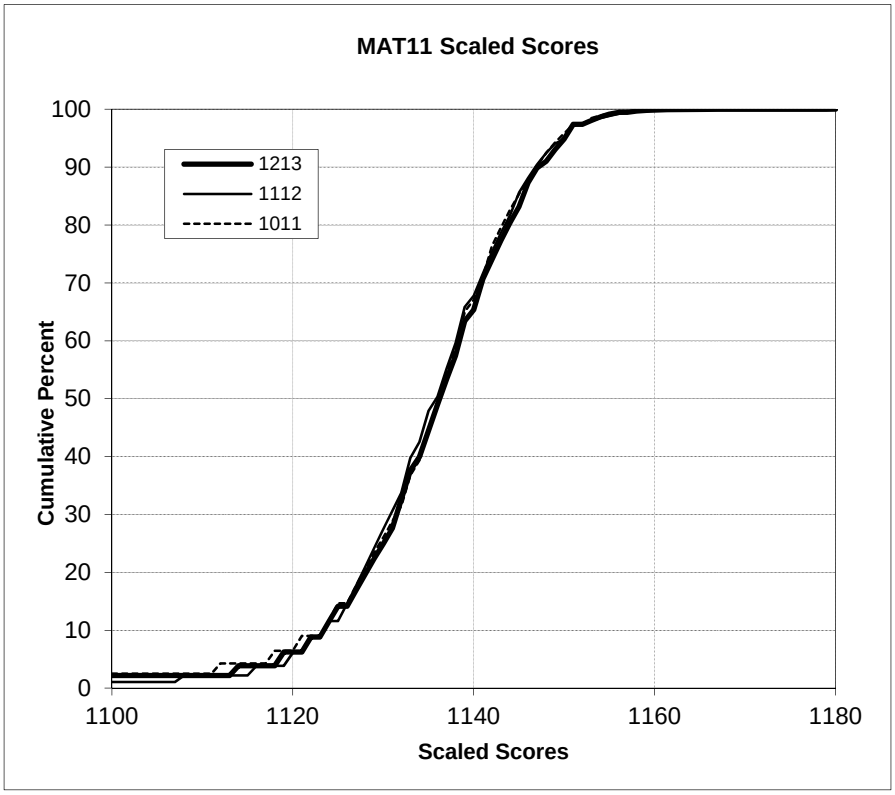


Figure N-8. 2012–13 NECAP: Cumulative Score Distribution
Reading Grade 3

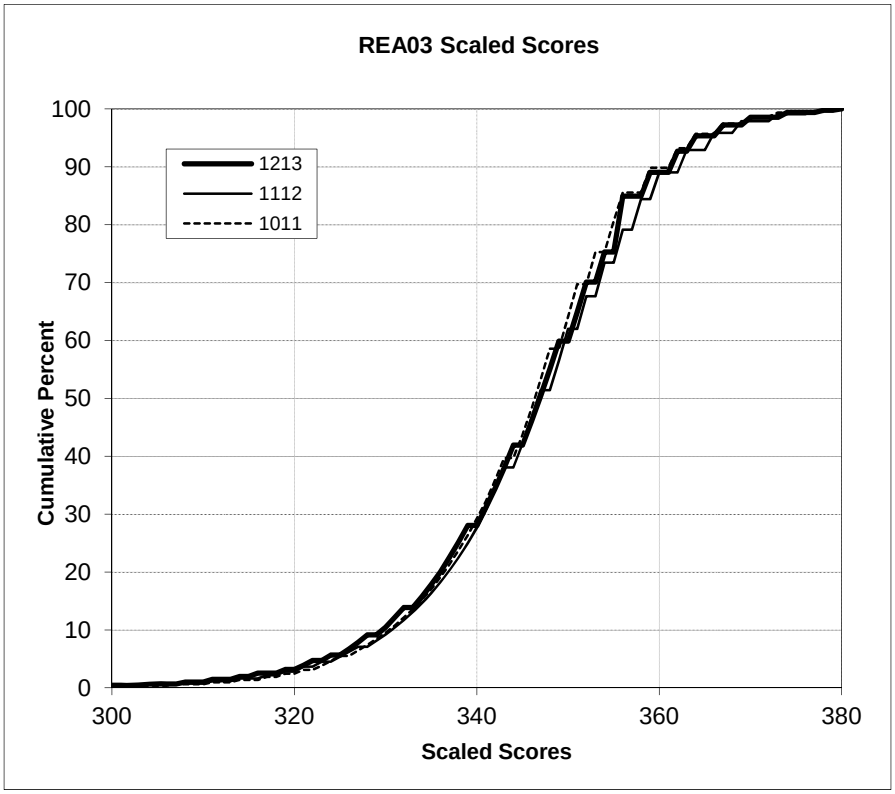


Figure N-9. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 4

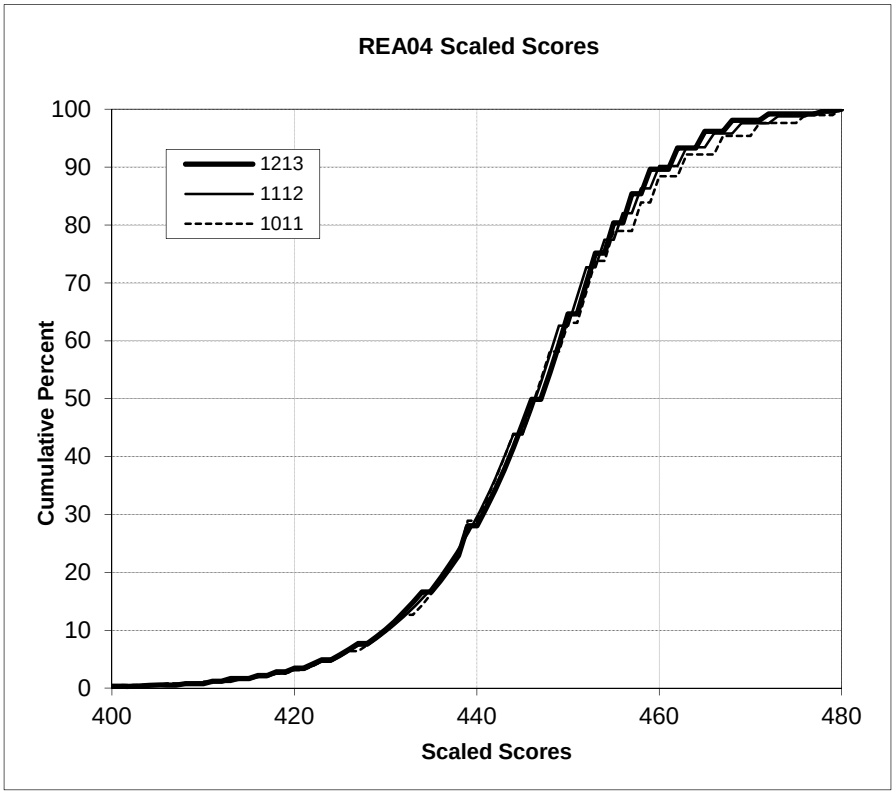


Figure N-10. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 5

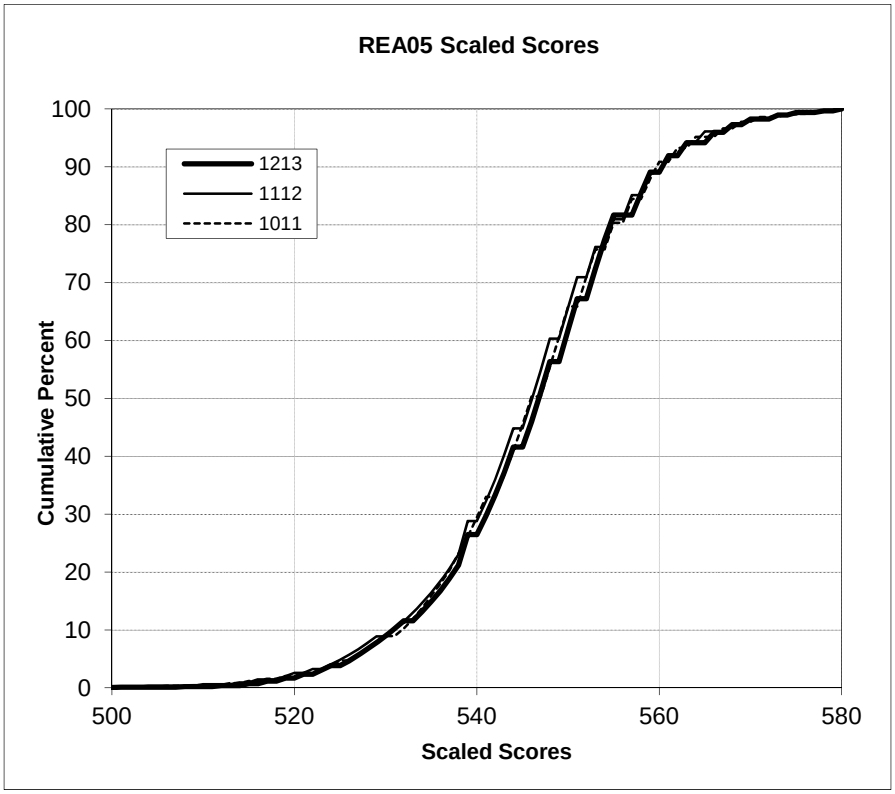


Figure N-11. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 6

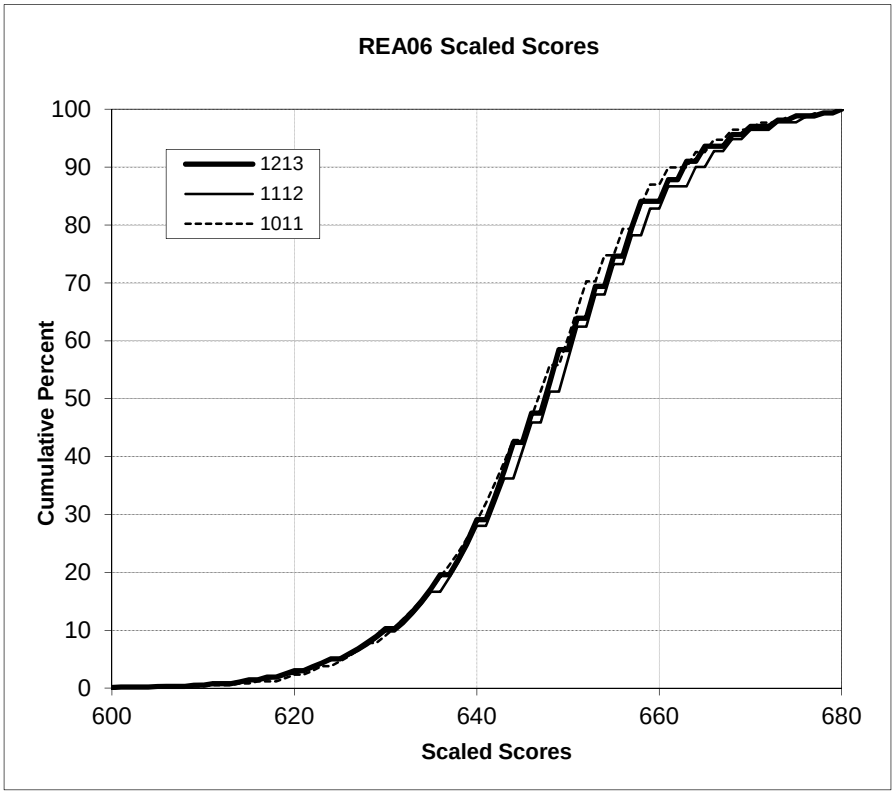


Figure N-12. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 7

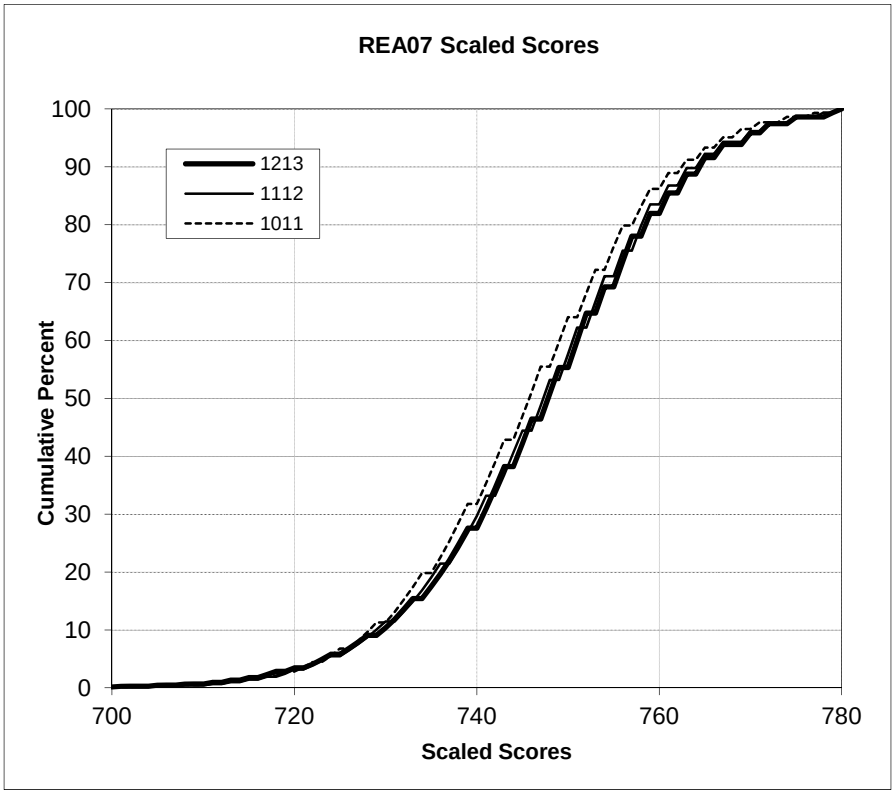


Figure N-13. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 8

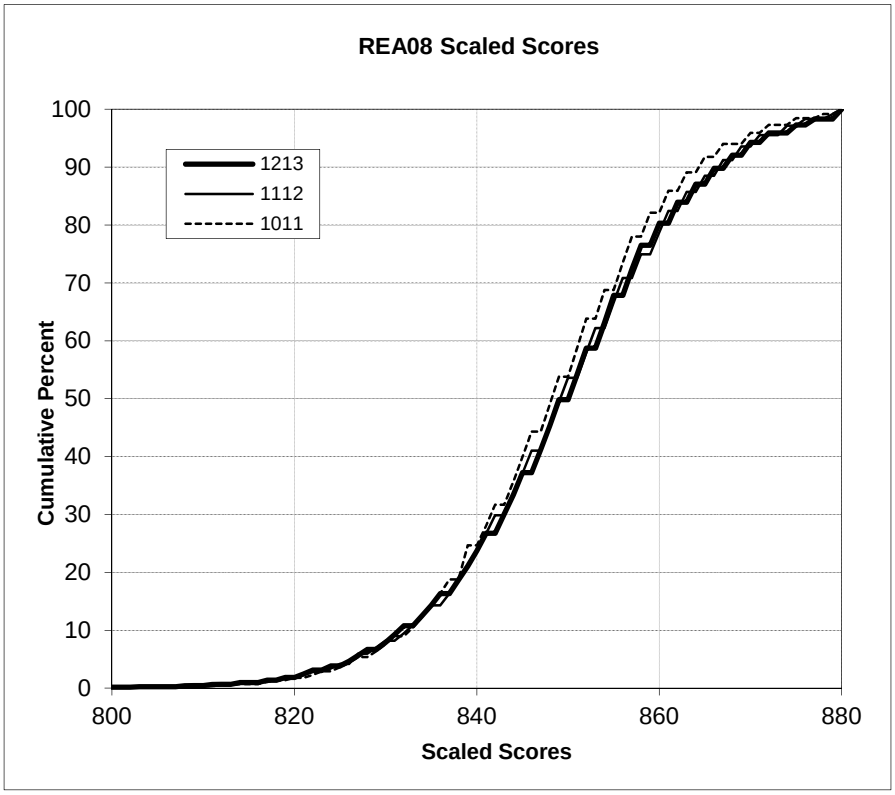


Figure N-14. 2012–13 NECAP: Cumulative Score Distribution Reading Grade 11

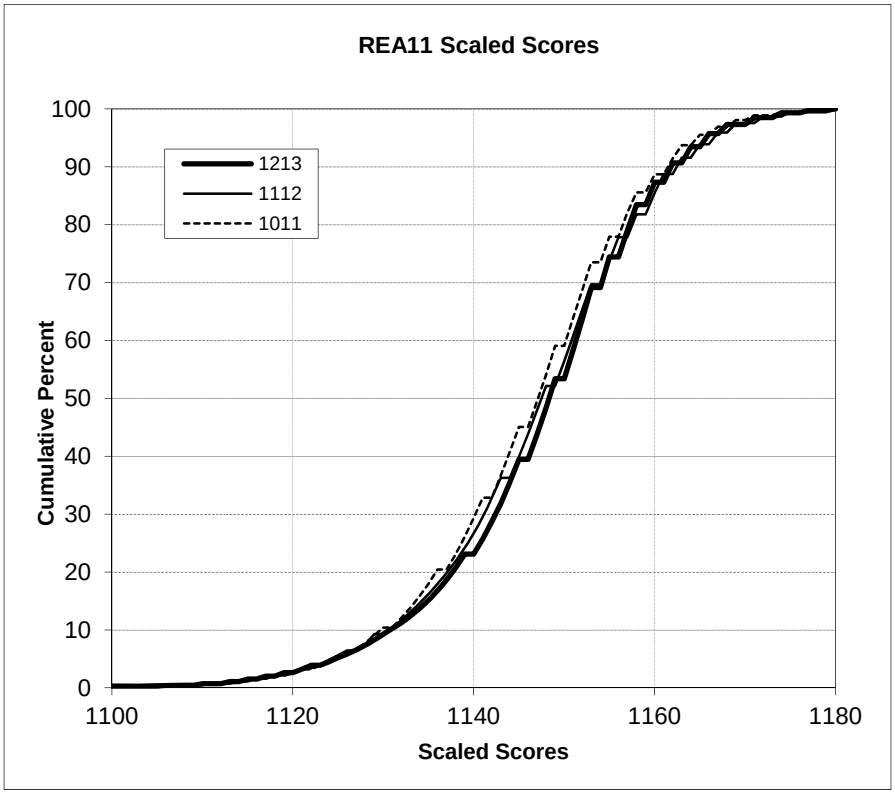


Figure N-15. 2012–13 NECAP: Cumulative Score Distribution Writing Grade 5

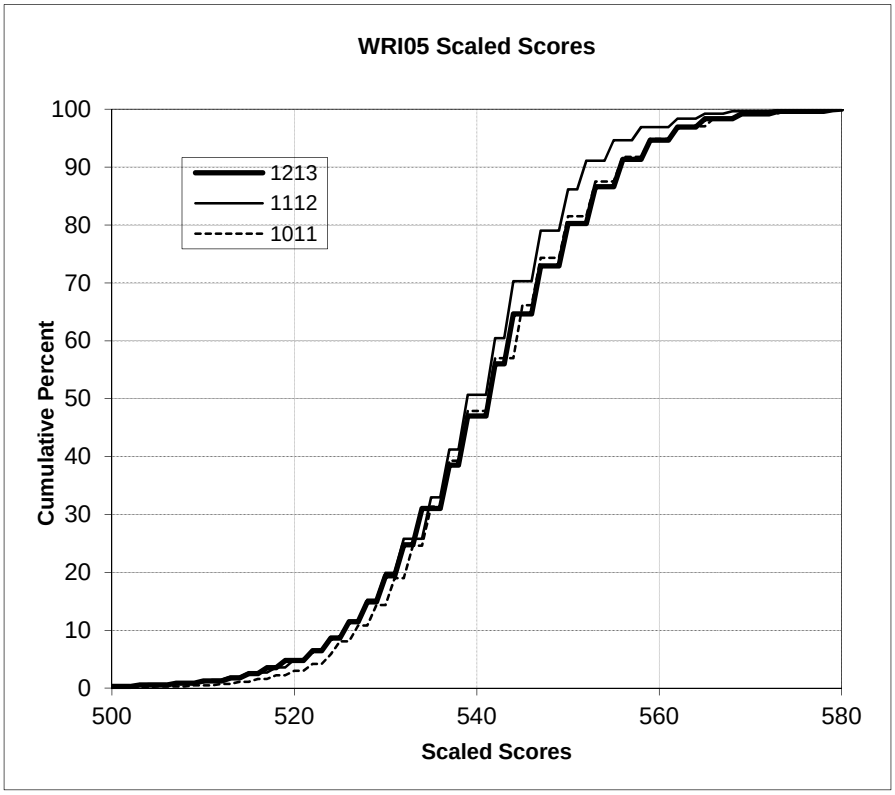


Figure N-16. 2012–13 NECAP: Cumulative Score Distribution Writing Grade 8

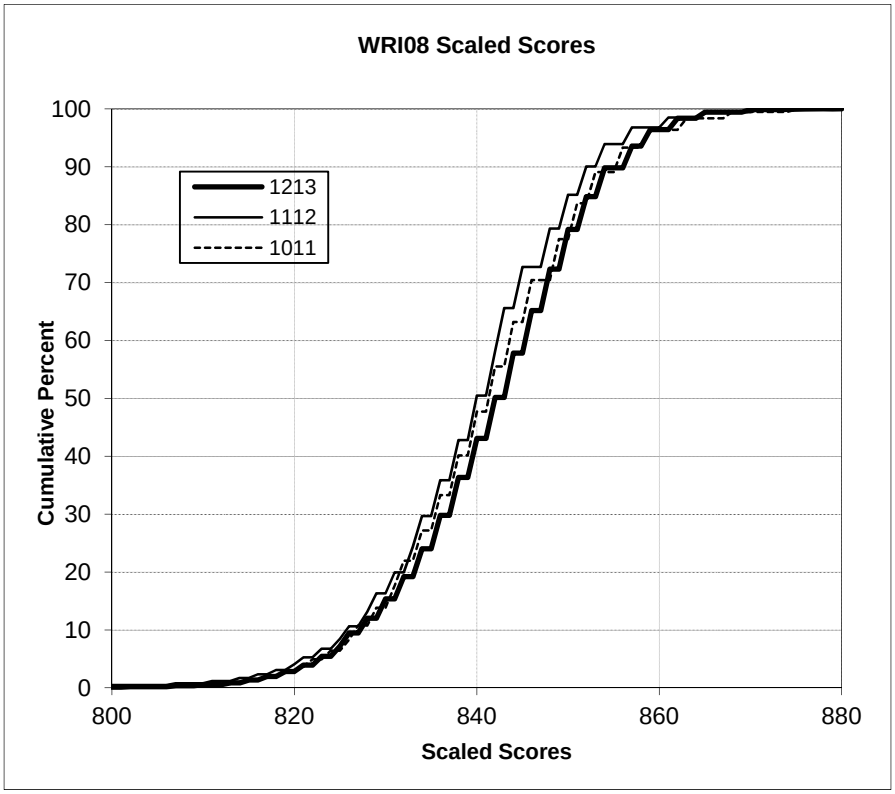


Figure N-17. 2012–13 NECAP: Cumulative Score Distribution
Writing Grade 11

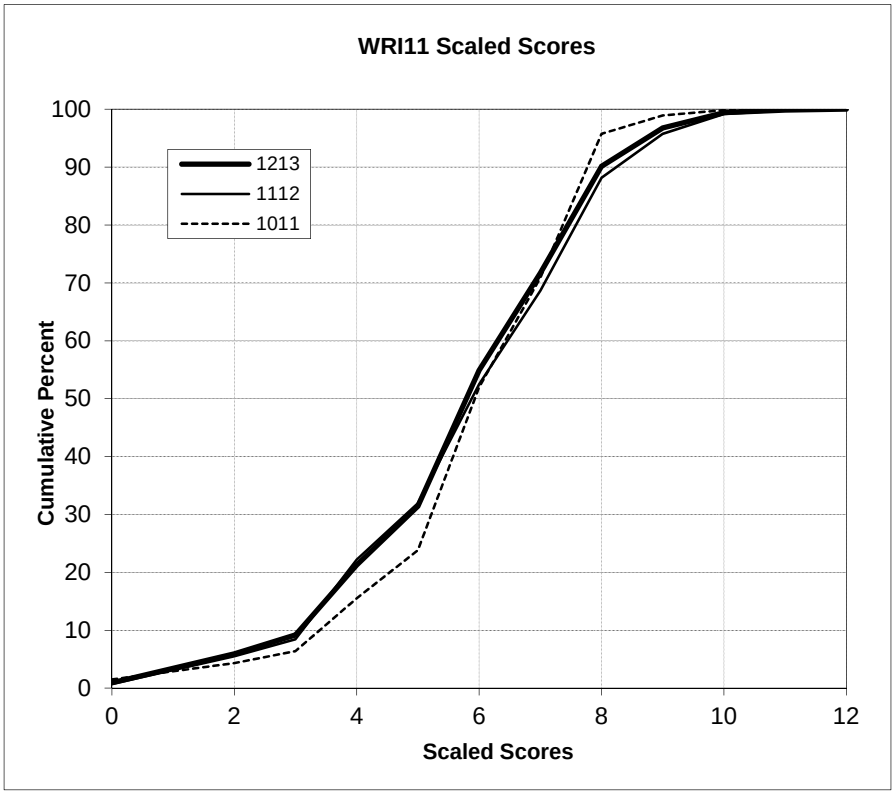


Table N-1. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 3

Scaled Score	*Percentage*	*Cumulative Percentage*
300	0.16	0.16
301	0.00	0.16
302	0.14	0.30
303	0.00	0.30
304	0.00	0.30
305	0.00	0.30
306	0.00	0.30
307	0.17	0.47
308	0.00	0.47
309	0.00	0.47
310	0.00	0.47
311	0.27	0.74
312	0.00	0.74
313	0.41	1.15
314	0.00	1.15
315	0.00	1.15
316	0.49	1.64
317	0.00	1.64
318	0.52	2.17
319	0.63	2.80
320	0.00	2.80
321	0.71	3.51
322	0.71	4.23
323	0.80	5.03
324	0.81	5.84
325	0.90	6.74
326	0.96	7.71
327	1.00	8.70
328	1.04	9.75
329	1.11	10.86
330	1.11	11.97
331	2.51	14.48
332	1.43	15.91
333	1.48	17.39
334	1.53	18.92
335	3.47	22.38
336	1.76	24.14
337	1.98	26.12
338	4.03	30.16
339	4.50	34.66
340	0.00	34.66
341	4.88	39.54
342	2.61	42.16
343	2.71	44.86
344	5.72	50.59
345	2.90	53.49
346	3.02	56.51
347	3.16	59.67
348	3.04	62.72
349	6.03	68.75
350	2.99	71.74
351	3.07	74.81
352	6.14	80.96
353	0.00	80.96
354	2.84	83.80
355	2.74	86.54
356	0.00	86.54
357	2.59	89.12
358	2.36	91.48
359	2.19	93.67
360	0.00	93.67
361	1.92	95.60
362	0.00	95.60
363	1.39	96.98
364	0.00	96.98
365	1.18	98.16
366	0.00	98.16
367	0.80	98.96
368	0.00	98.96
369	0.00	98.96
370	0.58	99.54
371	0.00	99.54
372	0.00	99.54
373	0.00	99.54
374	0.32	99.86
375	0.00	99.86
376	0.00	99.86
377	0.00	99.86
378	0.00	99.86
379	0.00	99.86
380	0.14	100.00

Table N-2. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 4

Scaled Score	*Percentage*	*Cumulative Percentage*
400	0.25	0.25
401	0.23	0.48
402	0.00	0.48
403	0.00	0.48
404	0.00	0.48
405	0.00	0.48
406	0.23	0.72
407	0.00	0.72
408	0.00	0.72
409	0.30	1.01
410	0.00	1.01
411	0.36	1.38
412	0.00	1.38
413	0.43	1.80
414	0.00	1.80
415	0.48	2.29
416	0.00	2.29
417	0.49	2.78
418	0.58	3.36
419	0.00	3.36
420	0.63	3.99
421	0.61	4.60
422	0.70	5.29
423	0.00	5.29
424	0.76	6.05
425	0.78	6.84
426	0.80	7.64
427	0.86	8.49
428	1.10	9.60
429	1.15	10.74
430	2.51	13.25
431	0.00	13.25
432	2.77	16.02
433	1.55	17.57
434	1.59	19.17
435	1.71	20.87
436	1.75	22.62
437	1.90	24.52
438	4.14	28.66
439	2.18	30.84
440	2.21	33.05
441	2.31	35.36
442	2.57	37.93
443	5.16	43.09
444	2.83	45.91
445	2.87	48.78
446	3.03	51.81
447	2.96	54.76
448	2.98	57.75
449	2.93	60.68
450	3.10	63.77
451	3.25	67.02
452	3.20	70.23
453	2.97	73.20
454	2.98	76.17
455	3.00	79.17
456	2.97	82.14
457	2.72	84.86
458	0.00	84.86
459	2.59	87.44
460	2.50	89.94
461	0.00	89.94
462	2.23	92.17
463	1.98	94.15
464	0.00	94.15
465	1.85	96.00
466	0.00	96.00
467	0.00	96.00
468	1.37	97.36
469	0.00	97.36
470	0.00	97.36
471	1.11	98.47
472	0.00	98.47
473	0.00	98.47
474	0.78	99.25
475	0.00	99.25
476	0.00	99.25
477	0.00	99.25
478	0.00	99.25
479	0.48	99.73
480	0.27	100.00

Table N-3. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 5

Scaled Score	*Percentage*	*Cumulative Percentage*
500	0.49	0.49
501	0.00	0.49
502	0.00	0.49
503	0.00	0.49
504	0.00	0.49
505	0.00	0.49
506	0.00	0.49
507	0.00	0.49
508	0.35	0.84
509	0.00	0.84
510	0.00	0.84
511	0.00	0.84
512	0.00	0.84
513	0.47	1.31
514	0.00	1.31
515	0.00	1.31
516	0.00	1.31
517	0.62	1.93
518	0.00	1.93
519	0.83	2.76
520	0.00	2.76
521	0.99	3.74
522	0.00	3.74
523	1.06	4.80
524	1.16	5.96
525	0.00	5.96
526	1.16	7.12
527	1.18	8.30
528	1.33	9.63
529	1.40	11.03
530	1.46	12.49
531	1.50	13.99
532	3.11	17.10
533	0.00	17.10
534	1.65	18.75
535	3.30	22.05
536	1.81	23.86
537	1.95	25.80
538	3.69	29.49
539	4.06	33.55
540	1.97	35.52
541	2.10	37.62
542	4.06	41.69
543	2.19	43.87
544	4.36	48.23
545	4.26	52.49
546	2.19	54.68
547	4.47	59.15
548	2.26	61.42
549	4.40	65.82
550	2.16	67.98
551	4.24	72.21
552	2.29	74.50
553	4.65	79.15
554	2.17	81.32
555	2.12	83.44
556	2.07	85.51
557	2.13	87.64
558	1.94	89.58
559	1.88	91.47
560	1.62	93.09
561	1.53	94.62
562	0.00	94.62
563	1.39	96.01
564	1.21	97.22
565	0.00	97.22
566	1.00	98.23
567	0.00	98.23
568	0.00	98.23
569	0.88	99.10
570	0.00	99.10
571	0.00	99.10
572	0.00	99.10
573	0.50	99.60
574	0.00	99.60
575	0.00	99.60
576	0.00	99.60
577	0.00	99.60
578	0.00	99.60
579	0.28	99.89
580	0.11	100.00

Table N-4. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 6

Scaled Score	*Percentage*	*Cumulative Percentage*
600	1.09	1.09
601	0.00	1.09
602	0.00	1.09
603	0.00	1.09
604	0.00	1.09
605	0.00	1.09
606	0.74	1.84
607	0.00	1.84
608	0.00	1.84
609	0.00	1.84
610	0.00	1.84
611	0.00	1.84
612	0.95	2.79
613	0.00	2.79
614	0.00	2.79
615	0.00	2.79
616	1.08	3.87
617	0.00	3.87
618	0.00	3.87
619	1.20	5.06
620	0.00	5.06
621	0.00	5.06
622	1.15	6.22
623	1.37	7.59
624	0.00	7.59
625	1.25	8.84
626	0.00	8.84
627	1.34	10.18
628	1.39	11.57
629	1.39	12.96
630	1.37	14.32
631	1.47	15.80
632	1.50	17.30
633	1.57	18.87
634	1.58	20.45
635	1.67	22.12
636	3.37	25.49
637	1.80	27.29
638	1.75	29.03
639	3.83	32.87
640	1.94	34.80
641	4.06	38.86
642	2.09	40.95
643	4.21	45.16
644	2.19	47.35
645	4.45	51.80
646	2.13	53.93
647	4.56	58.49
648	2.30	60.79
649	4.48	65.28
650	2.34	67.62
651	4.44	72.05
652	4.40	76.45
653	0.00	76.45
654	4.12	80.58
655	1.96	82.54
656	2.06	84.60
657	1.89	86.49
658	1.84	88.33
659	1.86	90.19
660	1.72	91.91
661	1.68	93.60
662	1.35	94.94
663	0.00	94.94
664	1.36	96.30
665	1.12	97.42
666	0.00	97.42
667	0.96	98.38
668	0.00	98.38
669	0.70	99.07
670	0.00	99.07
671	0.00	99.07
672	0.48	99.55
673	0.00	99.55
674	0.00	99.55
675	0.00	99.55
676	0.30	99.85
677	0.00	99.85
678	0.00	99.85
679	0.00	99.85
680	0.15	100.00

Table N-5. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 7

Scaled Score	*Percentage*	*Cumulative Percentage*
700	0.70	0.70
701	0.00	0.70
702	0.64	1.34
703	0.00	1.34
704	0.00	1.34
705	0.00	1.34
706	0.00	1.34
707	0.00	1.34
708	0.00	1.34
709	0.00	1.34
710	0.00	1.34
711	0.00	1.34
712	0.00	1.34
713	0.86	2.20
714	0.00	2.20
715	0.00	2.20
716	0.00	2.20
717	0.00	2.20
718	1.19	3.39
719	0.00	3.39
720	0.00	3.39
721	1.37	4.76
722	0.00	4.76
723	0.00	4.76
724	1.47	6.23
725	1.64	7.88
726	0.00	7.88
727	1.75	9.63
728	1.66	11.30
729	0.00	11.30
730	1.79	13.08
731	1.81	14.89
732	1.91	16.80
733	1.96	18.76
734	1.83	20.58
735	4.15	24.73
736	2.11	26.85
737	2.07	28.91
738	4.36	33.27
739	4.35	37.62
740	2.33	39.95
741	2.11	42.06
742	4.60	46.66
743	4.32	50.98
744	2.13	53.11
745	4.35	57.46
746	4.25	61.71
747	2.06	63.77
748	4.22	67.99
749	3.98	71.98
750	1.84	73.82
751	3.83	77.65
752	1.79	79.44
753	3.63	83.06
754	1.66	84.72
755	1.69	86.41
756	1.58	87.99
757	3.00	90.99
758	1.28	92.27
759	1.23	93.49
760	1.17	94.66
761	0.00	94.66
762	0.96	95.62
763	0.94	96.57
764	0.80	97.37
765	0.00	97.37
766	0.73	98.10
767	0.00	98.10
768	0.64	98.74
769	0.00	98.74
770	0.49	99.23
771	0.00	99.23
772	0.00	99.23
773	0.36	99.59
774	0.00	99.59
775	0.00	99.59
776	0.00	99.59
777	0.00	99.59
778	0.21	99.80
779	0.00	99.80
780	0.20	100.00

Table N-6. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 8

Scaled Score	*Percentage*	*Cumulative Percentage*
800	1.31	1.31
801	0.00	1.31
802	0.00	1.31
803	0.00	1.31
804	0.00	1.31
805	0.00	1.31
806	0.00	1.31
807	0.00	1.31
808	0.00	1.31
809	0.00	1.31
810	0.00	1.31
811	0.70	2.01
812	0.00	2.01
813	0.00	2.01
814	0.00	2.01
815	0.00	2.01
816	0.00	2.01
817	0.89	2.90
818	0.00	2.90
819	0.00	2.90
820	1.07	3.97
821	0.00	3.97
822	1.16	5.13
823	0.00	5.13
824	1.21	6.34
825	0.00	6.34
826	1.35	7.69
827	1.30	8.99
828	1.32	10.31
829	1.40	11.71
830	1.41	13.12
831	1.42	14.54
832	1.44	15.98
833	3.13	19.10
834	1.70	20.80
835	1.48	22.28
836	3.40	25.68
837	1.87	27.55
838	3.70	31.24
839	5.85	37.10
840	2.16	39.25
841	2.17	41.42
842	4.29	45.71
843	4.44	50.15
844	4.54	54.69
845	4.66	59.35
846	2.34	61.70
847	4.40	66.09
848	4.56	70.66
849	4.56	75.22
850	2.22	77.44
851	4.47	81.91
852	2.10	84.01
853	4.09	88.09
854	1.87	89.97
855	1.74	91.70
856	1.54	93.24
857	1.49	94.74
858	1.23	95.97
859	1.19	97.16
860	0.89	98.04
861	0.69	98.74
862	0.00	98.74
863	0.57	99.30
864	0.00	99.30
865	0.38	99.69
866	0.00	99.69
867	0.00	99.69
868	0.20	99.89
869	0.00	99.89
870	0.00	99.89
871	0.00	99.89
872	0.00	99.89
873	0.00	99.89
874	0.09	99.98
875	0.00	99.98
876	0.00	99.98
877	0.00	99.98
878	0.00	99.98
879	0.00	99.98
880	0.02	100.00

Table N-7. 2012–13 NECAP: Scaled Score Distribution
Mathematics Grade 11

Scaled Score	*Percentage*	*Cumulative Percentage*
1100	2.16	2.16
1101	0.00	2.16
1102	0.00	2.16
1103	0.00	2.16
1104	0.00	2.16
1105	0.00	2.16
1106	0.00	2.16
1107	0.00	2.16
1108	0.00	2.16
1109	0.00	2.16
1110	0.00	2.16
1111	0.00	2.16
1112	0.00	2.16
1113	0.00	2.16
1114	1.69	3.85
1115	0.00	3.85
1116	0.00	3.85
1117	0.00	3.85
1118	0.00	3.85
1119	2.39	6.24
1120	0.00	6.24
1121	0.00	6.24
1122	2.59	8.83
1123	0.00	8.83
1124	2.70	11.53
1125	2.61	14.14
1126	0.00	14.14
1127	2.76	16.90
1128	2.76	19.66
1129	2.71	22.37
1130	2.58	24.95
1131	2.64	27.58
1132	5.05	32.64
1133	4.95	37.59
1134	2.38	39.96
1135	4.56	44.52
1136	4.47	48.99
1137	4.36	53.35
1138	4.10	57.45
1139	5.96	63.41
1140	1.93	65.34
1141	5.31	70.65
1142	3.33	73.98
1143	3.34	77.32
1144	3.05	80.36
1145	2.77	83.14
1146	4.16	87.30
1147	2.58	89.87
1148	1.13	91.00
1149	2.04	93.04
1150	1.85	94.89
1151	2.50	97.39
1152	0.00	97.39
1153	0.67	98.06
1154	0.64	98.69
1155	0.45	99.14
1156	0.33	99.47
1157	0.00	99.47
1158	0.21	99.68
1159	0.13	99.81
1160	0.00	99.81
1161	0.13	99.94
1162	0.00	99.94
1163	0.00	99.94
1164	0.00	99.94
1165	0.05	99.99
1166	0.00	99.99
1167	0.00	99.99
1168	0.00	99.99
1169	0.00	99.99
1170	0.00	99.99
1171	0.00	99.99
1172	0.00	99.99
1173	0.00	99.99
1174	0.00	99.99
1175	0.00	99.99
1176	0.00	99.99
1177	0.00	99.99
1178	0.00	99.99
1179	0.00	99.99
1180	0.01	100.00

Table N-8. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 3

Scaled Score	*Percentage*	*Cumulative Percentage*
300	0.42	0.42
301	0.00	0.42
302	0.00	0.42
303	0.00	0.42
304	0.23	0.65
305	0.00	0.65
306	0.00	0.65
307	0.00	0.65
308	0.34	0.99
309	0.00	0.99
310	0.00	0.99
311	0.45	1.45
312	0.00	1.45
313	0.00	1.45
314	0.49	1.94
315	0.00	1.94
316	0.59	2.53
317	0.00	2.53
318	0.00	2.53
319	0.62	3.16
320	0.00	3.16
321	0.73	3.89
322	0.82	4.71
323	0.00	4.71
324	0.94	5.64
325	0.00	5.64
326	1.05	6.70
327	1.16	7.85
328	1.27	9.12
329	0.00	9.12
330	1.32	10.44
331	1.70	12.15
332	1.71	13.86
333	0.00	13.86
334	1.89	15.74
335	2.06	17.80
336	2.17	19.98
337	2.52	22.50
338	2.65	25.14
339	2.90	28.05
340	0.00	28.05
341	3.14	31.19
342	3.26	34.45
343	3.56	38.01
344	3.90	41.91
345	0.00	41.91
346	4.16	46.07
347	4.40	50.47
348	4.57	55.04
349	4.83	59.87
350	0.00	59.87
351	5.01	64.88
352	5.19	70.07
353	0.00	70.07
354	5.20	75.26
355	0.00	75.26
356	9.65	84.92
357	0.00	84.92
358	0.00	84.92
359	4.13	89.05
360	0.00	89.05
361	0.00	89.05
362	3.64	92.68
363	0.00	92.68
364	2.64	95.33
365	0.00	95.33
366	0.00	95.33
367	1.89	97.21
368	0.00	97.21
369	0.00	97.21
370	1.33	98.54
371	0.00	98.54
372	0.00	98.54
373	0.00	98.54
374	0.84	99.38
375	0.00	99.38
376	0.00	99.38
377	0.00	99.38
378	0.37	99.75
379	0.00	99.75
380	0.25	100.00

Table N-9. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 4

Scaled Score	*Percentage*	*Cumulative Percentage*
400	0.32	0.32
401	0.00	0.32
402	0.00	0.32
403	0.00	0.32
404	0.20	0.52
405	0.00	0.52
406	0.00	0.52
407	0.00	0.52
408	0.24	0.76
409	0.00	0.76
410	0.00	0.76
411	0.42	1.18
412	0.00	1.18
413	0.46	1.64
414	0.00	1.64
415	0.00	1.64
416	0.54	2.18
417	0.00	2.18
418	0.61	2.79
419	0.00	2.79
420	0.66	3.45
421	0.00	3.45
422	0.71	4.16
423	0.71	4.87
424	0.00	4.87
425	0.89	5.76
426	0.94	6.70
427	0.99	7.70
428	0.00	7.70
429	1.16	8.86
430	1.24	10.10
431	1.42	11.52
432	1.61	13.13
433	1.67	14.80
434	1.79	16.59
435	0.00	16.59
436	1.91	18.50
437	2.12	20.62
438	2.29	22.91
439	5.15	28.05
440	0.00	28.05
441	3.04	31.09
442	3.13	34.23
443	3.48	37.71
444	3.81	41.52
445	4.11	45.63
446	4.24	49.87
447	0.00	49.87
448	4.58	54.45
449	4.89	59.34
450	5.30	64.64
451	0.00	64.64
452	5.25	69.89
453	5.25	75.14
454	0.00	75.14
455	5.23	80.37
456	0.00	80.37
457	5.02	85.39
458	0.00	85.39
459	4.25	89.63
460	0.00	89.63
461	0.00	89.63
462	3.66	93.30
463	0.00	93.30
464	0.00	93.30
465	2.89	96.18
466	0.00	96.18
467	0.00	96.18
468	1.89	98.08
469	0.00	98.08
470	0.00	98.08
471	0.00	98.08
472	1.10	99.18
473	0.00	99.18
474	0.00	99.18
475	0.00	99.18
476	0.00	99.18
477	0.00	99.18
478	0.51	99.69
479	0.00	99.69
480	0.31	100.00

Table N-10. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 5

Scaled Score	*Percentage*	*Cumulative Percentage*
500	0.05	0.05
501	0.05	0.10
502	0.00	0.10
503	0.00	0.10
504	0.00	0.10
505	0.00	0.10
506	0.00	0.10
507	0.00	0.10
508	0.11	0.21
509	0.00	0.21
510	0.00	0.21
511	0.00	0.21
512	0.17	0.38
513	0.00	0.38
514	0.00	0.38
515	0.33	0.71
516	0.00	0.71
517	0.44	1.14
518	0.00	1.14
519	0.53	1.67
520	0.00	1.67
521	0.65	2.33
522	0.00	2.33
523	0.76	3.08
524	0.74	3.83
525	0.00	3.83
526	0.82	4.65
527	0.96	5.61
528	1.11	6.72
529	1.04	7.76
530	1.21	8.97
531	1.28	10.25
532	1.31	11.56
533	0.00	11.56
534	1.60	13.16
535	1.72	14.88
536	1.87	16.74
537	2.13	18.87
538	2.31	21.17
539	5.30	26.47
540	0.00	26.47
541	3.21	29.69
542	3.60	33.28
543	3.91	37.19
544	4.37	41.57
545	0.00	41.57
546	4.44	46.00
547	5.04	51.05
548	5.29	56.33
549	0.00	56.33
550	5.47	61.80
551	5.40	67.20
552	0.00	67.20
553	5.22	72.42
554	4.94	77.37
555	4.29	81.65
556	0.00	81.65
557	0.00	81.65
558	3.82	85.48
559	3.60	89.07
560	0.00	89.07
561	2.84	91.91
562	0.00	91.91
563	2.25	94.16
564	0.00	94.16
565	0.00	94.16
566	1.80	95.96
567	0.00	95.96
568	1.32	97.28
569	0.00	97.28
570	0.95	98.24
571	0.00	98.24
572	0.00	98.24
573	0.69	98.93
574	0.00	98.93
575	0.44	99.36
576	0.00	99.36
577	0.00	99.36
578	0.29	99.65
579	0.00	99.65
580	0.35	100.00

Table N-11. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 6

Scaled Score	*Percentage*	*Cumulative Percentage*
600	0.09	0.09
601	0.08	0.17
602	0.00	0.17
603	0.00	0.17
604	0.00	0.17
605	0.12	0.29
606	0.00	0.29
607	0.00	0.29
608	0.00	0.29
609	0.21	0.51
610	0.00	0.51
611	0.26	0.76
612	0.00	0.76
613	0.00	0.76
614	0.29	1.05
615	0.37	1.43
616	0.00	1.43
617	0.46	1.89
618	0.00	1.89
619	0.56	2.44
620	0.56	3.01
621	0.00	3.01
622	0.67	3.68
623	0.63	4.31
624	0.74	5.04
625	0.00	5.04
626	0.88	5.93
627	0.85	6.78
628	1.06	7.85
629	1.09	8.93
630	1.33	10.26
631	0.00	10.26
632	1.44	11.70
633	1.54	13.24
634	1.86	15.11
635	2.10	17.20
636	2.39	19.59
637	0.00	19.59
638	2.74	22.33
639	3.18	25.52
640	3.57	29.09
641	0.00	29.09
642	4.09	33.18
643	4.35	37.52
644	4.91	42.44
645	0.00	42.44
646	5.05	47.49
647	0.00	47.49
648	5.36	52.85
649	5.64	58.48
650	0.00	58.48
651	5.39	63.88
652	0.00	63.88
653	5.50	69.38
654	0.00	69.38
655	5.24	74.61
656	0.00	74.61
657	5.06	79.67
658	4.42	84.09
659	0.00	84.09
660	0.00	84.09
661	3.74	87.83
662	0.00	87.83
663	3.16	91.00
664	0.00	91.00
665	2.61	93.61
666	0.00	93.61
667	0.00	93.61
668	2.01	95.62
669	0.00	95.62
670	1.42	97.03
671	0.00	97.03
672	0.00	97.03
673	1.07	98.10
674	0.00	98.10
675	0.76	98.86
676	0.00	98.86
677	0.00	98.86
678	0.48	99.34
679	0.00	99.34
680	0.66	100.00

Table N-12. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 7

Scaled Score	*Percentage*	*Cumulative Percentage*
700	0.14	0.14
701	0.12	0.27
702	0.00	0.27
703	0.00	0.27
704	0.00	0.27
705	0.19	0.46
706	0.00	0.46
707	0.00	0.46
708	0.20	0.65
709	0.00	0.65
710	0.00	0.65
711	0.27	0.93
712	0.00	0.93
713	0.37	1.30
714	0.00	1.30
715	0.45	1.75
716	0.00	1.75
717	0.54	2.28
718	0.55	2.84
719	0.00	2.84
720	0.59	3.43
721	0.00	3.43
722	0.65	4.07
723	0.81	4.88
724	0.91	5.80
725	0.00	5.80
726	1.01	6.80
727	1.07	7.88
728	1.16	9.04
729	0.00	9.04
730	1.30	10.34
731	1.50	11.84
732	1.74	13.58
733	1.85	15.42
734	0.00	15.42
735	2.06	17.48
736	2.16	19.64
737	2.34	21.98
738	2.73	24.71
739	2.87	27.57
740	0.00	27.57
741	3.34	30.91
742	3.55	34.46
743	3.76	38.22
744	0.00	38.22
745	4.02	42.23
746	4.17	46.41
747	0.00	46.41
748	4.46	50.87
749	4.45	55.32
750	0.00	55.32
751	4.68	60.00
752	4.71	64.71
753	0.00	64.71
754	4.54	69.25
755	0.00	69.25
756	4.45	73.70
757	4.31	78.02
758	0.00	78.02
759	3.91	81.93
760	0.00	81.93
761	3.54	85.47
762	0.00	85.47
763	3.25	88.72
764	0.00	88.72
765	2.85	91.57
766	0.00	91.57
767	2.28	93.85
768	0.00	93.85
769	0.00	93.85
770	2.01	95.86
771	0.00	95.86
772	1.57	97.43
773	0.00	97.43
774	0.00	97.43
775	1.17	98.60
776	0.00	98.60
777	0.00	98.60
778	0.00	98.60
779	0.74	99.34
780	0.66	100.00

Table N-13. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 8

Scaled Score	*Percentage*	*Cumulative Percentage*
800	0.14	0.14
801	0.00	0.14
802	0.00	0.14
803	0.09	0.23
804	0.00	0.23
805	0.00	0.23
806	0.00	0.23
807	0.00	0.23
808	0.20	0.42
809	0.00	0.42
810	0.00	0.42
811	0.23	0.66
812	0.00	0.66
813	0.00	0.66
814	0.30	0.95
815	0.00	0.95
816	0.00	0.95
817	0.41	1.36
818	0.00	1.36
819	0.48	1.84
820	0.00	1.84
821	0.61	2.45
822	0.65	3.10
823	0.00	3.10
824	0.73	3.83
825	0.00	3.83
826	0.82	4.64
827	1.01	5.65
828	1.01	6.66
829	0.00	6.66
830	1.22	7.88
831	1.39	9.27
832	1.47	10.75
833	0.00	10.75
834	1.75	12.50
835	1.77	14.26
836	2.05	16.31
837	0.00	16.31
838	2.37	18.69
839	2.39	21.07
840	2.64	23.72
841	3.03	26.75
842	0.00	26.75
843	3.34	30.09
844	3.31	33.40
845	3.81	37.22
846	0.00	37.22
847	3.97	41.19
848	4.15	45.34
849	4.49	49.83
850	0.00	49.83
851	4.35	54.18
852	4.55	58.73
853	0.00	58.73
854	4.39	63.12
855	4.73	67.85
856	0.00	67.85
857	4.49	72.35
858	4.14	76.48
859	0.00	76.48
860	3.83	80.31
861	0.00	80.31
862	3.59	83.91
863	0.00	83.91
864	3.13	87.04
865	0.00	87.04
866	2.74	89.78
867	0.00	89.78
868	2.29	92.06
869	0.00	92.06
870	2.19	94.25
871	0.00	94.25
872	1.64	95.89
873	0.00	95.89
874	0.00	95.89
875	1.37	97.27
876	0.00	97.27
877	1.04	98.30
878	0.00	98.30
879	0.00	98.30
880	1.70	100.00

Table N-14. 2012–13 NECAP: Scaled Score Distribution
Reading Grade 11

Scaled Score	*Percentage*	*Cumulative Percentage*
1100	0.28	0.28
1101	0.00	0.28
1102	0.00	0.28
1103	0.00	0.28
1104	0.00	0.28
1105	0.00	0.28
1106	0.16	0.43
1107	0.00	0.43
1108	0.00	0.43
1109	0.00	0.43
1110	0.28	0.71
1111	0.00	0.71
1112	0.00	0.71
1113	0.34	1.05
1114	0.00	1.05
1115	0.45	1.50
1116	0.00	1.50
1117	0.52	2.01
1118	0.00	2.01
1119	0.61	2.62
1120	0.00	2.62
1121	0.61	3.23
1122	0.66	3.89
1123	0.00	3.89
1124	0.61	4.50
1125	0.72	5.22
1126	0.62	5.84
1127	0.69	6.53
1128	0.84	7.37
1129	0.87	8.24
1130	1.01	9.26
1131	0.97	10.23
1132	1.00	11.23
1133	1.16	12.39
1134	1.29	13.68
1135	1.43	15.11
1136	1.70	16.81
1137	1.84	18.64
1138	2.05	20.70
1139	2.39	23.08
1140	0.00	23.08
1141	2.63	25.72
1142	2.97	28.69
1143	3.16	31.85
1144	3.64	35.49
1145	3.97	39.46
1146	0.00	39.46
1147	4.29	43.75
1148	4.67	48.42
1149	4.96	53.38
1150	0.00	53.38
1151	5.03	58.41
1152	5.30	63.71
1153	5.44	69.15
1154	0.00	69.15
1155	5.26	74.41
1156	0.00	74.41
1157	4.63	79.04
1158	4.42	83.45
1159	0.00	83.45
1160	3.86	87.32
1161	0.00	87.32
1162	3.36	90.68
1163	0.00	90.68
1164	2.77	93.46
1165	0.00	93.46
1166	2.27	95.73
1167	0.00	95.73
1168	1.57	97.29
1169	0.00	97.29
1170	0.00	97.29
1171	1.20	98.50
1172	0.00	98.50
1173	0.00	98.50
1174	0.81	99.31
1175	0.00	99.31
1176	0.00	99.31
1177	0.41	99.72
1178	0.00	99.72
1179	0.00	99.72
1180	0.28	100.00

Table N-15. 2012–13 NECAP: Scaled Score Distribution
Writing Grade 5

Scaled Score	*Percentage*	*Cumulative Percentage*
500	0.30	0.30
501	0.00	0.30
502	0.00	0.30
503	0.23	0.54
504	0.00	0.54
505	0.00	0.54
506	0.00	0.54
507	0.30	0.83
508	0.00	0.83
509	0.00	0.83
510	0.39	1.23
511	0.00	1.23
512	0.00	1.23
513	0.52	1.75
514	0.00	1.75
515	0.76	2.51
516	0.00	2.51
517	1.01	3.52
518	0.00	3.52
519	1.25	4.77
520	0.00	4.77
521	0.00	4.77
522	1.67	6.45
523	0.00	6.45
524	2.19	8.64
525	0.00	8.64
526	2.83	11.46
527	0.00	11.46
528	3.41	14.87
529	0.00	14.87
530	4.52	19.39
531	0.00	19.39
532	5.36	24.74
533	0.00	24.74
534	6.28	31.03
535	0.00	31.03
536	0.00	31.03
537	7.50	38.52
538	0.00	38.52
539	8.47	46.99
540	0.00	46.99
541	0.00	46.99
542	9.01	56.00
543	0.00	56.00
544	8.63	64.63
545	0.00	64.63
546	0.00	64.63
547	8.29	72.92
548	0.00	72.92
549	0.00	72.92
550	7.33	80.25
551	0.00	80.25
552	0.00	80.25
553	6.36	86.60
554	0.00	86.60
555	0.00	86.60
556	4.72	91.32
557	0.00	91.32
558	0.00	91.32
559	3.32	94.64
560	0.00	94.64
561	0.00	94.64
562	2.28	96.92
563	0.00	96.92
564	0.00	96.92
565	1.43	98.35
566	0.00	98.35
567	0.00	98.35
568	0.00	98.35
569	0.84	99.19
570	0.00	99.19
571	0.00	99.19
572	0.00	99.19
573	0.44	99.63
574	0.00	99.63
575	0.00	99.63
576	0.00	99.63
577	0.00	99.63
578	0.00	99.63
579	0.26	99.89
580	0.11	100.00

Table N-16. 2012–13 NECAP: Scaled Score Distribution
Writing Grade 8

Scaled Score	*Percentage*	*Cumulative Percentage*
800	0.09	0.09
801	0.00	0.09
802	0.10	0.18
803	0.00	0.18
804	0.00	0.18
805	0.00	0.18
806	0.00	0.18
807	0.16	0.35
808	0.00	0.35
809	0.00	0.35
810	0.18	0.53
811	0.00	0.53
812	0.00	0.53
813	0.32	0.85
814	0.00	0.85
815	0.48	1.33
816	0.00	1.33
817	0.62	1.95
818	0.00	1.95
819	0.85	2.80
820	0.00	2.80
821	1.12	3.92
822	0.00	3.92
823	1.50	5.42
824	0.00	5.42
825	1.79	7.21
826	2.25	9.46
827	0.00	9.46
828	2.56	12.02
829	0.00	12.02
830	3.32	15.34
831	0.00	15.34
832	3.86	19.20
833	0.00	19.20
834	4.81	24.01
835	0.00	24.01
836	5.78	29.79
837	0.00	29.79
838	6.55	36.34
839	0.00	36.34
840	6.74	43.07
841	0.00	43.07
842	7.07	50.14
843	0.00	50.14
844	7.67	57.81
845	0.00	57.81
846	7.34	65.15
847	0.00	65.15
848	7.15	72.31
849	0.00	72.31
850	6.85	79.16
851	0.00	79.16
852	5.68	84.84
853	0.00	84.84
854	4.96	89.80
855	0.00	89.80
856	0.00	89.80
857	3.78	93.58
858	0.00	93.58
859	2.86	96.43
860	0.00	96.43
861	0.00	96.43
862	1.96	98.39
863	0.00	98.39
864	0.00	98.39
865	1.03	99.42
866	0.00	99.42
867	0.00	99.42
868	0.00	99.42
869	0.00	99.42
870	0.45	99.87
871	0.00	99.87
872	0.00	99.87
873	0.00	99.87
874	0.00	99.87
875	0.10	99.97
876	0.00	99.97
877	0.00	99.97
878	0.00	99.97
879	0.00	99.97
880	0.03	100.00

Table N-17. 2012–13 NECAP: Scaled Score Distribution
Writing Grade 11

Scaled Score	*Percentage*	*Cumulative Percentage*
0	0.92	0.92
2	5.00	5.93
3	3.26	9.19
4	12.08	21.27
5	10.15	31.42
6	23.48	54.91
7	16.79	71.70
8	18.44	90.14
9	6.63	96.77
10	2.63	99.40
11	0.44	99.85
12	0.15	100.00

APPENDIX O—CLASSICAL RELIABILTIES

Table O-1. 2012–13 NECAP: Subgroup Reliabilities
Mathematics

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	All Students	43,978	65	40.50	12.71	0.93	3.39
	Male	22,519	65	40.64	12.71	0.93	3.38
	Female	21,436	65	40.33	12.70	0.93	3.40
	Gender Not Reported	23	65	54.00	7.75	0.86	2.88
	Hispanic or Latino	3,567	65	33.31	13.13	0.93	3.45
	American Indian or Alaskan Native	260	65	34.59	12.02	0.92	3.48
	Asian	1,128	65	42.88	13.42	0.94	3.30
	Black or African American	1,679	65	31.77	13.15	0.93	3.45
	Native Hawaiian or Pacific Islander	57	65	39.35	11.85	0.92	3.41
	White (non-Hispanic)	36,207	65	41.64	12.18	0.92	3.38
	Two or More Races (non-Hispanic)	995	65	38.59	13.16	0.93	3.42
	No Primary Race/Ethnicity Reported	85	65	40.32	17.09	0.96	3.26
	Currently receiving LEP services	2,072	65	30.35	13.34	0.93	3.46
	Former LEP student - monitoring year 1	111	65	45.23	10.81	0.91	3.24
	Former LEP student - monitoring year 2	34	65	45.44	12.11	0.93	3.26
	LEP: All Other Students	41,761	65	40.99	12.47	0.93	3.38
	Students with an IEP	5,838	65	30.29	13.49	0.93	3.52
	IEP: All Other Students	38,140	65	42.06	11.83	0.92	3.36
	Economically Disadvantaged Students	18,758	65	35.63	12.69	0.93	3.47
	SES: All Other Students	25,220	65	44.12	11.45	0.92	3.31
	Migrant Students	28	65	29.46	12.69	0.92	3.54
	Migrant: All Other Students	43,950	65	40.51	12.71	0.93	3.39
	Students receiving Title 1 Services	11,113	65	34.88	12.73	0.93	3.47
	Title 1: All Other Students	32,865	65	42.40	12.12	0.92	3.35
	Plan 504	499	65	40.26	12.22	0.92	3.41
	Plan 504: All Other Students	43,479	65	40.50	12.71	0.93	3.39
4	All Students	43,558	65	41.15	12.72	0.93	3.47
	Male	22,451	65	41.03	12.97	0.93	3.46
	Female	21,081	65	41.28	12.44	0.92	3.46
	Gender Not Reported	26	65	39.62	12.73	0.92	3.53
	Hispanic or Latino	3,396	65	34.11	13.46	0.93	3.53
	American Indian or Alaskan Native	224	65	35.89	12.31	0.92	3.58
	Asian	1,085	65	43.16	13.42	0.94	3.37
	Black or African American	1,645	65	33.04	13.18	0.93	3.55
	Native Hawaiian or Pacific Islander	42	65	38.67	14.37	0.94	3.49
	White (non-Hispanic)	36,203	65	42.21	12.23	0.92	3.45
	Two or More Races (non-Hispanic)	901	65	39.51	12.85	0.93	3.50
	No Primary Race/Ethnicity Reported	62	65	36.60	13.51	0.93	3.56
	Currently receiving LEP services	1,663	65	28.75	13.13	0.93	3.53
	Former LEP student - monitoring year 1	419	65	39.46	11.28	0.90	3.53
	Former LEP student - monitoring year 2	126	65	45.39	10.79	0.90	3.34
	LEP: All Other Students	41,350	65	41.66	12.46	0.92	3.46
	Students with an IEP	6,026	65	28.66	12.95	0.92	3.57
	IEP: All Other Students	37,532	65	43.16	11.48	0.91	3.44

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
4	Economically Disadvantaged Students	18,159	65	36.09	12.79	0.92	3.56
	SES: All Other Students	25,399	65	44.77	11.36	0.91	3.38
	Migrant Students	24	65	32.92	10.79	0.89	3.61
	Migrant: All Other Students	43,534	65	41.16	12.72	0.93	3.47
	Students receivng Title 1 Services	10,449	65	35.83	12.72	0.92	3.56
	Title 1: All Other Students	33,109	65	42.83	12.25	0.92	3.43
	Plan 504	625	65	40.03	12.04	0.91	3.52
	Plan 504: All Other Students	42,933	65	41.17	12.73	0.93	3.46
5	All Students	44,020	66	37.55	14.67	0.93	3.97
	Male	22,569	66	37.27	14.81	0.93	3.97
	Female	21,406	66	37.85	14.51	0.93	3.97
	Gender Not Reported	45	66	35.51	13.39	0.91	3.94
	Hispanic or Latino	3,339	66	29.46	14.18	0.92	3.93
	American Indian or Alaskan Native	238	66	31.29	14.76	0.93	3.95
	Asian	1,043	66	41.22	15.63	0.94	3.90
	Black or African American	1,590	66	28.18	13.69	0.92	3.90
	Native Hawaiian or Pacific Islander	47	66	36.36	15.65	0.94	3.92
	White (non-Hispanic)	36,732	66	38.70	14.30	0.92	3.96
	Two or More Races (non-Hispanic)	936	66	35.16	15.00	0.93	3.99
	No Primary Race/Ethnicity Reported	95	66	32.13	14.40	0.92	3.99
	Currently receiving LEP services	1,325	66	21.43	12.31	0.91	3.73
	Former LEP student - monitoring year 1	317	66	34.93	13.52	0.91	4.04
	Former LEP student - monitoring year 2	400	66	36.27	13.23	0.91	4.00
	LEP: All Other Students	41,978	66	38.09	14.46	0.92	3.97
	Students with an IEP	6,506	66	23.98	12.51	0.91	3.82
	IEP: All Other Students	37,514	66	39.90	13.71	0.92	3.96
	Economically Disadvantaged Students	18,326	66	31.57	13.93	0.92	3.98
	SES: All Other Students	25,694	66	41.81	13.66	0.92	3.92
	Migrant Students	19	66	26.42	15.58	0.94	3.81
	Migrant: All Other Students	44,001	66	37.55	14.66	0.93	3.97
	Students receiving Title 1 Services	9,906	66	30.66	13.77	0.92	3.95
	Title 1: All Other Students	34,114	66	39.55	14.31	0.92	3.96
	Plan 504	724	66	34.65	13.79	0.91	4.04
	Plan 504: All Other Students	43,296	66	37.60	14.68	0.93	3.97
6	All Students	44,582	66	35.80	14.90	0.93	4.02
	Male	22,933	66	35.46	15.09	0.93	4.01
	Female	21,586	66	36.15	14.68	0.92	4.04
	Gender Not Reported	63	66	39.81	14.38	0.93	3.89
	Hispanic or Latino	3,330	66	27.49	14.41	0.92	3.96
	American Indian or Alaskan Native	259	66	28.79	15.02	0.93	3.94
	Asian	1,091	66	40.03	16.15	0.94	4.00
	Black or African American	1,638	66	25.77	13.92	0.92	3.89
	Native Hawaiian or Pacific Islander	34	66	33.06	17.06	0.94	4.09
	White (non-Hispanic)	37,165	66	36.97	14.47	0.92	4.02
	Two or More Races (non-Hispanic)	932	66	33.55	15.76	0.94	4.01
	No Primary Race/Ethnicity Reported	133	66	34.86	15.82	0.94	3.98
	Currently receiving LEP services	1,294	66	19.67	12.41	0.91	3.65
	Former LEP student - monitoring year 1	233	66	29.73	12.89	0.90	4.06

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
6	Former LEP student - monitoring year 2	342	66	33.39	14.14	0.92	4.11
	LEP: All Other Students	42,713	66	36.34	14.70	0.93	4.02
	Students with an IEP	6,731	66	20.43	11.91	0.90	3.68
	IEP: All Other Students	37,851	66	38.53	13.66	0.91	4.01
	Economically Disadvantaged Students	18,364	66	29.83	14.12	0.92	3.99
	SES: All Other Students	26,218	66	39.98	13.97	0.92	3.99
	Migrant Students	16	66	24.19	13.40	0.92	3.72
	Migrant: All Other Students	44,566	66	35.81	14.90	0.93	4.02
	Students receiving Title 1 Services	8,434	66	28.59	14.09	0.92	3.96
	Title 1: All Other Students	36,148	66	37.48	14.57	0.92	4.02
	Plan 504	802	66	32.55	13.42	0.91	4.02
	Plan 504: All Other Students	43,780	66	35.86	14.92	0.93	4.02
7	All Students	44,518	66	32.52	14.68	0.93	3.85
	Male	23,086	66	32.50	14.92	0.93	3.84
	Female	21,372	66	32.52	14.41	0.93	3.85
	Gender Not Reported	60	66	37.42	14.09	0.92	3.89
	Hispanic or Latino	3,160	66	24.43	13.09	0.92	3.63
	American Indian or Alaskan Native	242	66	25.97	13.89	0.93	3.67
	Asian	961	66	38.54	15.65	0.94	3.92
	Black or African American	1,649	66	23.16	12.76	0.92	3.57
	Native Hawaiian or Pacific Islander	44	66	31.57	15.25	0.94	3.84
	White (non-Hispanic)	37,564	66	33.54	14.46	0.93	3.86
	Two or More Races (non-Hispanic)	790	66	30.66	14.82	0.93	3.80
	No Primary Race/Ethnicity Reported	108	66	31.54	16.17	0.94	3.84
	Currently receiving LEP services	1,231	66	17.20	10.81	0.91	3.31
	Former LEP student - monitoring year 1	124	66	30.27	13.03	0.91	3.90
	Former LEP student - monitoring year 2	198	66	32.13	13.93	0.92	3.89
	LEP: All Other Students	42,965	66	32.96	14.55	0.93	3.86
	Students with an IEP	6,834	66	17.97	10.65	0.90	3.36
	IEP: All Other Students	37,684	66	35.16	13.74	0.92	3.89
	Economically Disadvantaged Students	17,934	66	26.39	13.31	0.92	3.69
	SES: All Other Students	26,584	66	36.65	14.11	0.92	3.90
	Migrant Students	17	66	18.82	10.84	0.91	3.34
	Migrant: All Other Students	44,501	66	32.52	14.68	0.93	3.85
	Students receiving Title 1 Services	5,690	66	25.43	12.99	0.92	3.64
	Title 1: All Other Students	38,828	66	33.56	14.63	0.93	3.87
	Plan 504	918	66	30.46	13.45	0.92	3.84
	Plan 504: All Other Students	43,600	66	32.56	14.70	0.93	3.85
8	All Students	44,586	66	35.46	14.69	0.93	3.95
	Male	23,019	66	35.10	14.87	0.93	3.93
	Female	21,492	66	35.83	14.50	0.93	3.96
	Gender Not Reported	75	66	39.33	13.51	0.92	3.79
	Hispanic or Latino	2,950	66	26.66	14.11	0.93	3.85
	American Indian or Alaskan Native	261	66	29.36	13.45	0.91	3.94
	Asian	945	66	39.34	15.60	0.94	3.92
	Black or African American	1,606	66	25.90	14.19	0.93	3.84
	Native Hawaiian or Pacific Islander	35	66	33.69	16.11	0.94	3.91
	White (non-Hispanic)	37,885	66	36.53	14.33	0.92	3.95

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
8	Two or More Races (non-Hispanic)	775	66	34.03	14.86	0.93	3.97
	No Primary Race/Ethnicity Reported	129	66	33.50	15.88	0.94	3.95
	Currently receiving LEP services	1,174	66	19.01	12.00	0.91	3.53
	Former LEP student - monitoring year 1	72	66	32.04	15.59	0.93	4.00
	Former LEP student - monitoring year 2	144	66	37.12	13.12	0.91	3.93
	LEP: All Other Students	43,196	66	35.90	14.51	0.93	3.95
	Students with an IEP	6,825	66	20.13	11.82	0.91	3.59
	IEP: All Other Students	37,761	66	38.23	13.40	0.91	3.92
	Economically Disadvantaged Students	17,248	66	29.12	13.89	0.92	3.91
	SES: All Other Students	27,338	66	39.45	13.75	0.92	3.91
	Migrant Students	14	66	24.50	13.01	0.92	3.76
	Migrant: All Other Students	44,572	66	35.46	14.69	0.93	3.95
	Students receiving Title 1 Services	4,492	66	27.28	13.80	0.92	3.88
	Title 1: All Other Students	40,094	66	36.37	14.51	0.93	3.95
	Plan 504	1,060	66	32.61	13.32	0.91	3.95
	Plan 504: All Other Students	43,526	66	35.53	14.72	0.93	3.95
11	All Students	31,249	64	25.99	14.69	0.94	3.58
	Male	15,935	64	26.52	15.34	0.94	3.60
	Female	15,304	64	25.46	13.95	0.94	3.54
	Gender Not Reported	10	64	12.20	7.76	0.86	2.95
	Hispanic or Latino	2,624	64	17.51	11.81	0.92	3.26
	American Indian or Alaskan Native	96	64	22.64	13.31	0.93	3.46
	Asian	808	64	29.43	16.08	0.95	3.65
	Black or African American	1,251	64	17.16	11.55	0.92	3.24
	Native Hawaiian or Pacific Islander	43	64	24.26	12.99	0.93	3.47
	White (non-Hispanic)	25,719	64	27.30	14.57	0.94	3.61
	Two or More Races (non-Hispanic)	501	64	24.92	15.46	0.95	3.51
	No Primary Race/Ethnicity Reported	207	64	15.81	12.41	0.94	3.09
	Currently receiving LEP services	674	64	10.14	7.82	0.89	2.60
	Former LEP student - monitoring year 1	127	64	14.87	9.33	0.89	3.03
	Former LEP student - monitoring year 2	124	64	19.47	12.04	0.93	3.29
	LEP: All Other Students	30,324	64	26.42	14.61	0.94	3.59
	Students with an IEP	4,392	64	12.42	8.85	0.89	2.89
	IEP: All Other Students	26,857	64	28.21	14.25	0.94	3.63
	Economically Disadvantaged Students	9,045	64	19.37	12.46	0.93	3.36
	SES: All Other Students	22,204	64	28.69	14.67	0.94	3.63
	Migrant Students	6	64				
	Migrant: All Other Students	31,243	64	25.99	14.69	0.94	3.58
	Students receiving Title 1 Services	2,495	64	18.64	12.48	0.93	3.34
	Title 1: All Other Students	28,754	64	26.63	14.69	0.94	3.59
	Plan 504	668	64	22.51	13.61	0.93	3.51
	Plan 504: All Other Students	30,581	64	26.07	14.70	0.94	3.58

Table O-2. 2012–13 NECAP: Subgroup Reliabilities
Reading

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	All Students	43,834	52	32.80	8.74	0.88	3.04
	Male	22,451	52	31.77	8.91	0.88	3.05
	Female	21,360	52	33.88	8.42	0.87	3.02
	Gender Not Reported	23	52	40.74	3.65	0.57	2.38
	Hispanic or Latino	3,497	52	28.92	9.33	0.88	3.22
	American Indian or Alaskan Native	260	52	29.03	8.56	0.86	3.18
	Asian	1,101	52	33.93	8.88	0.88	3.04
	Black or African American	1,656	52	28.45	9.42	0.88	3.23
	Native Hawaiian or Pacific Islander	57	52	32.28	8.31	0.85	3.19
	White (non-Hispanic)	36,182	52	33.39	8.47	0.87	3.01
	Two or More Races (non-Hispanic)	995	52	32.08	8.92	0.88	3.12
	No Primary Race/Ethnicity Reported	86	52	32.62	10.64	0.91	3.11
	Currently receiving LEP services	1,948	52	26.00	9.40	0.87	3.33
	Former LEP student - monitoring year 1	111	52	36.22	6.30	0.78	2.93
	Former LEP student - monitoring year 2	34	52	37.47	6.65	0.81	2.86
	LEP: All Other Students	41,741	52	33.11	8.58	0.88	3.03
	Students with an IEP	5,817	52	24.11	9.55	0.88	3.28
	IEP: All Other Students	38,017	52	34.13	7.79	0.85	2.98
	Economically Disadvantaged Students	18,684	52	29.70	9.04	0.88	3.17
	SES: All Other Students	25,150	52	35.11	7.73	0.86	2.92
	Migrant Students	28	52	26.36	9.91	0.89	3.22
	Migrant: All Other Students	43,806	52	32.81	8.73	0.88	3.04
	Students receiving Title 1 Services	11,461	52	29.40	8.93	0.87	3.19
	Title 1: All Other Students	32,373	52	34.01	8.34	0.87	2.98
	Plan 504	500	52	32.28	7.97	0.85	3.05
	Plan 504: All Other Students	43,334	52	32.81	8.74	0.88	3.04
4	All Students	43,433	52	34.62	8.99	0.89	3.05
	Male	22,381	52	33.50	9.24	0.89	3.09
	Female	21,026	52	35.80	8.55	0.88	2.98
	Gender Not Reported	26	52	37.96	8.62	0.88	3.02
	Hispanic or Latino	3,332	52	29.98	9.82	0.89	3.28
	American Indian or Alaskan Native	224	52	31.47	9.23	0.88	3.22
	Asian	1,054	52	35.61	9.23	0.90	2.99
	Black or African American	1,629	52	29.79	9.73	0.89	3.28
	Native Hawaiian or Pacific Islander	42	52	35.02	9.66	0.90	3.09
	White (non-Hispanic)	36,191	52	35.28	8.65	0.88	3.00
	Two or More Races (non-Hispanic)	901	52	33.59	9.14	0.88	3.12
	No Primary Race/Ethnicity Reported	60	52	34.92	10.14	0.90	3.20
	Currently receiving LEP services	1,544	52	24.98	9.74	0.88	3.41
	Former LEP student - monitoring year 1	419	52	33.60	7.16	0.81	3.10
	Former LEP student - monitoring year 2	126	52	38.37	6.74	0.83	2.74
	LEP: All Other Students	41,344	52	34.98	8.78	0.88	3.02
	Students with an IEP	6,018	52	24.46	9.53	0.87	3.39
	IEP: All Other Students	37,415	52	36.25	7.74	0.86	2.94
	Economically Disadvantaged Students	18,089	52	31.21	9.43	0.88	3.22
	SES: All Other Students	25,344	52	37.05	7.79	0.86	2.88

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
4	Migrant Students	22	52	27.18	8.83	0.86	3.26
	Migrant: All Other Students	43,411	52	34.62	8.99	0.89	3.05
	Students receiving Title 1 Services	10,616	52	30.94	9.23	0.88	3.23
	Title 1: All Other Students	32,817	52	35.81	8.58	0.88	2.97
	Plan 504	626	52	34.42	8.10	0.86	3.08
	Plan 504: All Other Students	42,807	52	34.62	9.00	0.89	3.05
5	All Students	43,892	52	29.91	8.63	0.89	2.89
	Male	22,498	52	28.68	8.57	0.89	2.82
	Female	21,349	52	31.20	8.50	0.88	2.91
	Gender Not Reported	45	52	31.09	7.77	0.87	2.84
	Hispanic or Latino	3,272	52	25.31	9.47	0.90	3.01
	American Indian or Alaskan Native	237	52	25.99	8.73	0.89	2.94
	Asian	1,015	52	31.40	9.02	0.90	2.92
	Black or African American	1,571	52	25.73	9.41	0.89	3.05
	Native Hawaiian or Pacific Islander	47	52	31.40	6.99	0.83	2.90
	White (non-Hispanic)	36,720	52	30.50	8.31	0.88	2.86
	Two or More Races (non-Hispanic)	935	52	29.17	8.83	0.89	2.90
	No Primary Race/Ethnicity Reported	95	52	28.26	9.58	0.91	2.90
	Currently receiving LEP services	1,206	52	19.51	9.05	0.88	3.08
	Former LEP student - monitoring year 1	318	52	28.49	7.61	0.84	3.03
	Former LEP student - monitoring year 2	399	52	28.78	7.96	0.85	3.03
	LEP: All Other Students	41,969	52	30.23	8.44	0.88	2.88
	Students with an IEP	6,503	52	20.41	8.69	0.88	2.96
	IEP: All Other Students	37,389	52	31.56	7.47	0.85	2.85
	Economically Disadvantaged Students	18,260	52	26.66	8.88	0.89	2.96
	SES: All Other Students	25,632	52	32.22	7.64	0.86	2.82
	Migrant Students	17	52	24.00	10.44	0.90	3.27
	Migrant: All Other Students	43,875	52	29.91	8.63	0.89	2.89
	Students receivng Title 1 Services	9,977	52	26.29	8.78	0.89	2.98
	Title 1: All Other Students	33,915	52	30.97	8.29	0.88	2.85
	Plan 504	724	52	28.94	7.61	0.86	2.83
	Plan 504: All Other Students	43,168	52	29.92	8.64	0.89	2.89
6	All Students	44,482	52	31.77	8.25	0.88	2.82
	Male	22,886	52	30.29	8.28	0.89	2.77
	Female	21,533	52	33.33	7.92	0.87	2.81
	Gender Not Reported	63	52	35.78	6.35	0.82	2.67
	Hispanic or Latino	3,277	52	27.97	9.18	0.89	2.98
	American Indian or Alaskan Native	259	52	27.06	9.27	0.90	2.92
	Asian	1,067	52	33.26	8.78	0.89	2.85
	Black or African American	1,616	52	27.35	9.13	0.89	3.02
	Native Hawaiian or Pacific Islander	34	52	31.29	9.39	0.91	2.89
	White (non-Hispanic)	37,165	52	32.31	7.91	0.88	2.78
	Two or More Races (non-Hispanic)	931	52	30.80	8.78	0.89	2.89
	No Primary Race/Ethnicity Reported	133	52	32.01	9.22	0.90	2.85
	Currently receiving LEP services	1,185	52	21.40	8.90	0.88	3.08
	Former LEP student - monitoring year 1	233	52	29.83	7.36	0.83	3.00
	Former LEP student - monitoring year 2	342	52	30.35	7.64	0.85	2.93
	LEP: All Other Students	42,722	52	32.08	8.05	0.88	2.80

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
6	Students with an IEP	6,739	52	22.20	8.42	0.88	2.97
	IEP: All Other Students	37,743	52	33.48	6.94	0.84	2.75
	Economically Disadvantaged Students	18,298	52	28.70	8.53	0.88	2.92
	SES: All Other Students	26,184	52	33.92	7.32	0.86	2.72
	Migrant Students	14	52	26.64	7.44	0.85	2.85
	Migrant: All Other Students	44,468	52	31.77	8.25	0.88	2.82
	Students receiving Title 1 Services	8,478	52	28.24	8.50	0.88	2.94
	Title 1: All Other Students	36,004	52	32.60	7.96	0.88	2.78
	Plan 504	803	52	30.01	7.13	0.85	2.79
	Plan 504: All Other Students	43,679	52	31.80	8.26	0.88	2.82
7	All Students	44,427	52	33.08	9.13	0.88	3.19
	Male	23,049	52	31.34	9.12	0.88	3.15
	Female	21,318	52	34.96	8.75	0.87	3.17
	Gender Not Reported	60	52	36.70	8.48	0.87	3.11
	Hispanic or Latino	3,099	52	28.31	9.47	0.88	3.31
	American Indian or Alaskan Native	242	52	29.14	9.58	0.88	3.26
	Asian	938	52	35.24	9.38	0.89	3.17
	Black or African American	1,625	52	28.08	9.85	0.88	3.35
	Native Hawaiian or Pacific Islander	43	52	32.23	9.08	0.88	3.11
	White (non-Hispanic)	37,584	52	33.69	8.84	0.87	3.17
	Two or More Races (non-Hispanic)	787	52	32.15	9.63	0.88	3.32
	No Primary Race/Ethnicity Reported	109	52	32.24	11.30	0.91	3.30
	Currently receiving LEP services	1,107	52	21.81	9.17	0.87	3.35
	Former LEP student - monitoring year 1	124	52	32.72	7.46	0.81	3.25
	Former LEP student - monitoring year 2	198	52	33.23	8.12	0.84	3.27
	LEP: All Other Students	42,998	52	33.37	8.95	0.87	3.19
	Students with an IEP	6,847	52	22.58	8.64	0.86	3.19
	IEP: All Other Students	37,580	52	35.00	7.82	0.84	3.13
	Economically Disadvantaged Students	17,882	52	29.40	9.25	0.88	3.26
	SES: All Other Students	26,545	52	35.56	8.15	0.86	3.10
	Migrant Students	15	52	25.20	11.69	0.91	3.48
	Migrant: All Other Students	44,412	52	33.09	9.13	0.88	3.19
	Students receiving Title 1 Services	5,677	52	28.94	9.14	0.87	3.28
	Title 1: All Other Students	38,750	52	33.69	8.97	0.87	3.18
	Plan 504	918	52	31.87	7.91	0.84	3.17
	Plan 504: All Other Students	43,509	52	33.11	9.15	0.88	3.19
8	All Students	44,520	52	31.60	9.01	0.88	3.09
	Male	22,986	52	30.03	8.99	0.89	3.02
	Female	21,458	52	33.27	8.72	0.88	3.06
	Gender Not Reported	76	52	34.74	8.01	0.88	2.82
	Hispanic or Latino	2,896	52	27.09	9.14	0.88	3.23
	American Indian or Alaskan Native	262	52	27.69	8.62	0.86	3.17
	Asian	915	52	33.74	9.49	0.89	3.08
	Black or African American	1,589	52	26.90	9.52	0.88	3.28
	Native Hawaiian or Pacific Islander	35	52	29.49	10.69	0.90	3.30
	White (non-Hispanic)	37,917	52	32.13	8.79	0.88	3.06
	Two or More Races (non-Hispanic)	776	52	31.14	9.25	0.88	3.14
	No Primary Race/Ethnicity Reported	130	52	30.85	10.71	0.92	3.08

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
8	Currently receiving LEP services	1,065	52	21.27	8.28	0.84	3.31
	Former LEP student - monitoring year 1	72	52	31.07	8.57	0.86	3.21
	Former LEP student - monitoring year 2	144	52	32.78	8.28	0.86	3.13
	LEP: All Other Students	43,239	52	31.85	8.88	0.88	3.08
	Students with an IEP	6,853	52	21.76	8.32	0.86	3.13
	IEP: All Other Students	37,667	52	33.39	7.91	0.85	3.02
	Economically Disadvantaged Students	17,219	52	27.88	8.95	0.88	3.16
	SES: All Other Students	27,301	52	33.94	8.22	0.87	3.00
	Migrant Students	13	52	24.38	8.73	0.87	3.15
	Migrant: All Other Students	44,507	52	31.60	9.01	0.88	3.09
	Students receiving Title 1 Services	4,438	52	27.24	9.01	0.87	3.21
	Title 1: All Other Students	40,082	52	32.08	8.88	0.88	3.07
	Plan 504	1,063	52	30.17	8.07	0.86	3.06
	Plan 504: All Other Students	43,457	52	31.63	9.03	0.88	3.09
11	All Students	31,234	52	33.98	9.38	0.90	2.96
	Male	15,940	52	32.41	9.53	0.90	2.95
	Female	15,286	52	35.62	8.92	0.90	2.89
	Gender Not Reported	8	52				
	Hispanic or Latino	2,575	52	29.22	10.39	0.91	3.18
	American Indian or Alaskan Native	96	52	31.99	8.89	0.89	2.99
	Asian	782	52	33.93	10.56	0.92	3.00
	Black or African American	1,238	52	29.17	10.07	0.90	3.17
	Native Hawaiian or Pacific Islander	43	52	33.28	8.00	0.87	2.89
	White (non-Hispanic)	25,796	52	34.77	8.93	0.89	2.91
	Two or More Races (non-Hispanic)	502	52	33.61	8.75	0.88	2.97
	No Primary Race/Ethnicity Reported	202	52	24.99	11.99	0.93	3.10
	Currently receiving LEP services	583	52	16.20	7.73	0.83	3.23
	Former LEP student - monitoring year 1	126	52	26.33	6.80	0.77	3.27
	Former LEP student - monitoring year 2	125	52	29.82	7.03	0.81	3.03
	LEP: All Other Students	30,400	52	34.37	9.07	0.90	2.94
	Students with an IEP	4,433	52	23.41	9.46	0.89	3.11
	IEP: All Other Students	26,801	52	35.73	8.13	0.88	2.86
	Economically Disadvantaged Students	9,038	52	30.05	9.92	0.90	3.11
	SES: All Other Students	22,196	52	35.58	8.65	0.89	2.87
	Migrant Students	6	52				
	Migrant: All Other Students	31,228	52	33.98	9.38	0.90	2.96
	Students receiving Title 1 Services	2,338	52	30.00	10.22	0.90	3.17
	Title 1: All Other Students	28,896	52	34.30	9.23	0.90	2.94
	Plan 504	668	52	32.60	8.39	0.87	2.98
	Plan 504: All Other Students	30,566	52	34.01	9.40	0.90	2.95

Table O-3. 2012–13 NECAP: Subgroup Reliabilities
Writing

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
5	All Students	43,832	34	19.42	4.85	0.76	2.40
	Male	22,461	34	18.13	4.79	0.76	2.35
	Female	21,326	34	20.77	4.54	0.73	2.36
	Gender Not Reported	45	34	20.82	4.63	0.74	2.37
	Hispanic or Latino	3,274	34	17.79	5.28	0.78	2.46
	American Indian or Alaskan Native	237	34	17.26	5.02	0.77	2.43
	Asian	1,017	34	20.87	4.91	0.76	2.38
	Black or African American	1,571	34	17.95	5.08	0.77	2.44
	Native Hawaiian or Pacific Islander	47	34	20.60	4.24	0.65	2.51
	White (non-Hispanic)	36,656	34	19.61	4.75	0.75	2.39
	Two or More Races (non-Hispanic)	935	34	19.15	4.85	0.75	2.40
	No Primary Race/Ethnicity Reported	95	34	18.78	5.24	0.77	2.54
	Currently receiving LEP services	1,205	34	15.15	5.60	0.78	2.62
	Former LEP student - monitoring year 1	317	34	19.70	3.96	0.67	2.27
	Former LEP student - monitoring year 2	400	34	19.76	4.50	0.74	2.30
	LEP: All Other Students	41,910	34	19.54	4.78	0.75	2.39
	Students with an IEP	6,482	34	14.26	4.79	0.74	2.43
	IEP: All Other Students	37,350	34	20.31	4.27	0.70	2.32
	Economically Disadvantaged Students	18,222	34	17.77	4.90	0.76	2.41
	SES: All Other Students	25,610	34	20.59	4.46	0.72	2.34
	Migrant Students	17	34	17.35	4.51	0.70	2.47
	Migrant: All Other Students	43,815	34	19.42	4.85	0.76	2.40
	Students receiving Title 1 Services	9,964	34	17.72	4.87	0.75	2.41
	Title 1: All Other Students	33,868	34	19.92	4.73	0.75	2.38
	Plan 504	722	34	18.08	4.32	0.70	2.36
	Plan 504: All Other Students	43,110	34	19.44	4.86	0.76	2.40
8	All Students	44,366	34	20.99	5.35	0.76	2.61
	Male	22,890	34	19.34	5.35	0.75	2.66
	Female	21,401	34	22.75	4.76	0.74	2.44
	Gender Not Reported	75	34	24.79	4.19	0.70	2.28
	Hispanic or Latino	2,894	34	19.08	5.32	0.74	2.70
	American Indian or Alaskan Native	258	34	18.55	4.91	0.72	2.62
	Asian	916	34	22.61	5.15	0.76	2.51
	Black or African American	1,578	34	19.17	5.32	0.74	2.70
	Native Hawaiian or Pacific Islander	35	34	21.20	5.96	0.75	2.99
	White (non-Hispanic)	37,788	34	21.19	5.30	0.76	2.60
	Two or More Races (non-Hispanic)	769	34	20.76	5.36	0.75	2.65
	No Primary Race/Ethnicity Reported	128	34	21.42	7.08	0.80	3.16
	Currently receiving LEP services	1,067	34	16.35	5.36	0.72	2.83
	Former LEP student - monitoring year 1	72	34	21.25	5.52	0.77	2.62
	Former LEP student - monitoring year 2	144	34	21.69	4.27	0.69	2.40
	LEP: All Other Students	43,083	34	21.10	5.30	0.76	2.60
	Students with an IEP	6,790	34	15.19	4.95	0.69	2.78
	IEP: All Other Students	37,576	34	22.04	4.71	0.73	2.46
	Economically Disadvantaged Students	17,121	34	18.95	5.30	0.74	2.69
	SES: All Other Students	27,245	34	22.28	4.97	0.75	2.50

continued

Grade	Description	Number of Students	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
8	Migrant Students	13	34	16.38	4.48	0.56	2.96
	Migrant: All Other Students	44,353	34	20.99	5.35	0.76	2.61
	Students receiving Title 1 Services	4,423	34	18.77	5.39	0.75	2.71
	Title 1: All Other Students	39,943	34	21.24	5.29	0.76	2.60
	Plan 504	1,057	34	19.36	4.98	0.70	2.73
	Plan 504: All Other Students	43,309	34	21.03	5.36	0.76	2.61
11	All Students	31,183	12	6.18	2.02		
	Male	15,901	12	5.79	2.09		
	Female	15,274	12	6.58	1.86		
	Gender Not Reported	8	12				
	Hispanic or Latino	2,570	12	5.61	2.02		
	American Indian or Alaskan Native	97	12	5.82	1.81		
	Asian	784	12	6.50	2.09		
	Black or African American	1,238	12	5.62	1.97		
	Native Hawaiian or Pacific Islander	42	12	6.55	1.47		
	White (non-Hispanic)	25,752	12	6.26	2.00		
	Two or More Races (non-Hispanic)	500	12	6.22	1.94		
	No Primary Race/Ethnicity Reported	200	12	4.38	2.35		
	Currently receiving LEP services	579	12	3.79	2.06		
	Former LEP student - monitoring year 1	128	12	5.50	1.57		
	Former LEP student - monitoring year 2	125	12	5.61	1.55		
	LEP: All Other Students	30,351	12	6.23	1.99		
	Students with an IEP	4,409	12	4.27	1.86		
	IEP: All Other Students	26,774	12	6.49	1.87		
	Economically Disadvantaged Students	9,021	12	5.53	1.99		
	SES: All Other Students	22,162	12	6.44	1.97		
	Migrant Students	6	12				
	Migrant: All Other Students	31,177	12	6.18	2.02		
	Students receiving Title 1 Services	2,330	12	5.85	2.01		
	Title 1: All Other Students	28,853	12	6.20	2.02		
	Plan 504	667	12	5.81	1.79		
	Plan 504: All Other Students	30,516	12	6.18	2.02		

Table O-4. 2012–13 NECAP Reliabilities by Reporting Category—Mathematics

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	DP	9	10	5.97	2.48	0.75	1.25
	FA	8	10	5.65	2.55	0.65	1.50
	GM	8	10	5.38	2.14	0.58	1.38
	NO	30	35	23.50	7.02	0.88	2.39
4	DP	8	10	5.93	2.25	0.60	1.42
	FA	9	10	5.55	2.57	0.67	1.47
	GM	11	13	7.54	2.92	0.70	1.59
	NO	27	32	22.13	6.45	0.87	2.30

continued

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
5	DP	8	10	5.73	2.77	0.69	1.56
	FA	9	13	7.02	3.36	0.69	1.88
	GM	9	13	7.08	3.03	0.68	1.71
	NO	22	30	17.72	6.93	0.85	2.64
6	DP	6	10	5.14	2.61	0.61	1.63
	FA	8	13	7.15	3.27	0.70	1.78
	GM	13	17	9.22	4.00	0.71	2.16
	NO	21	26	14.29	6.53	0.87	2.39
7	DP	6	10	4.39	2.53	0.65	1.49
	FA	15	20	10.89	4.95	0.82	2.11
	GM	12	17	7.71	3.97	0.73	2.07
	NO	15	19	9.52	4.60	0.82	1.96
8	DP	6	10	5.45	2.44	0.63	1.48
	FA	20	26	15.80	6.30	0.84	2.49
	GM	13	17	7.55	3.87	0.73	2.01
	NO	9	13	6.66	3.52	0.74	1.80
11	DP	5	10	3.18	2.65	0.70	1.45
	FA	20	25	11.24	5.52	0.85	2.14
	GM	15	19	8.71	5.26	0.85	2.05
	NO	6	10	2.86	2.47	0.71	1.32

Table O-5. 2012–13 NECAP Reliabilities by Reporting Category—Reading

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
3	LA	10	16	8.22	2.97	0.69	1.66
	LI	12	18	11.26	3.43	0.73	1.80
	TI	10	16	9.85	2.76	0.64	1.64
	TL	12	18	9.62	3.74	0.77	1.79
	WV	12	18	13.33	3.34	0.71	1.81
4	LA	9	15	8.28	2.61	0.66	1.52
	LI	14	20	13.47	4.45	0.79	2.06
	TI	10	16	9.42	3.33	0.72	1.75
	TL	13	19	12.33	3.72	0.75	1.88
	WV	11	17	12.87	2.96	0.70	1.62
5	LA	10	22	9.46	3.68	0.76	1.79
	LI	14	20	12.58	3.79	0.74	1.92
	TI	12	21	11.29	3.62	0.75	1.82
	TL	12	21	10.75	3.82	0.75	1.90
	WV	10	10	7.86	2.13	0.72	1.13
6	LA	12	21	11.69	3.76	0.77	1.82
	LI	12	21	12.40	3.44	0.73	1.79
	TI	12	21	12.04	3.71	0.76	1.82
	TL	12	21	12.05	3.51	0.74	1.78
	WV	10	10	7.68	1.95	0.65	1.16

continued

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
7	LA	12	24	13.64	4.42	0.77	2.13
	LI	12	18	12.28	3.53	0.69	1.97
	TI	12	21	12.52	4.23	0.74	2.17
	TL	12	21	13.40	3.80	0.74	1.92
	WV	10	10	7.16	2.17	0.65	1.29
8	LA	14	23	12.86	4.20	0.75	2.11
	LI	10	19	11.18	3.67	0.75	1.85
	TI	12	21	10.99	4.10	0.76	2.01
	TL	12	21	13.05	3.91	0.76	1.93
	WV	10	10	7.56	2.13	0.67	1.23
11	LA	11	23	13.49	4.53	0.81	1.99
	LI	14	20	13.79	3.78	0.78	1.78
	TI	13	22	13.78	4.03	0.78	1.90
	TL	12	21	13.50	4.29	0.81	1.89
	WV	9	9	6.71	2.03	0.67	1.16

Table O-6. 2012–13 NECAP Reliabilities by Reporting Category—Writing

Grade	Reporting Category	Number of Items	Raw Score			Alpha	Standard Error
			Maximum	Mean	Standard Deviation		
5	ER	1	12	4.79	1.75		
	MC	10	10	7.93	1.96	0.66	1.14
	SR	3	12	6.69	2.23	0.77	1.07
8	ER	1	12	6.74	1.97		
	MC	10	10	6.88	2.12	0.62	1.31
	SR	3	12	7.37	2.31	0.79	1.05
11	LR	1	12	6.18	2.02		

APPENDIX P—INTERRATER AGREEMENT

Table P-1. 2012–13 Item-Level Interrater Agreement Statistics—Mathematics

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
3	119701	2	870	98.74	1.26	0.96	0.00
	119731	2	869	99.42	0.58	0.98	0.00
	119747	2	862	98.26	1.74	0.96	0.00
	119780	3	875	95.43	4.34	0.95	0.23
	119784	2	865	99.65	0.35	0.99	0.00
	119852	2	865	99.31	0.69	0.98	0.00
	119862	3	874	98.74	1.14	0.99	0.11
	119920	3	864	94.79	5.21	0.96	0.00
	124305	3	872	92.43	7.34	0.91	0.23
	139695	3	879	93.63	6.37	0.89	0.00
	145250	2	866	96.42	3.58	0.93	0.00
	145252	2	881	99.77	0.23	0.99	0.00
	145505	3	873	93.13	6.87	0.93	0.00
	168409	2	872	98.97	1.03	0.98	0.00
	168416	3	875	98.51	1.49	0.99	0.00
	203056	3	877	96.12	3.76	0.97	0.11
	203063	2	872	99.54	0.46	0.99	0.00
	203129	2	864	98.73	1.27	0.97	0.00
	203198	3	878	98.63	1.14	0.99	0.23
	215701	3	874	92.45	7.32	0.90	0.23
4	120071	2	861	98.95	1.05	0.98	0.00
	120083	3	873	93.24	5.96	0.92	0.80
	120131	2	870	99.08	0.92	0.97	0.00
	120159	2	863	98.38	1.62	0.96	0.00
	120199	2	866	98.85	1.15	0.98	0.00
	120203	3	879	84.98	14.45	0.85	0.57
	120221	2	878	98.63	1.37	0.96	0.00
	120222	3	868	98.27	1.73	0.97	0.00
	120284	3	863	89.46	9.97	0.90	0.58
	124723	3	869	92.87	6.67	0.93	0.46
	124737	2	869	98.96	1.04	0.98	0.00
	139450	3	867	90.66	9.23	0.94	0.12
	139487	3	883	97.73	2.27	0.97	0.00
	139527	3	871	98.05	1.95	0.98	0.00
	145085	2	874	99.08	0.92	0.97	0.00
	168457	2	869	98.39	1.61	0.97	0.00
	169111	2	869	99.08	0.92	0.98	0.00
	201364	3	881	94.78	4.88	0.96	0.34
	201383	2	876	98.97	1.03	0.98	0.00
	201425	3	875	89.26	9.83	0.90	0.91
5	120639	3	870	94.02	5.63	0.96	0.34
	120659	2	870	97.13	2.87	0.93	0.00
	120694	3	874	93.82	5.95	0.95	0.23
	120746	3	880	87.61	11.93	0.89	0.45
	120783	2	868	98.27	1.73	0.96	0.00
	120791	3	860	90.93	8.26	0.93	0.81

continued

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
5	120850	2	871	99.08	0.92	0.98	0.00
	120857	5	885	93.79	5.54	0.97	0.68
	124804	2	922	97.94	2.06	0.95	0.00
	144664	5	916	86.46	12.01	0.94	1.53
	145102	2	886	98.08	1.92	0.96	0.00
	145571	3	874	89.47	10.41	0.94	0.11
	145942	5	872	84.40	14.22	0.94	1.38
	167743	3	869	96.89	2.76	0.97	0.35
	167794	5	878	89.64	7.74	0.95	2.62
	167878	2	883	97.51	2.49	0.95	0.00
6	119203	3	884	96.27	3.73	0.97	0.00
	119205	3	886	93.45	5.98	0.95	0.56
	119227	2	899	97.55	2.45	0.95	0.00
	119229	3	914	89.06	10.83	0.92	0.11
	119248	2	896	97.32	2.68	0.93	0.00
	119263	2	891	97.64	2.36	0.95	0.00
	119281	3	882	97.39	2.27	0.97	0.34
	119319	2	896	97.54	2.46	0.94	0.00
	119341	3	891	83.39	16.16	0.86	0.45
	125128	3	865	95.61	4.16	0.95	0.23
	139370	2	889	99.10	0.90	0.98	0.00
	167831	5	948	81.33	17.72	0.94	0.84
	167982	5	928	82.54	16.59	0.94	0.75
	202623	2	894	99.11	0.89	0.98	0.00
	202642	5	960	78.02	15.73	0.92	6.35
	202876	5	897	82.39	16.16	0.93	1.45
7	120413	3	870	97.82	2.18	0.97	0.00
	120458	5	883	92.75	6.57	0.98	0.68
	120468	2	882	99.09	0.91	0.98	0.00
	125344	5	902	89.36	10.20	0.95	0.44
	140280	2	875	99.09	0.91	0.98	0.00
	140292	5	881	88.54	10.33	0.95	1.14
	145144	3	857	98.02	1.98	0.98	0.00
	145157	2	878	96.81	3.19	0.93	0.00
	150893	3	862	95.01	4.52	0.95	0.46
	169224	2	873	97.14	2.86	0.94	0.00
	169226	5	917	79.39	18.10	0.92	2.40
	169486	3	874	84.10	14.87	0.86	1.03
	169533	2	882	98.19	1.81	0.96	0.00
	204567	3	878	95.79	3.87	0.97	0.34
	204667	3	884	93.78	6.11	0.95	0.11
	204925	2	882	98.19	1.81	0.96	0.00
8	120876	3	894	96.53	3.36	0.96	0.11
	120932	2	895	99.11	0.89	0.97	0.00
	120936	3	880	96.48	3.52	0.98	0.00
	121002	5	929	79.98	17.55	0.93	2.37
	121016	5	914	84.46	13.89	0.94	1.64
	121031	3	890	93.93	6.07	0.95	0.00

continued

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
8	121040	2	899	89.10	10.90	0.74	0.00
	121061	2	872	97.59	2.41	0.94	0.00
	125498	2	900	98.44	1.56	0.97	0.00
	125528	5	907	87.43	11.91	0.97	0.66
	139865	2	885	97.18	2.82	0.93	0.00
	139885	3	888	95.16	4.28	0.95	0.56
	139894	5	948	80.91	18.46	0.92	0.63
	145393	3	876	95.89	3.54	0.94	0.57
	146483	3	873	92.78	7.22	0.93	0.00
	204961	2	885	98.42	1.58	0.97	0.00
11	119460	5	576	88.72	10.76	0.94	0.52
	119464	2	603	99.00	1.00	0.98	0.00
	119489	5	596	88.59	6.54	0.92	4.87
	119553	2	581	96.39	3.61	0.92	0.00
	119562	2	618	99.03	0.97	0.98	0.00
	119603	2	605	99.01	0.99	0.97	0.00
	119620	3	591	94.75	4.74	0.95	0.51
	119623	5	627	88.36	11.16	0.94	0.48
	125790	3	584	91.61	8.39	0.92	0.00
	125822	2	589	98.64	1.36	0.97	0.00
	125854	2	577	99.13	0.87	0.98	0.00
	140033	3	556	85.79	11.87	0.83	2.34
	140089	2	584	99.49	0.51	0.99	0.00
	141322	5	599	88.48	10.52	0.95	1.00
	145234	2	599	98.50	1.50	0.97	0.00
	145461	2	614	99.67	0.33	0.99	0.00
	169588	2	601	99.17	0.83	0.98	0.00
	169595	3	582	95.88	3.78	0.97	0.34
	169711	3	586	97.10	2.90	0.97	0.00
	169792	3	560	94.29	4.64	0.95	1.07
	205985	2	602	99.34	0.66	0.99	0.00
	210417	2	623	97.11	2.89	0.94	0.00

Table P-2. 2012–13 Item-Level Interrater Agreement Statistics—Reading

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
3	117801	5	900	95.33	4.33	0.96	0.33
	117802	5	897	98.66	1.34	0.99	0.00
	171335	5	914	69.47	29.32	0.81	1.09
	171554	5	911	70.25	27.88	0.74	1.65
	208096	5	888	98.65	1.35	0.99	0.00
	208103	5	905	61.99	35.25	0.80	2.76
4	118026	5	876	71.35	26.26	0.87	2.28
	118028	5	876	98.40	1.60	0.98	0.00
	171975	5	912	74.67	23.25	0.89	2.19

continued

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
4	208080	5	883	69.88	28.54	0.76	1.59
	208727	5	872	98.39	1.49	0.99	0.11
	208740	5	908	65.75	32.49	0.75	1.76
5	171848	5	883	66.59	32.39	0.84	1.02
	209380	5	876	63.58	34.47	0.74	1.60
	209383	5	880	69.43	29.43	0.79	1.14
	209558	5	873	68.61	30.93	0.85	0.46
	209634	5	871	69.35	30.08	0.69	0.57
	209640	5	880	62.84	33.41	0.72	3.52
6	172140	5	958	68.79	30.38	0.80	0.84
	208991	5	885	67.01	31.41	0.78	1.58
	208995	5	886	64.90	33.86	0.75	1.13
	209151	5	900	66.44	33.22	0.75	0.33
	209153	5	935	63.64	34.12	0.75	2.03
	209299	5	892	74.33	25.11	0.79	0.56
7	147224	5	958	60.65	36.01	0.76	3.34
	147637	5	894	68.90	29.75	0.80	1.23
	173286	5	986	61.26	37.02	0.79	2.13
	173293	5	1,062	66.76	29.94	0.88	3.11
	207649	5	909	63.59	34.87	0.77	1.54
	207654	5	916	62.12	35.48	0.80	2.29
8	172373	5	909	57.87	40.15	0.75	1.65
	207754	5	915	61.64	33.33	0.75	4.70
	207757	5	903	59.47	39.09	0.78	1.33
	207782	5	946	59.09	38.16	0.77	2.22
	207852	5	901	64.48	33.63	0.77	1.78
	207854	5	888	58.33	38.63	0.77	2.93
11	147388	5	651	60.68	36.87	0.75	2.30
	208528	5	634	59.31	37.85	0.74	2.21
	208545	5	609	58.29	36.62	0.77	5.09
	208700	5	615	57.72	39.02	0.72	2.93
	208744	5	651	64.82	33.79	0.77	1.38
	208747	5	594	55.89	41.92	0.72	2.02

Table P-3. 2012–13 Item-Level Interrater Agreement Statistics by Grade—Writing

Grade	Item	Number of		Percent		Correlation	Percent of Third Scores
		Score Categories	Responses Scored Twice	Exact	Adjacent		
5	150155	7	43,119	59.25	36.97	0.68	3.72
	150161	5	947	69.48	29.25	0.78	1.27
	150201	5	909	62.82	34.87	0.72	2.31
	150231	5	884	70.48	28.39	0.78	1.13
8	150778	5	865	65.09	33.41	0.75	1.50
	150794	5	898	68.82	30.51	0.78	0.67
	150813	7	43,340	57.06	40.02	0.74	2.73
	150830	5	906	60.26	38.30	0.70	1.32
11	211263	7	30,248	58.19	38.92	0.77	2.68

APPENDIX Q—DECISION ACCURACY AND CONSISTENCY RESULTS

Table Q-1. 2012–13 NECAP: Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Overall and Conditional on Performance Level

Content Area	*Grade*	*Overall*	*Kappa*	*Conditional on Level*			
				Substantially Below Proficient	*Partially Proficient*	*Proficient*	*Proficient with Distinction*
Mathematics	3	0.81 (0.74)	0.63	0.85 (0.79)	0.72 (0.63)	0.83 (0.77)	0.86 (0.76)
	4	0.82 (0.74)	0.63	0.83 (0.76)	0.69 (0.59)	0.82 (0.77)	0.90 (0.81)
	5	0.81 (0.74)	0.63	0.83 (0.77)	0.62 (0.51)	0.84 (0.79)	0.89 (0.80)
	6	0.81 (0.74)	0.63	0.84 (0.77)	0.62 (0.51)	0.83 (0.77)	0.89 (0.81)
	7	0.80 (0.73)	0.62	0.82 (0.76)	0.64 (0.53)	0.82 (0.76)	0.90 (0.82)
	8	0.81 (0.74)	0.63	0.85 (0.79)	0.66 (0.55)	0.83 (0.78)	0.87 (0.77)
	11	0.84 (0.78)	0.68	0.89 (0.86)	0.73 (0.63)	0.89 (0.83)	0.80 (0.59)
Reading	3	0.78 (0.70)	0.54	0.77 (0.65)	0.66 (0.55)	0.81 (0.77)	0.85 (0.71)
	4	0.78 (0.70)	0.56	0.78 (0.67)	0.69 (0.59)	0.79 (0.74)	0.87 (0.74)
	5	0.80 (0.72)	0.58	0.76 (0.61)	0.75 (0.66)	0.81 (0.75)	0.86 (0.74)
	6	0.79 (0.71)	0.55	0.76 (0.62)	0.72 (0.62)	0.81 (0.76)	0.86 (0.72)
	7	0.78 (0.70)	0.54	0.77 (0.64)	0.68 (0.58)	0.80 (0.75)	0.86 (0.73)
	8	0.80 (0.73)	0.56	0.74 (0.58)	0.69 (0.59)	0.82 (0.77)	0.87 (0.75)
	11	0.79 (0.72)	0.58	0.78 (0.65)	0.72 (0.63)	0.76 (0.70)	0.89 (0.80)
Writing	5	0.67 (0.56)	0.37	0.70 (0.53)	0.66 (0.58)	0.63 (0.54)	0.80 (0.59)
	8	0.68 (0.58)	0.37	0.69 (0.50)	0.61 (0.51)	0.70 (0.63)	0.81 (0.59)

Table Q-2. 2012–13 NECAP: Summary of Decision Accuracy (and Consistency) Results by Content Area and Grade—Conditional on Cutpoint

Content Area	*Grade*	*Substantially Below Proficient / Partially Proficient*			*Partially Proficient / Proficient*			*Proficient / Proficient with Distinction*		
		Accuracy (consistency)	*False*		*Accuracy (consistency)*	*False*		*Accuracy (consistency)*	*False*	
			Positive	*Negative*		*Positive*	*Negative*		*Positive*	*Negative*
Mathematics	3	0.96 (0.94)	0.02	0.02	0.93 (0.90)	0.04	0.03	0.93 (0.90)	0.05	0.03
	4	0.96 (0.94)	0.02	0.02	0.93 (0.90)	0.04	0.03	0.93 (0.90)	0.05	0.02
	5	0.95 (0.93)	0.03	0.03	0.93 (0.90)	0.04	0.03	0.94 (0.91)	0.04	0.02
	6	0.95 (0.93)	0.03	0.03	0.93 (0.90)	0.04	0.03	0.93 (0.90)	0.04	0.02
	7	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.03	0.94 (0.91)	0.04	0.02
	8	0.95 (0.93)	0.03	0.03	0.93 (0.90)	0.04	0.03	0.93 (0.91)	0.04	0.02
	11	0.93 (0.90)	0.04	0.03	0.94 (0.91)	0.04	0.03	0.98 (0.98)	0.01	0.00
Reading	3	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04	0.92 (0.89)	0.06	0.02
	4	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04	0.91 (0.88)	0.06	0.03
	5	0.97 (0.96)	0.01	0.02	0.91 (0.87)	0.05	0.04	0.92 (0.89)	0.05	0.03
	6	0.97 (0.95)	0.01	0.02	0.91 (0.87)	0.05	0.05	0.92 (0.89)	0.06	0.02
	7	0.96 (0.94)	0.02	0.02	0.91 (0.87)	0.05	0.04	0.91 (0.88)	0.06	0.02
	8	0.98 (0.97)	0.01	0.02	0.92 (0.89)	0.04	0.04	0.91 (0.87)	0.06	0.03
	11	0.97 (0.96)	0.01	0.02	0.92 (0.89)	0.04	0.04	0.91 (0.87)	0.06	0.03
Writing	5	0.93 (0.90)	0.03	0.04	0.84 (0.78)	0.10	0.07	0.90 (0.86)	0.08	0.02
	8	0.94 (0.92)	0.02	0.04	0.85 (0.79)	0.08	0.07	0.89 (0.85)	0.09	0.02

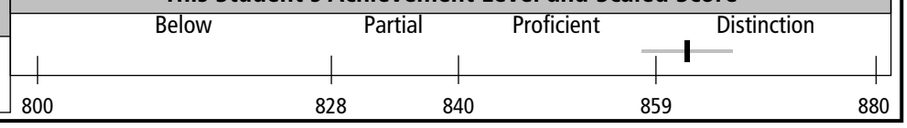
APPENDIX R—SAMPLE REPORTS

Student	Grade	School	District	State
Garrett J Swandal	8	Demonstration School 3	Demonstration District A	RI

Fall 2012 - Beginning of Grade 8 NECAP Test Results

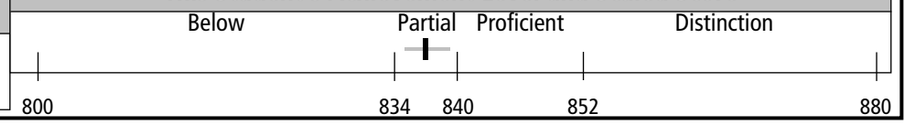
Content Area	Achievement Level	Scaled Score
Reading	Proficient with Distinction	862

This Student's Achievement Level and Scaled Score



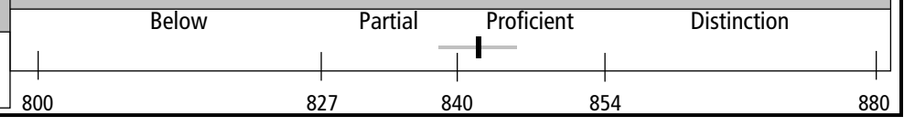
Content Area	Achievement Level	Scaled Score
Mathematics	Partially Proficient	837

This Student's Achievement Level and Scaled Score



Content Area	Achievement Level	Scaled Score
Writing	Proficient	842

This Student's Achievement Level and Scaled Score



Interpretation of Graphic Display

The line (**I**) 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 8 Students by School, District, and State

	Reading				Mathematics				Writing			
	Student	School	District	State	Student	School	District	State	Student	School	District	State
Proficient with Distinction	✓		27%	25%			16%	16%			15%	16%
Proficient			47%	52%			46%	42%	✓		50%	49%
Partially Proficient			17%	16%	✓		10%	18%			24%	26%
Substantially Below Proficient			9%	7%			28%	24%			11%	8%

This Student's Performance in Content Area Subcategories

Reading		Possible Points	Student	Average Points Earned: School	District	State	Students at Beginning of Proficient
Word ID/ Vocabulary		10	**7**		6.9	7.1	5.2-7.7
Type of Text*	Literary	21	**16**		12.7	13.2	8.7-12.5
	Informational	21	**17**		11.0	11.2	5.9-10.0
Level of Comprehension*	Initial Understanding	19	**16**		11.1	11.4	6.8-10.5
	Analysis and Interpretation	23	**17**		12.6	13.0	7.8-12.0

Mathematics	Possible Points	Student	Average Points Earned: School	District	State	Students at Beginning of Proficient
Numbers and Operations	13	**4**		6.0	6.2	3.6-7.2
Geometry and Measurement	17	**3**		7.0	7.1	3.9-7.9
Functions and Algebra	26	**15**		14.9	15.2	12.2-17.2
Data, Statistics, and Probability	10	**3**		5.1	5.2	3.5-6.4

Writing	Possible Points	Student	Average Points Earned: School	District	State	Students at Beginning of Proficient
Multiple Choice	10	**5**		6.8	6.8	5.3-7.9
Short Responses	12	**8**		7.4	7.5	5.9-8.0
Extended Response	12	**8**		6.8	6.9	6.5-6.5

*With the exception of Word ID/Vocabulary items, reading items are reported in two ways - Type of Text and Level of Comprehension.

CONFIDENTIAL

Fall 2012 - Beginning of Grade 04 NECAP Tests

Grade 04 Students in 2012-2013

Item Analysis Report — Reading

School: Demonstration School 1
District: Demonstration District A
State: Maine
Code: DEMA-DEM1

		Released Items												Total Test Results							
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12	Subcategory Points Earned							
	Content Strand	WV	WV	IA	WV	IA	II	II	WV	LI	LA	LA	LA	Word ID/Vocabulary	Literary	Informational	Initial Understanding	Analysis & Interpretation	Total Points Earned	Scaled Score	Achievement Level
	GE Code	3-3	3-3	3-8	3-1	3-8	3-7	3-7	3-2	3-4	3-5	3-5	3-5								
	Depth of Knowledge Code	1	1	2	1	2	1	2	1	2	2	2	2								
	Item Type	MC	MC	MC	MC	MC	MC	CR	MC	MC	MC	MC	CR								
	Correct MC Response	A	C	D	B	B	A		D	D	B	A									
Name/Student ID	Total Possible Points	1	1	1	1	1	1	4	1	1	1	1	4	17	19	16	20	15	52		
Abidshah, Kinzha G	D04100013	B	+	+	+	D	+	2	+	+	+	+	1	9	11	7	12	6	27	437	2
Aden, Nicholas W	D04100049	+	+	+	+	A	+	2	A	+	+	+	2	13	15	11	16	10	39	450	3
Angwin, Jordan S	D04100054	C	+	+	C	C	B	1	+	+	A	+	3	13	16	8	13	11	37	448	3
Blair, Brandon D	D04100039	D	+	+	+	+	+	2	+	+	+	+	1	13	10	8	10	8	31	441	3
Blaskovich, Kiley J	D04100011	+	+	+	+	D	D	2	+	+	A	+	1	15	12	6	11	7	33	443	3
Brooks, Colton	D04100007	+	+	+	+	+	+	3	+	+	+	+	1	17	15	13	18	10	45	462	4
Brown, Shannon K	D04100063	+	+	C	C	A	+	2	B	+	+	+	2	10	13	6	11	8	29	439	2
Cabrera, Jeffrey	D04100060	+	+	+	+	+	D	3	+	+	+	+	1	15	12	9	12	9	36	446	3
Calhoun, Roosevelt P	D04100031	+	+	+	+	+	+	2	A	+	C	B	1	11	10	12	15	7	33	443	3
Campbell, Andrea J	D04100016	C	A	+	C	C	D	2	A	B	C	+	0	9	4	4	5	3	17	425	1
Fowler, Shane	D04100015	+	+	+	+	D	B	0	+	+	A	+	2	12	9	5	7	7	26	436	2
French, April M	D04100023	+	+	+	+	A	+	3	+	+	+	+	1	16	16	12	19	9	44	459	4
Kelley, Brett B	D04100037	+	+	+	+	+	+	2	+	+	+	+	2	12	14	11	15	10	37	448	3
King, Wesley R	D04100068	+	+	+	C	+	C	2	+	+	+		1	13	13	8	13	8	34	444	3
Lefrancois, Kaleb L	D04100030	D	+	+	+	+	+	2	+	+	+	+	2	11	14	11	14	11	36	446	3
Malik, Doren W	D04100010	+	A	+	+	C	+	0	A	+	+	+	1	9	9	5	7	7	23		A
Mclaren, Andrew J	D04100019	+	+	+	+	+	+	3	+	+	+	+	1	15	15	13	18	10	43	457	4
Muff, Michael	D04100014	C	A	B	+	+	+	2	+	+	+	+	0	8	9	8	11	6	25		L
Munsey, Joseph S	D04100034	C	+	A	+	D	C	2	+	+	D	+	1	9	10	5	9	6	24	433	2
Murrell, Abigail M	D04100055	+	D	+	+	+	+	2	+	+	+	+	2	15	14	10	14	10	39	450	3
Nathans, Gage B	D04100053	+	+	+	+	A	+	0	+	+	+	+	2	14	10	8	10	8	32	442	3
Oliva, Joseph A	D04100035	C	+	+	C	C	+	1	A	B	+	B	0	8	3	5	4	4	16	423	1
Oquendo, April M	D04100027	+	+	+	+	+	+	3	+	+	A	+	3	13	14	10	15	9	37	448	3
Otash, Wayne J	D04100050	+	+	+	+	+	+	3	+	+	+	+	2	15	16	13	18	11	44	459	4
Pelcastremartinez, Jennif	D04100062	+	+	B	C	C	B		+	+	+	+		10	8	3	8	3	21	430	1
Pellegri, Matthew M	D04100052	B	+	A	+	A	+		+	+	+	+	1	12	9	5	8	6	26	436	2
Porter, Travis E	D04100043		+	+	C	C	+		A	A	D	+		1	3	4	5	2	8	404	1
Schaeffer, Matthew	D04100064	+	+	+	+	C	+	3	+	+	+	+	2	13	17	11	17	11	41	453	3
Simpson, Justina C	D04100009	C	+	B	C	A	D	1	+	B	+	+	1	6	4	3	3	4	13	418	1
Skaw, Matthew E	D04100002	C	+	B	A	D	+	0	A	C	+	B	0	5	8	3	7	4	16	423	1

CONFIDENTIAL

Fall 2012 - Beginning of Grade 04 NECAP Tests

Grade 04 Students in 2012-2013

Item Analysis Report — Reading

School: Demonstration School 1
District: Demonstration District A
State: Maine
Code: DEMA-DEM1

		Released Items																	Total Test Results							
	Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12						Subcategory Points Earned					Total Points Earned	Scaled Score	Achievement Level
	Content Strand	WV	WV	IA	WV	IA	II	II	WV	LI	LA	LA	LA						Word ID/Vocabulary	Literary	Informational	Initial Understanding	Analysis & Interpretation			
	GE Code	3-3	3-3	3-8	3-1	3-8	3-7	3-7	3-2	3-4	3-5	3-5	3-5													
	Depth of Knowledge Code	1	1	2	1	2	1	2	1	2	2	2	2													
	Item Type	MC	MC	MC	MC	MC	MC	CR	MC	MC	MC	MC	CR													
	Correct MC Response	A	C	D	B	B	A		D	D	B	A														
Name/Student ID	Total Possible Points	1	1	1	1	1	1	4	1	1	1	1	4						17	19	16	20	15	52		
Travaglini, Alexa L	D04100047	+	+	+	+	+	+	3	C	+	A	+	1						13	13	13	17	9	39	450	3
Tucker, Brandon M	D04100057	+	+	+	+	+	+	1	+	+	+	+	1						13	15	7	14	8	35	445	3
Vaz, Aaron	D04100051	+	+	+	+	A	+	2	+	+	+	+	2						14	9	10	11	8	33	443	3
Willey, Shane R	D04100028	B	+	+	+	+	+	0	+	+	+	+	2						12	15	9	13	11	36	446	3

Released Item Number	1	2	3	4	5	6	7	8	9	10	11	12										
Percent Correct/Average Score: Group	63	94	81	72	44	72	1.7	75	84	72	88	1.3						11.8	11.5	8.2	11.9	7.8
Percent Correct/Average Score: School	63	94	81	72	44	72	1.7	75	84	72	88	1.3						11.8	11.5	8.2	11.9	7.8
Percent Correct/Average Score: District	65	90	79	77	42	77	1.9	73	85	66	90	1.3						12.2	12.2	8.6	12.9	7.9
Percent Correct/Average Score: State	71	87	80	84	56	79	2.2	72	88	66	85	1.5						12.6	12.1	9.2	13.2	8.1

About The New England Common Assessment Program

This report highlights results from the Fall 2012 Beginning of Grade New England Common Assessment Program (NECAP) tests. The NECAP tests are administered to students in Maine, 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 four 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 that 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 schools' and districts' evaluation of their ongoing improvement efforts.

Fall 2012
Beginning of Grade 5
NECAP Tests

Grade 4 Students in 2011-2012

School Results

School: Demonstration School 1

District: Demonstration District A

Code: DEMOA-DEMO1

Fall 2012 - Beginning of Grade 5 NECAP Tests
Grade 4 Students in 2011-2012
Grade Level Summary Report

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

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																		
With an approved accommodation																		
Current LEP Students																		
With an approved accommodation																		
IEP Students																		
With an approved accommodation																		
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 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				28	3	11	15	54	4	14	6	21	541	54	13	48	19	20	541	6,100	19	53	19	10	546
MATH				27	3	11	13	48	5	19	6	22	541	53	15	40	23	23	541	6,103	21	44	16	19	544
WRITING				26	0	0	14	54	5	19	7	27	538	52	4	46	25	25	538	6,086	14	37	35	14	540

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.
Note: Some numbers may have been left blank because fewer than ten (10) students were tested.

Fall 2012 - Beginning of Grade 5 NECAP Tests

Grade 4 Students in 2011-2012

Reading Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

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.
(Scaled Score 556–580)

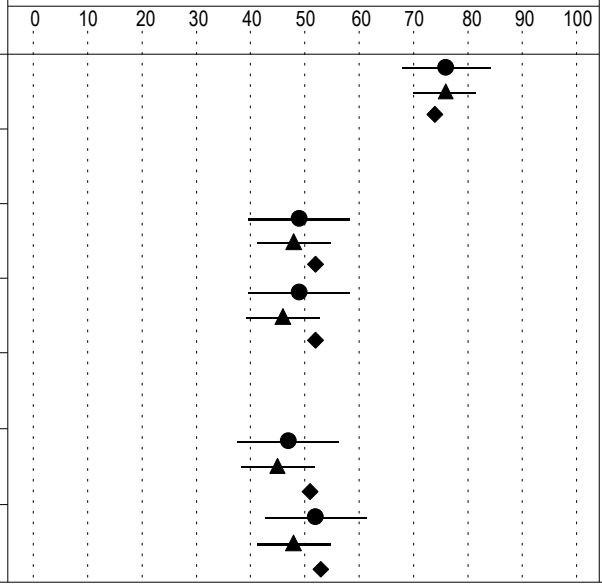
Proficient (Level 3)
Student's performance demonstrates an ability to read and comprehend grade-appropriate text. Student is able to analyze and interpret literary and informational text. Student makes and supports relevant assertions by referencing text. Student uses vocabulary strategies and breadth of vocabulary knowledge to read and comprehend text.
(Scaled Score 540–555)

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.
(Scaled Score 530–539)

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.
(Scaled Score 500–529)

	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													
2010-11				23	2	9	15	65	3	13	3	13	544
2011-12				25	1	4	8	32	9	36	7	28	535
2012-13				**28**	**3**	**11**	**15**	**54**	**4**	**14**	**6**	**21**	**541**
Cumulative Total				76	6	8	38	50	16	21	16	21	540
District													
2010-11				48	6	13	26	54	9	19	7	15	542
2011-12				51	3	6	16	31	21	41	11	22	537
2012-13				**54**	**7**	**13**	**26**	**48**	**10**	**19**	**11**	**20**	**541**
Cumulative Total				153	16	10	68	44	40	26	29	19	540
State													
2010-11				6,190	1,251	20	3,236	52	1,137	18	566	9	546
2011-12				6,120	1,152	19	3,119	51	1,195	20	654	11	545
2012-13				**6,100**	**1,135**	**19**	**3,237**	**53**	**1,137**	**19**	**591**	**10**	**546**
Cumulative Total				18,410	3,538	19	9,592	52	3,469	19	1,811	10	546

Subtopic	Total Possible Points
Word ID/Vocabulary	25
Type of Text	
Literary	49
Informational	56
Level of Comprehension	
Initial Understanding	52
Analysis & Interpretation	53



Fall 2012 - Beginning of Grade 5 NECAP Tests
Grade 4 Students in 2011-2012
Disaggregated Reading Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

REPORTING CATEGORIES	School													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				**28**	**3**	**11**	**15**	**54**	**4**	**14**	**6**	**21**	**541**	**54**	**13**	**48**	**19**	**20**	**541**	**6,100**	**19**	**53**	**19**	**10**	**546**
Gender																									
Male				10	0	0	5	50	3	30	2	20	540	22	9	41	32	18	541	3,087	13	54	21	12	544
Female				17	3	18	10	59	0	0	4	24	543	31	16	55	6	23	542	3,012	24	52	17	7	548
Not Reported				1										1						1					
Race/Ethnicity																									
Hispanic or Latino				0										0						89	20	58	17	4	548
Not Hispanic or Latino																									
American Indian or Alaskan Native				0										1						14	0	36	43	21	535
Asian				1										1						107	22	41	21	15	545
Black or African American				0										1						108	11	45	24	19	542
Native Hawaiian or Pacific Islander				1										1						5					
White				23	3	13	14	61	2	9	4	17	545	44	14	55	14	18	543	5,577	19	54	18	9	547
Two or more races				1										2						159	16	52	19	14	545
No Race/Ethnicity Reported				2										4						41	15	37	27	22	541
LEP Status																									
Current LEP student				1										1						81	4	32	35	30	536
Former LEP student - monitoring year 1				1										1						59	19	61	20	0	549
Former LEP student - monitoring year 2				1										1						18	39	44	17	0	553
All Other Students				25	2	8	14	56	4	16	5	20	542	51	12	49	20	20	542	5,942	19	53	18	10	546
IEP																									
Students with an IEP				5										10	0	10	30	60	527	866	1	22	33	43	532
All Other Students				23	3	13	15	65	3	13	2	9	545	44	16	57	16	11	545	5,234	22	58	16	4	549
SES																									
Economically Disadvantaged Students				11	2	18	3	27	2	18	4	36	536	21	14	29	29	29	537	2,651	10	48	25	16	542
All Other Students				17	1	6	12	71	2	12	2	12	545	33	12	61	12	15	544	3,449	25	57	13	5	550
Migrant																									
Migrant Students				1										1						6					
All Other Students				27	3	11	14	52	4	15	6	22	541	53	13	47	19	21	541	6,094	19	53	19	10	546

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 2012 - Beginning of Grade 5 NECAP Tests

Grade 4 Students in 2011-2012

Mathematics Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

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.
(Scaled Score 554–580)

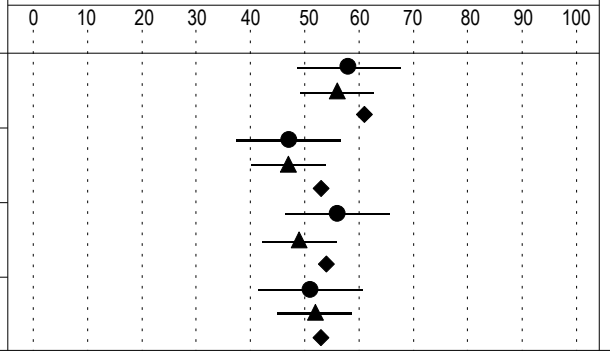
Proficient (Level 3)
Student's problem solving demonstrates logical reasoning with appropriate explanations that include both words and proper mathematical notation. Student uses a variety of strategies that are often systematic. Computational errors do not interfere with communicating understanding. Student demonstrates conceptual understanding of most aspects of the grade level expectations.
(Scaled Score 540–553)

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.
(Scaled Score 533–539)

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.
(Scaled Score 500–532)

	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													
2010-11				22	2	9	13	59	1	5	6	27	542
2011-12				25	1	4	8	32	5	20	11	44	532
2012-13				**27**	**3**	**11**	**13**	**48**	**5**	**19**	**6**	**22**	**541**
Cumulative Total				74	6	8	34	46	11	15	23	31	538
District													
2010-11				50	2	4	23	46	9	18	16	32	537
2011-12				49	5	10	17	35	7	14	20	41	535
2012-13				**53**	**8**	**15**	**21**	**40**	**12**	**23**	**12**	**23**	**541**
Cumulative Total				152	15	10	61	40	28	18	48	32	538
State													
2010-11				6,191	1,221	20	2,832	46	1,001	16	1,137	18	544
2011-12				6,064	1,175	19	2,814	46	994	16	1,081	18	544
2012-13				**6,103**	**1,302**	**21**	**2,675**	**44**	**977**	**16**	**1,149**	**19**	**544**
Cumulative Total				18,358	3,698	20	8,321	45	2,972	16	3,367	18	544

Subtopic	Total Possible Points
Numbers & Operations	73
Geometry & Measurement	33
Functions & Algebra	31
Data, Statistics, & Probability	25



Fall 2012 - Beginning of Grade 5 NECAP Tests
Grade 4 Students in 2011-2012
Disaggregated Mathematics Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

REPORTING CATEGORIES	School													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				**27**	**3**	**11**	**13**	**48**	**5**	**19**	**6**	**22**	**541**	**53**	**15**	**40**	**23**	**23**	**541**	**6,103**	**21**	**44**	**16**	**19**	**544**
Gender																									
Male				11	1	9	4	36	2	18	4	36	537	23	22	30	22	26	540	3,090	20	43	16	21	543
Female				15	2	13	9	60	3	20	1	7	545	29	10	48	24	17	542	3,012	22	45	16	17	545
Not Reported				1										1						1					
Race/Ethnicity																									
Hispanic or Latino				0										0						88	19	50	15	16	544
Not Hispanic or Latino																									
American Indian or Alaskan Native				0										1						14	0	29	7	64	530
Asian				2										2						110	28	39	15	17	546
Black or African American				0										1						108	10	39	19	32	538
Native Hawaiian or Pacific Islander				1										1						5					
White				21	3	14	11	52	3	14	4	19	543	42	19	43	24	14	544	5,577	22	44	16	18	544
Two or more races				1										2						160	16	33	23	28	540
No Race/Ethnicity Reported				2										4						41	10	37	17	37	536
LEP Status																									
Current LEP student				2										2						81	2	31	27	40	535
Former LEP student - monitoring year 1				1										1						59	19	59	12	10	545
Former LEP student - monitoring year 2				1										1						18	44	44	11	0	552
All Other Students				23	3	13	12	52	2	9	6	26	541	49	16	41	18	24	541	5,945	22	44	16	19	544
IEP																									
Students with an IEP				4										9						867	2	18	21	59	531
All Other Students				23	3	13	13	57	5	22	2	9	545	44	18	48	23	11	545	5,236	25	48	15	12	546
SES																									
Economically Disadvantaged Students				11	1	9	1	9	5	45	4	36	534	21	14	14	38	33	535	2,653	11	40	20	28	540
All Other Students				16	2	13	12	75	0	0	2	13	546	32	16	56	13	16	545	3,450	29	46	13	11	547
Migrant																									
Migrant Students				1										1						6					
All Other Students				26	3	12	13	50	4	15	6	23	541	52	15	40	21	23	541	6,097	21	44	16	19	544

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 2012 - Beginning of Grade 5 NECAP Tests

Grade 4 Students in 2011-2012

Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

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.
(Scaled Score 555–580)

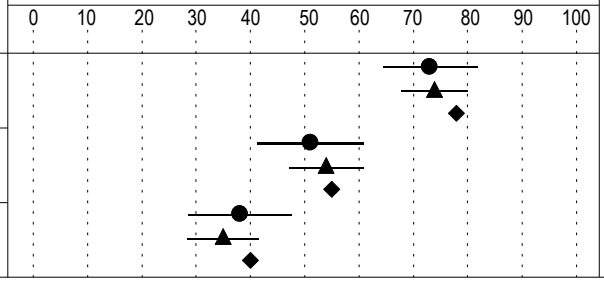
Proficient (Level 3)
Student's writing demonstrates an ability to respond to prompt/task. Focus is clear and maintained throughout the response. Response is organized with a beginning, middle, and end with appropriate transitions. Details are sufficiently elaborated to support focus. Sentence structures and language use are varied. Response demonstrates control of conventions; errors may occur but do not interfere with meaning.
(Scaled Score 540–554)

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.
(Scaled Score 527–539)

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.
(Scaled Score 500–526)

	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													
2010-11				22	3	14	6	27	10	45	3	14	538
2011-12				25	0	0	5	20	11	44	9	36	529
2012-13				**26**	**0**	**0**	**14**	**54**	**5**	**19**	**7**	**27**	**538**
Cumulative Total				73	3	4	25	34	26	36	19	26	535
District													
2010-11				46	5	11	12	26	21	46	8	17	536
2011-12				47	1	2	11	23	22	47	13	28	531
2012-13				**52**	**2**	**4**	**24**	**46**	**13**	**25**	**13**	**25**	**538**
Cumulative Total				145	8	6	47	32	56	39	34	23	535
State													
2010-11				6,175	767	12	2,376	38	2,419	39	613	10	541
2011-12				5,959	503	8	2,249	38	2,328	39	879	15	539
2012-13				**6,086**	**840**	**14**	**2,246**	**37**	**2,154**	**35**	**846**	**14**	**540**
Cumulative Total				18,220	2,110	12	6,871	38	6,901	38	2,338	13	540

Subtopic	Total Possible Points
Multiple Choice	10
Short Responses	12
Extended Response	12



Fall 2012 - Beginning of Grade 5 NECAP Tests
Grade 4 Students in 2011-2012
Disaggregated Writing Results

School: Demonstration School 1
District: Demonstration District A
State: Vermont
Code: DEMOA-DEMO1

REPORTING CATEGORIES	School													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				**26**	**0**	**0**	**14**	**54**	**5**	**19**	**7**	**27**	**538**	**52**	**4**	**46**	**25**	**25**	**538**	**6,086**	**14**	**37**	**35**	**14**	**540**
Gender																									
Male				10	0	0	3	30	2	20	5	50	531	22	0	32	32	36	533	3,075	8	31	42	20	537
Female				15	0	0	11	73	2	13	2	13	542	29	7	59	17	17	541	3,010	20	43	29	7	544
Not Reported				1										1						1					
Race/Ethnicity																									
Hispanic or Latino				0										0						89	12	44	34	10	541
Not Hispanic or Latino																									
American Indian or Alaskan Native				0										1						14	0	21	36	43	528
Asian				1										1						107	20	34	35	12	541
Black or African American				0										1						108	8	32	35	24	536
Native Hawaiian or Pacific Islander				1										1						5					
White				21	0	0	13	62	3	14	5	24	540	42	5	55	19	21	540	5,563	14	37	35	13	541
Two or more races				1										2						159	10	35	35	20	538
No Race/Ethnicity Reported				2										4						41	5	20	49	27	533
LEP Status																									
Current LEP student				1										1						80	5	25	38	33	532
Former LEP student - monitoring year 1				1										1						59	14	47	36	3	543
Former LEP student - monitoring year 2				1										1						18	50	33	17	0	553
All Other Students				23	0	0	13	57	4	17	6	26	538	49	4	47	24	24	538	5,929	14	37	35	14	540
IEP																									
Students with an IEP				4										9						861	1	7	38	54	526
All Other Students				22	0	0	14	64	5	23	3	14	541	43	5	56	26	14	541	5,225	16	42	35	7	543
SES																									
Economically Disadvantaged Students				10	0	0	3	30	2	20	5	50	530	20	5	30	30	35	534	2,646	7	30	42	21	536
All Other Students				16	0	0	11	69	3	19	2	13	542	32	3	56	22	19	540	3,440	19	42	31	8	544
Migrant																									
Migrant Students				1										1						6					
All Other Students				25	0	0	14	56	4	16	7	28	538	51	4	47	24	25	538	6,080	14	37	35	14	540

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 2012 NECAP Tests
School Summary
2011-2012 Students

School: Demonstration School 1
District: Demonstration District A
State: New Hampshire
Code: DEM-DEA-DEMO1

Reading	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
	N	N	N	N	Level 4		Level 3		Level 2		Level 1		Mean Score
					N	%	N	%	N	%	N	%	
Demonstration School 1				**305**	**66**	**22**	**160**	**52**	**48**	**16**	**31**	**10**	
Beginning of Grade 3				41	4	10	24	59	6	15	7	17	343
Beginning of Grade 4				40	7	18	21	53	10	25	2	5	446
Beginning of Grade 5				48	12	25	24	50	7	15	5	10	547
Beginning of Grade 6				38	7	18	21	55	7	18	3	8	649
Beginning of Grade 7				51	10	20	27	53	6	12	8	16	746
Beginning of Grade 8				45	12	27	26	58	6	13	1	2	851
Beginning of Grade 11				42	14	33	17	40	6	14	5	12	1147

Mathematics	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
	N	N	N	N	Level 4		Level 3		Level 2		Level 1		Mean Score
					N	%	N	%	N	%	N	%	
Demonstration School 1				**306**	**68**	**22**	**125**	**41**	**47**	**15**	**66**	**22**	
Beginning of Grade 3				40	7	18	19	48	8	20	6	15	343
Beginning of Grade 4				41	14	34	14	34	5	12	8	20	446
Beginning of Grade 5				48	13	27	20	42	6	13	9	19	547
Beginning of Grade 6				38	11	29	18	47	4	11	5	13	647
Beginning of Grade 7				51	14	27	16	31	10	20	11	22	744
Beginning of Grade 8				45	9	20	22	49	8	18	6	13	843
Beginning of Grade 11				43	0	0	16	37	6	14	21	49	1134

Writing	Enrolled	NT Approved	NT Other	Tested	Achievement Level								
	N	N	N	N	Level 4		Level 3		Level 2		Level 1		Mean Score
					N	%	N	%	N	%	N	%	
Demonstration School 1				**134**	**12**	**9**	**53**	**40**	**57**	**43**	**12**	**9**	
Beginning of Grade 5				47	7	15	17	36	19	40	4	9	541
Beginning of Grade 8				44	5	11	24	55	13	30	2	5	842
Beginning of Grade 11				43	0	0	12	28	25	58	6	14	5.5

Level 4 = Proficient with Distinction; Level 3 = Proficient; Level 2 = Partially Proficient; Level 1 = Substantially Below Proficient

Achievement Level Summary

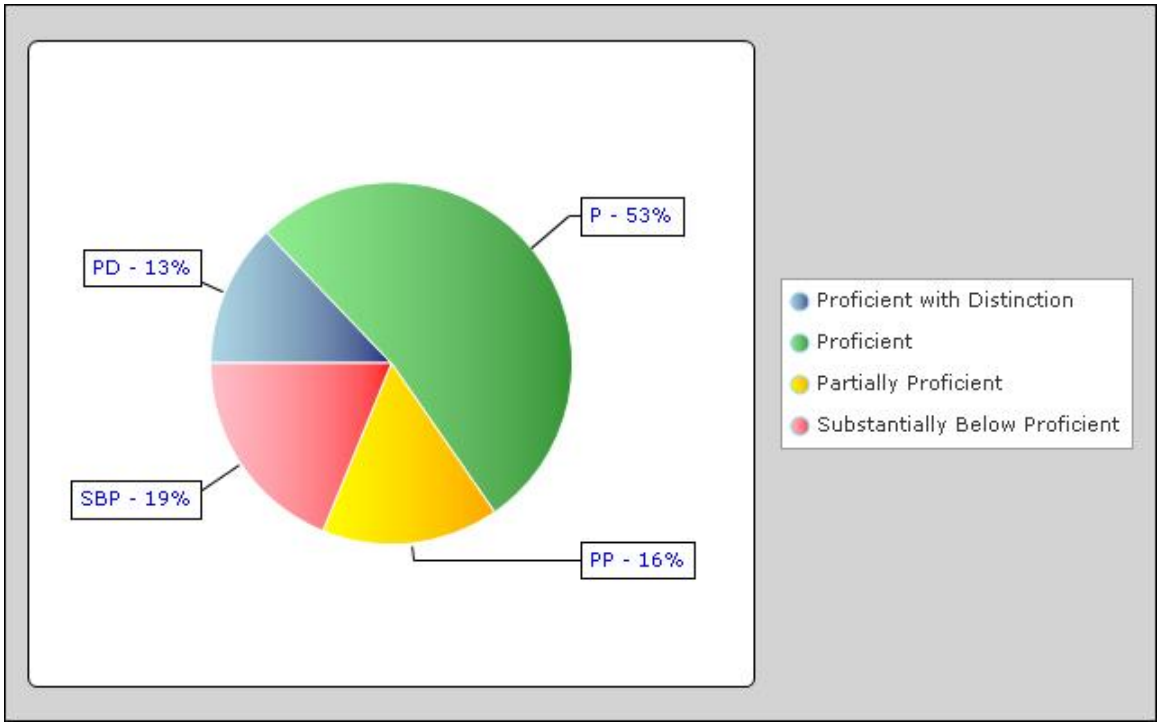
District: Demonstration District A

School: Demonstration School 1

Grade: 04

Date: 2/28/2013 10:10:14 AM

Reading



Achievement Level	Count	Percentage %*
Proficient with Distinction	4	13
Proficient	17	53
Partially Proficient	5	16
Substantially Below Proficient	6	19

*Percentages may not total exactly 100% due to applied rounding.

Reading Released Items Summary Data

District: Demonstration District A
School: Demonstration School 1
Grade: 04
Date: 2/28/2013 10:10:33 AM

Multiple Choice

Released Item	Content Strand	GE Code	Correct (#)	A (#)	B (#)	C (#)	D (#)	IR (#)	Correct Response
1	WV	3-3	20	20	3	6	2	1	A
2	WV	3-3	30	1	0	30	1	0	C
3	IA	3-8	26	2	3	1	26	0	D
4	WV	3-1	23	1	23	8	0	0	B
5	IA	3-8	14	7	14	6	5	0	B
6	II	3-7	23	23	3	2	4	0	A
8	WV	3-2	24	6	1	1	24	0	D
9	LI	3-4	27	1	3	1	27	0	D
10	LA	3-5	23	5	23	2	2	0	B
11	LA	3-5	28	28	3	0	0	1	A

Open Response

Released Item	Content Strand	GE Code	Point Value	Average Score
7	II	3-7	4	1.7
12	LA	3-5	4	1.3

CONFIDENTIAL

Longitudinal Data Report

Student Name
Ashley Anzuoni

Year	Enrolled Grade	School Name	Administration	Test Name	Content Area	Score	Achievement Level
1011	06	Demonstration School 1	NECAP Fall 2010	Grade 06 Mathematics	mat	613	Substantially Below Proficient
1011	06	Demonstration School 1	NECAP Fall 2010	Grade 06 Reading	rea	627	Substantially Below Proficient
1112	06	Demonstration School 1	NECAP Fall 2011	Grade 06 Mathematics	mat	646	Proficient
1112	06	Demonstration School 1	NECAP Fall 2011	Grade 06 Reading	rea	655	Proficient
1213	06	Demonstration School 1	NECAP Fall 2012	Grade 06 Mathematics	mat	640	Proficient
1213	06	Demonstration School 1	NECAP Fall 2012	Grade 06 Reading	rea	658	Proficient

Note: This report returns as many years of NECAP data as are available for this student beginning with 08-09.

APPENDIX S—DECISION RULES

Data Analysis and Static Reporting Decision Rules
NECAP
Fall 12-13 Administration

This document specifies rules for data analysis and static reporting requirements. 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 Departments 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

NECAP is administered in the fall and spring. This document incorporates fall and spring rules so that changes are carried to future administrations. In the fall, students are reported based on the current year fall school /district (referred to as testing school/district) and prior year spring school/district (referred to as teaching school/district). In the spring, students are reported based on the spring school/district (referred to as testing school/district). In the spring, students are not reported based on the teaching school. Rules pertaining to the teaching school/district can be ignored for spring administrations. For more information regarding discode, schcode, sprdiscode, sprschcode, senddiscode, and sprsenddiscode, please refer to the data processing specifications and demographic data specification.

This document is the official rules for the current reporting administration.

A. Fall Tests Administered:

Grade	Subject	Test Type	Test items used for Scaling
03	Reading	Operational	Common
03	Math	Operational	Common
04	Reading	Operational	Common
04	Math	Operational	Common
05	Reading	Operational	Common
05	Math	Operational	Common
05	Writing	Operational	Common
06	Reading	Operational	Common
06	Math	Operational	Common
07	Reading	Operational	Common
07	Math	Operational	Common
08	Reading	Operational	Common
08	Math	Operational	Common
08	Writing	Operational	Common
11	Reading	Operational	Common
11	Math	Operational	Common
11	Writing	Operational	Common

B. Spring Tests Administered

Grade	Subject	Test items used for Scaling	Item Reporting Categories (Subtopic and Subcategory Source)
04	Science	Common	Cat3
08	Science	Common	Cat3
11	Science	Common	Cat3

C. Reports Produced:

1. Student Report

   a. Testing School District

I. Parent Copy

II. School Copy

2. Interactive Reporting (Only the data analysis requirements are outlined in this document)
   a. Item Analysis
   b. Achievement Level Summary
   c. Item Information
   d. Student Longitudinal
3. Grade Level School/District/State Results
   a. Testing School District
   b. Teaching School District – District and School Levels only (Fall Only)
4. School/District/State Summary (School Level is produced in the Fall Only)
   a. Testing School District
   b. Teaching School District – District and School Levels only (Fall Only)
5. Writing Prompt CDs

D. Files Produced*:*

1. Preliminary State Results
2. State Student Released Item Data
3. State Student Raw Data
4. State Student Scored Data
5. District Student Data
6. School Student Data
7. Common Item Information
8. Grade Level Results Report Disaggregated and Historical Data
9. State Standard Deviations and Average Scaled Scores
10. Grade Level Results Report Participation Category Data
11. Grade Level Results Report Subtopic Data
12. Summary Results Data
13. Released Item Percent Responses Data
14. Invalidated Students Original Score
15. Student Questionnaire Summary
16. TCTA Questionnaire Raw Data
17. TCTA Questionnaire Frequency Distribution
18. Scaled Score Lookup
19. Grade 11Writing Score Distribution Results
20. Grade 11Writing Historical Subtopic Results
21. Subtopic Average Points Earned (For Program Management)
22. Item Stats for Inquiry Task Items (For Program Management)

23. Memo Shipping files (For Program Management)
24. Released Item Info File (For Program Management)
25. CD print file.

E. School Type*:*

Testing School Type: SchType **Teaching School Type: sprSchType (Fall Only)**	**Source: ICORE SubTypeID**	**Description**	**States**
PUB	1,12,13	Public School	ME, NH, RI, VT
CHA	11	Charter School	NH, RI, ME
PSP	19	Public Special Purpose	ME
PSE	15	Public Special Education	ME
INS	7	Institution	VT
OTH	9	Other	VT
OOD	4	Out-of-District Private Providers	NH
OUT	8	Out Placement	RI
PSN	23	Private Special Purpose	ME
BIG	6	Private with >60% Publicly Funded	ME
PRI	3	Private School	RI, VT

School Type Impact on Data Analysis and Reporting				
Level	**Testing**		**Teaching (Fall Only)**	
	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 BIG, PSN, PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	n/a	n/a
School	Do not exclude any students based on school type 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 BIG, PSN, PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.	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 BIG, PSN, PRI, OOD, OUT, INS, or OTH schools. Always print tested year state data.
District	For OUT, OOD,	Generate a report for each	For OUT, OOD,	Generate a report for each

	BIG, and PSN schools, aggregate using the sending district. If OUT, OOD, BIG, or PSN student does not have a sending district, do not include in aggregations. Do not include students tested at PRI, INS, or OTH schools	district with at least one student enrolled using the tested district aggregate denominator. Always report tested year state data.	BIG, and PSN teaching schools, aggregate using the spring sending district. If OUT, OOD, BIG, or PSN teaching school student does not have a teaching sending district, do not include in aggregations. Do not include students taught at PRI, INS, or OTH schools	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 and ME.	Always report testing year state data.	n/a	n/a

F. Student Status

StuStatus	Description
1	Homeschooled
2	Privately Funded
3	Exchange Student
4	Excluded State
0	Publicly Funded

StuStatus impact on Data Analysis and Reporting		
Level	**Impact on Analysis**	**Impact on Reporting**
Student	n/a	School and District data will be blank for students with a StuStatus value of 1,2 or 3. Always print tested year state data. For StuStatus values of 1, 2, and 3 print the description from the table above for the school and district names.
School	Exclude all students with a StuStatus value of 1, 2 or 3.	Students with a StuStatus value of 1, 2 or 3 are excluded from Interactive Reporting.
District	Exclude all students with a StuStatus value of 1, 2 or 3.	n/a
State	Exclude all students with a StuStatus value of 1, 2, 3 or 4.	n/a.

G. 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
Content Area Cumulative Total Enrollment, Not tested, Tested, Number and Percent at each achievement level, mean score	Suppress all cumulative total data if at least one reported year has fewer than 10 tested students. The reported years are 1011 , 1112, and 1213.

H. 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.

I. Other Information

1. NH, RI, and VT participate in NECAP testing for Grades 03-08 and 11. ME only participates in NECAP testing for Grades 03-08.
2. Grade 12 students are allowed to participate in the NECAP Grade 11 test under the following circumstances: RI students trying to improve prior NECAP score, and NH, RI, and VT students taking the NECAP Grade 11 test for the first time.
   a. RI students trying to improve are identified as StuGrade=12 and Grade=11. They only receive a student report. They are not listed on a roster or included in any aggregations. Do not print tested school and district aggregate data on the student report.
   b. For students taking NECAP for the first time the StuGrade in the student demographics file will be 11 and the remaining decision rules apply.
3. Plan504 data not available for NH and VT; therefore 504 Plan section will be suppressed for NH and VT.
4. To calculate Title1 data for writing using Title1rea variable.
5. Title 1 data are not available for VT; therefore Title 1 section will be suppressed for VT.
6. Title 1 Science data are not available for NH; therefore, Title 1 section will be suppressed for NH on Science specific reports. Title 1 Reading and Math data are available for NH and should not be suppressed.
7. Testing level is defined by the variables discode and schcode. Teaching level is defined by the variables sprdiscode and sprschcode. Every student will have testing district and school codes. In the fall, some students will have a teaching school code and some students will have a teaching district code. In the spring, no students will have a teaching school/district.
8. A non-public district code is a district code associated with a school that is type BIG, PSN, PRI, OOD, OUT, INS, or OTH. Non-public testing sending district codes will be ignored. . For example: For RI, senddiscode of 88 is ignored. For NH, senddiscode of 000 is ignored.

9. Only students with a testing school type of OUT, OOD, BIG, or PSN are allowed to have a testing sending district code. Testing sending district codes will be blanked for students at any other testing school types.

10. Only students with a teaching school type of OUT, OOD, BIG, or PSN are allowed to have a spring sending district code. Spring sending district codes will be blanked for students at any other teaching school types.

11. If students have a teaching district code and no teaching school, then ignore teaching district codes that are associated with schools that are BIG, PSN, PRI, OOD, OUT, INS, or OTH.

II. 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

   a. If a student links to the demographic file has content area not tested status of "Not Tested State Approved Alternate Assessment" is identified as "Not Tested State Approved Alternate Assessment" for the content area.

   b. If a student is identified as receiving an alternate assessment achievement level, then the student's record will be updated as outlined in the NECAP1213StudentDemographicFileDescription.doc.

2. Not Tested State Approved First Year LEP (reading and writing only)

   a. If a student links to the demographic file has content area not tested status of "Not Tested State Approved First Year LEP" or does not link to the demographic file has content area "First Year LEP blank or partially blank reason" marked, then the student is identified as "Not Tested State Approved First Year LEP".

3. Not Tested State Approved Special Consideration

   a. If a student links to the demographic data file has content area "Not Tested State Approved Special Consideration" indicated or does not link to the demographic data file and has content area "Special Consideration blank or partially blank reason" marked, then the student is identified as "Not Tested State Approved Special Consideration".

4. Not Tested State Approved Withdrew After

   a. If a student links to the demographic data file has content area not tested status of "Not Tested Withdrew After" and at least one content area session was not attempted or does not link to the demographic file has content area "Withdrew After blank or partially

blank reason" marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Withdrew After". For grade 11 writing, only use session 1 attempt status.

5. Not Tested State Approved Enrolled After

   a. If a student links to the demographic data file has content area not tested status of "Not Tested Enrolled After" and at least one content area session was not attempted or does not link to the demographic file has content area "Enrolled After blank or partially blank reason" marked and at least one content area session was not attempted, then the student is identified as "Not Tested State Approved Enrolled After". 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.

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 Enrolled After
5. Not Tested State Approved Withdrew After
6. Not Tested Other

E. Special Circumstances by content area

1. Item invalidation flags are provided to the DOE during data processing test clean up. The item invalidation flag variables are initially set using the rules below. The final values used for reporting are provided back to Measured Progress by the DOE and used in reporting..

   a. If reaaccomM2 is marked, then mark reaInvSes1, reaInvSes2, and reaInvSes3.
   b. If reaaccomM3 is marked, then mark reaInvSes1, reaInvSes2 and reaInvSes3.
   c. If mataccomM1 is marked then mark matInvSes1NC.
   d. If mataccomM3 is marked, then mark matInvSes1, matInvSes2, and matInvSes3.
   e. If wriaccomM3 is marked, then mark wriInvSes1 and wriInvSes2.
   f. If sciaccomM1 is marked, then mark sciInvSes3.
   g. If sciaccomM3 is marked, then mark sciInvSes1, sciInvSes2, and sciInvSes3.

2. A student is identified as content area tested if the student does not have any content area not tested reasons identified. Tested students are categorized in one of the four tested participation statuses: "Tested Damaged SRB", "Tested with Non-Standard Accommodations", "Tested Incomplete", and "Tested".

   a. Students with a common item response of 'X' are identified as "Tested Damaged SRB".
   b. Students identified as content area tested, are not identified as "Tested Damaged SRB", and have at least one of the content area invalidation session flags marked will be identified as "Tested with Non-Standard Accommodations". Grade 11 writing use only session 1 invalidation flag.
   c. Students identified as content area tested, are not identified as "Tested Damaged SRB", and not identified as "Tested with Non-Standard Accommodations" and did not attempt all sessions in the test are considered to be "Tested Incomplete."

d. All other tested students are identified as “Tested”.

3. For students identified as “Tested Damaged SRB”, the content area subcategories with at least one damaged item will not be reported. The school and district averages will be suppressed for the impacted subcategories on the student report. These students are excluded from all raw score aggregations (item, subcategory, and total raw score). They are included in participation, achievement level, and scaled score aggregations.

4. 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

5. 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).

6. 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.

7. Item scores for students with an invalidation flag marked and have a not tested status will be blanked out based on the invalidation flag. For example, if the student is identified as “Not Tested: State Approved Alternate Assessment” and has ReaInvSes1 marked, then all reading session 1 item responses will be reported as a blank.

F. Student Participation Status Hierarchy by content area

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 Enrolled After
5. Not Tested: State Approved Withdrew After
6. Not Tested: Other
7. Tested Damaged SRB
8. Tested with Non-Standard Accommodations
9. Tested Incomplete
10. Tested

G. Student Participation Summary

Participation Status	**Description**	**Raw Score (*)**	**Scaled Score (&)**	**Ach. Level**	**Student Report Ach. Level Text**	**Roster Ach. Level Text**
Z	Tested Damaged SRB(**)	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
A	Tested	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
B	Tested Incomplete(%)	✓	✓	✓	Substantially Below Proficient, Partially Proficient, Proficient, or Proficient with Distinction	1,2,3, or 4
C	Tested with Non-Standard	✓	✓	✓	Substantially Below Proficient, Partially	1,2,3, or 4

	Accommodations (%%)		Proficient, Proficient, or Proficient with Distinction	
D	Not Tested State Approved Alternate Assessment	✓	Alternate Assessment	A
E	Not Tested State Approved First Year LEP (Reading and Writing only)	✓	First Year LEP	L
F	Not Tested State Approved Enrolled After	✓	Fall: Enrolled After October 1 Spring: Enrolled After May 11	E
G	Not Tested State Approved Withdrew After	✓	Fall: Withdrew After October 1 Spring: Withdrew After May 11	W
H	Not Tested State Approved Special Consideration	✓	Special Consideration	S
I	Not Tested Other	✓	Not Tested	N

If a student has a participation status of Alternate Assessment for all subjects assessed at the grade level, a Parent Letter is not produced.

(*) Raw scores are not printed on student report for students with a not tested status.

(**) Raw scores for Tested damaged SRB students will be reported based on the set of non-damaged items. Subcategory scores will not be reported if it includes a damaged item.

(%) Tested incomplete students will be identified on the student report with a footnote.

(%%) Tested with Non-standard accommodations students will be identified on student report with a footnote. The invalidated items will be stored as a '-'for item analysis.

(&) Grade 11 writing students do not receive a scaled score. The writing achievement level is determined by the total common writing prompt score.

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. All mean raw scores are rounded to the nearest tenth.
4. Content Area Subcategories: Average Points Earned (student report): round to the nearest tenth.
5. 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, 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.

3. Students identified as Tested Damaged SRB are excluded from all raw score aggregations (item, subcategory, and total raw score). They are included in participation, achievement level, and scaled score aggregations.

C. Raw scores

1. For all analyses, non-response for an item by a tested student is treated as a score of 0. Items identified as damaged (response of 'X') will be excluded for student identified as "Tested Damaged SRB".

2. Content Area Total Points: Sum the points 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 all writing prompt scores, the final score of record is the sum of scorer 1 and scorer 2. If both scorers give the student a B, then the final score is B. If both scorers give the student an O or F, then the final score is 0.

E. Scaling

1. Scale Form creation

Scaling is accomplished by defining the unique set of test forms for the grade/subject. This is accomplished as follows:

a. Translate each form and position into the unique item number assigned to the form/position.

b. Order the items by

I. Type – multiple-choice, short-answer, constructed- response, extended-response, writing prompt.

II. Form – common, then by ascending form number.

III. Position

c. If an item number is on a form, then set the value for that item number to '1', otherwise set to '.'. Set the Exception field to '0' to indicate this is an original test form.

d. If an item number contains an 'X' (item is not included in scaling) then set the item number to '.'. Set the Exception field to '1' to indicate this is not an original test form.

e. Compress all of the item numbers together into one field in the order defined in step II to create the test for the student.

f. Select the distinct set of tests from the student data and order them by the exception field and the descending test field.

g. Check to see if the test has already been assigned a scale form by looking in the tblScaleForm table. If the test exists then assign the existing scale form. Otherwise assign the next available scale form number. All scale form numbering starts at 01 and increments by 1 up to 99.

2. Scaled Score assignment

a. Psychometrics provides data analysis with a lookup table for each scale form. The lookup table contains the raw score and the resulting scaled score.

F. SubTopic Item Scores

1. Identify the Subtopic

a. Fall:

I. The variable ContentFramework from the IABS export contains the data needed to calculate Content Strand, GLE code, subtopics, and subcategories.

i The Content Strand is stored as Standard. Except for Writing, Standard and RepCat are calculated using the third portion of ContentFramework and Reporting Category GLE Codes.doc provided by PM. For Grade 11 writing, Standard and RepCat are calculated using Writing Grade 11 by Form and Genre.doc provided by PM. For Grades 05 and 08 writing use Writing Content Strand Info.xls provided by PM.

ii The GLE Code is stored as TargetCode and is calculated by content area. For all content area remove leading zeros.

– Reading: Concatenate the second and third portions of ContentFramework separated by a dash.

– Math: Concatenate the second and fourth portions of ContentFramework separated by a dash.

– Writing: For Grades 05 and 08 concatenate the second and third portions of ContentFramework separated by a dash.

– Writing: For Grade 11 Writing PM provided a list of the appropriate target codes.

II. The variable Process Framework contains Depth of Knowledge code.

III. The variable type in IABS is the source for the Item Type, except the writing prompt item type is reported as "ER".

IV. PM provided Data Analysis with 2010NECAP_IABS_ReleasedItemsMAT.xls, 2010NECAP_IABS_ReleasedItemsREA.xls, and 2010NECAP_IABS_ReleasedItemsWRI which contain the released item orders for Math, Reading, and Writing respectively.

b. Spring: NECAP science item information is stored in IABS, except for inquiry items.

I. Program management provided Data Analysis with "UPDATED Copy of NECAP1112SInquiryIREF.xls" which contains the item order, domain, assessment target, DOK, item type, and maximum possible points for the inquiry items. Inquiry items are administered in session 3.

i Item numbers are created for inquiry items using the convention [2 Digit Grade][2 Digit Test Year][Inquiry Item Order] where 14 and 18 are the 2 digit grades for 4 and 8 respectively.

II. Program management provided Data Analysis with "IABS Export Codes for NECAP SCI Reporting.doc" which contains the crosswalk between IABS item information and reporting.

III. Program management provided Data Analysis with "2010 IABS_Released ItemsSCI for Tara.xls" which contains released item order. Inquiry items are listed at the end in the order they are in the test booklet.

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. For grade 11 writing, the subtopic score is the final score of record for the common writing prompt.

3. Content Area Subtopic (grade level results report): Subtopic scores are based on all unique common and matrix items.

   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 tenth. For the lower bound and upper bound round the Percent of Total Possible Points +/- Rounded Standard Error to the nearest hundredth.

   c. For grade11writing calculate the mean and standard deviation of each writing prompt score. These averages and standard deviations will be linked historically to the 1112 averages and standard deviations by writing type.

G. Grade 11 Writing prompt Score Distribution

1. Calculate number and percent of students at each score point by school, district, and state according to schtype and stustatus inclusion rules.

H. 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.

5. For NECAP 1213 Fall grades 05 and 08 writing cumulative total will be calculated using 0809, 1011, and 1213 data.

I. Participation

1. For participation calculate the number and percent of students in each of the following categories by school, district, and state according to schtype and stustatus decision rules.

2. Note that a student is tested with approved accommodations if one is tested, has a non-M accommodation marked, and does not have the M2 or M3 accommodation marked for that subject.

   a. For Students Enrolled, Students Tested, and Students Not Tested the denominator will be the number of students enrolled

   b. For Students Tested with approved Accommodations, Current LEP Students Tested (LEP=1), and IEP Students Tested the denominator will be the number of students tested.

   c. For Current LEP Students Tested with approved accommodations (LEP=1 the denominator will be the number of current LEP students tested.

d. For IEP Students Tested with approved accommodations the denominator will be the number of IEP students tested.

e. For Students Not Tested State Approved and Not Tested Other the denominator will be the number of students not tested.

f. For Students Not Tested Alternate Assessment, First Year LEP, Withdrew After October 1, Enrolled After October 1, and Special Considerations the denominator will be the number of students not tested state approved.

J. 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. Add and subtract one standard error of measurement to get the range and round to the nearest tenth.

2. Grade 11 writing Average Points Earned Students at Proficient Level will be reported as '7'.

3. For grades 05 and 08 writing prompt select all students across the states with the Y40 scaled score. Average the prompt scores across the students. Add and subtract one standard error to get the range and round to the nearest tenth.

K. 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. Students with the common writing prompt score of F or O will also receive annotations of FF and OO respectively.

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 single digit tested grade

c. For school and district name do the following.

I. For students with a stustatus value of 0 or 4, print the abbreviated tested school and district ICORE name based on school type decision rules.

II. Otherwise, for the school and district names print the "Description" in the StuStatus table presented earlier in this document.

d. Print "ME", "NH","RI", or "VT" for state.

2. Test Results by content area

a. Always display the cut scores in the graphic display.

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.

VI. 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.

3. 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 student column blank

   b. For percent of students with achievement level by school, district, and state print aggregate data based on student status, StuGrade, school type and minimum N rules.

4. 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. Print school, district, and state aggregate data for subcategories based on student status, StuGrade, school type and minimum N rules.

      II. For students identified as “Tested Damaged SRB” do not report student, school, and district aggregate data for subcategories that have at least one damaged item. Print Points Possible and state aggregate data.

      III. Otherwise, always print student subcategory scores

      IV. 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.

5. Writing Annotations

   a. Beginning in 1213, writing annotations will not print on the student reports.

6. Footer information

   a. Footnotes

      I. If the student received a participation status of “Tested with a non-standard accommodation” for any content area then print “**Student received no credit for parts of the test that were administered under non-standard conditions.”

      II. If the student received a participations status of “Tested Incomplete” for any content area then print “§This score should be viewed with caution because the student did not complete all parts of the test.”

      III. If both footnotes should appear, the print I. above II.

   b. For NH the SAU, district, and school codes should appear at the bottom right of the page separated by ‘-‘.

   c. For ME, RI, and VT district and school codes should appear at the bottom right of the page separated by ‘-‘.

B. Grade Level School/District/State Results

1. Reports are run by testing state, testing district, testing school using the aggregate school and district codes described in the school type table.

2. Fall Only: Reports are also run by teaching district, and teaching school using the aggregate school and district codes described in the school type table.
3. Exclude students based on stugrade=12, student status, school type and participation status decision rules for aggregations.
4. The reports will be collated as follows:
    a. Page 1 is the Title page.
    b. Page 2 is the Participation Results
    c. Page 3 is the Reading Historical and Subtopic Results
    d. Page 4 is the Reading Disaggregated Results
    e. Page 5 is the Math Historical and Subtopic Results
    f. Page 6 is the Math Disaggregated Results
    g. Page 7 is the Writing Historical and Subtopic Results (for Grades 05 and 08 only)
    h. Page 8 is the Writing Disaggregated Results (for Grades 05 and 08 only)
    i. Page 7 is the Writing Historical Results (for Grade 11only)
    j. Page 8 is the Writing Subtopic Results Page (for Grade 11only)
    k. Page 9 is the Writing Score Distribution Results (for Grade 11 only)
    l. Page 10 is the Writing Disaggregated Results (for Grade 11 only)
5. Report Header Information
    a. "Fall YYYY Beginning of Grade XX NECAP Tests" where XX is the single digit grade level and YYYY is the year, will print as the title.
    b. Teaching level reports will have the following subtitle: "Grade XX-1 Students in (YYYY-1)-(YYYY)".
    c. Testing level reports will have the following subtitle: "Grade XX Students in (YYYY)-(YYYY+1)".
    d. Use abbreviated school and district name from ICORE based on school type decision rules.
    e. Print "Maine", "New Hampshire", "Rhode Island", or "Vermont" to reference the state. The state graphic is printed on the first page.
    f. For NH print SAU, district, and school codes separated by '-' for Code on first page for school level. Print SAU and district codes separated by '-'for the district level. Print the full state name for the state level.
    g. For ME, RI, and VT print district and school codes separated by '-' for Code on first page for the school level. Print the district code for the district level. Print the full state name for the state level.
6. For achievement level and participation category data if the number of students in an achievement level or participation category does not equal 0, and the percent of students is 0 then format the percent as <1.
7. 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.
8. 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.

9. Report Section: Historical NECAP Results by content area

a. For tested 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 scores based on minimum N-size and school type decision rules.

c. Bold current year data.

10. 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

11. Report Section: Historical Score Comparison(Grade 11 Writing Only)

a. For testing and teaching level reports always print based on minimum N-size and school type decision rules

b. The subtopic associated with the current year common prompt will appear first. The other subtopics will appear in the following GE sort order.

I. Response to Literary Text

II. Response to Informational Text

III. Reflective Essay

IV. Report

V. Persuasive Essay

VI. Proceedure

12. Report Section: Common Prompt Score Distribution(Grade 11 Writing Only)

b. For testing and teaching level reports always print based on minimum N-size and school type decision rules

13. 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 scores based on minimum N-size and school type decision rules.

C. School/District/State Summary(School Level is run in the Fall Only)

1. Report Header Information

a. Use abbreviated school and district name from ICORE based on school type decision rules.

b. Print “Maine”, “New Hampshire”, “Rhode Island”, or “Vermont” to reference the state.

c. For NH print SAU, district, and school codes separated by ‘-‘ for Code on first page for school level. Print SAU and district codes separated by ‘-‘for the district level. Print the full state name for the state level.

d. For ME, RI, and VT print district and school codes separated by '-' for Code on first page for the school level. Print the district code for the district level. Print the full state name for the state level.

2. Reports are run by testing state, testing district, testing school (Fall Only) using the aggregate school and district codes described in the school type table

3. Fall Only: Reports are also run by teaching district, and teaching school using the aggregate school and district codes described in the school type table.

4. Exclude students based on StuGrade=12, student status, school type and participation status decision rules for aggregations.

5. For achievement level and participation category data if the number of students in an achievement level or participation category does not equal 0, and the percent of students is 0 then format the percent as <1.

6. 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 scores across the grades is not calculated.

7. 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 scores across the grades is not calculated.

8. Printed Grade Column

   a. For the all grades row, display the school, district, or state name.

   b. For grades 3-8 and 11 rows print Beginning of Grade X.

D. Writing Prompt CD

1. The bookletnumber associated with the writing prompt score will be stored in tblStuDemo.

V. Data Requirements Interactive Reporting

A. Student Level

1. Refer to Sections II and III. D for decision rules on how student test data will be stored.

2. Students will be loaded into the Interactive System based off of the Interactive flag in tblStuDemo. Students with Interactive flag set to 0 will not be loaded into the system. Students with Interactive set to 1 will be loaded.

   a. Students with StuStatus value of 1, 2 or 3 or RI StuGrade=12 will have the Interactive flag set to 0.

   b. All others will have Interactive=1.

3. The Included flag will determine which students are included in school level aggregations. Students with Included=0 are excluded from all aggregations. Students with Included=2 will be included in Performance Level aggregations and excluded from raw score aggregations (item, subcategory, and total raw score). Students with Included=1 will be included in all school level aggregations.

   a. Students with a Not Tested Participation Status, StuStatus=1, 2, or 3, or RI StuGrade=12 will have their Included flag set to 0.

   b. Students who do fall into the above group and have Participation Status of Tested Damaged SRB will have their Included flag set to 2.

   c. All other students will have their Included flag set to 1.

4. Longitudinal Data

a. Only students with a valid StudentID and Interactive flag=1 will be loaded.

b. The complete achievement level name or not tested reason will be stored .

B. Aggregate Level

1. Data Analysis will compute Item Averages for the whole group only at the testing and teaching (Fall only) School and District Levels.

2. Data Analysis will compute Item Averages for all of the filter combinations that exist at the State Level.

3. Data Analysis will create a lookup table with all of the possible filter combinations. It will contain the variable Filter with length 5. Each position represents one of the filter variables. It will contain all the possible combinations of the values plus nulls for when variables are not selected. The first position will be Gender, second Ethnic, third IEP, fourth LEP, and fifth EconDis.

4. Data Analysis will compute Item Averages, Achievement Level Summary, and Item Summary data for the filter combinations for a sample of schools for quality assurance reveiw.

a. For this sample, percents will be rounded to the nearest whole number and open response average scores will be rounded to the nearest tenth.

b. For the Item Summary data, item responses other than A, B, C, and D will be counted in the IR column.

C. Released Item Data

1. In 1213 there will be 22 items included in the released item document posted on the supporting documents tab for grade 11 reading. Only 17 of these will have data released on the on the interactive system and in the student level csv files posted on the web. The last 5 items will not have data included in the system or files.

VI. Data File Rules

In the file names GR refers to the two digit grade (03-08, 11), YYYY refers to the year, DDDDD refers to the district code, and SS refers to two letter state code. Refer to the tables at the end of this section for filenames and layouts. Teaching level data files will be produced in the Fall Only.

A. 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

B. State Student Released Item Data

1. A CSV file will be created for each state for grades 3-8 and one for grade 11.

2. One CSV file will be created for each state in the Spring.

3. Accommodation Flags

a. If the student has at least 1 standard accommodation marked (excluding M) for a given subject then set [sub]STDaccom flag to ‘1’. Otherwise set it to ‘0’.

b. For each group of accommodations (S, T, P, R, and O) if a student has any accommodation in that group marked set [sub]Accom[group]=’1’. Otherwise set it to ‘0’.

c. If a student has the M2 accommodation marked, then set [sub]AccomM2=’1’. Otherwise set it to ‘0’.

d. If a student has the M3 accommodation marked, then set [sub]AccomM3=’1’. Otherwise set it to ‘0’.

4. Exclusion Rules

   a. NH: If the student has a testing school type of ‘PRI’ or StuStatus is 1,2 ,3,or 4 then exclude the student

   b. RI: If testing school type is PRI and teaching school type is PRI or blank, then exclude the student.

   c. VT: Do not exclude any students

   d. ME: If the student has a StuStatus is 1,2 ,3,or 4 then exclude the student

C. State Student Raw Data

1. A CSV file will be created for each state by grade span. The grade spans are 3-4, 5-8, and 11. In the spring, all grades will be combined.

2. If the student has at least 1 standard accommodation marked (excluding M) for a given subject then set [sub]STDaccom flag to ‘1’. Otherwise set it to ‘0’.

3. Exclusion Rules

   a. NH: If the student has a testing school type of ‘PRI’ or StuStatus is 1,2, 3,or 4 then exclude the student

   b. RI: If testing school type is PRI and teaching school type is PRI or blank, then exclude the student.

   c. VT: Do not exclude any students

   d. ME: If the student has a StuStatus is 1,2 ,3,or 4 then exclude the student.

D. State Student Scored Data

1. A CSV file will be created for each state including all grades.

2. Exclusion Rules

   a. NH: If the student has a testing school type of ‘PRI’ or StuStatus is 1,2 3,or 4 then exclude the student

   b. RI: If testing school type is PRI and teaching school type is PRI or blank, then exclude the student.

   c. VT: Do not exclude any students

   d. ME: If the student has a StuStatus is 1,2 ,3,or 4 then exclude the student.

E. District Student Data

1. Testing and teaching CSV files will be created for each state and grade and district.

2. Students with the Discode or SendDiscode will be in the district grade specific CSV file for the testing year.

3. Fall Only: Students with a sprDiscode or sprSendDiscode will be in the district grade specific CSV file for the teaching year.

4. For ME, NH, and RI only public school districts will receive district data files. (Districts with at least one school with schoolsubtypeID=1, 11, 19, or 15 in ICORE)

5. Accommodation Flags

   a. If the student has at least 1 standard accommodation marked (excluding M) for a given subject then set [sub]STDaccom flag to ‘1’. Otherwise set it to ‘0’.

   b. For each group of accommodations (S, T, P, R, and O) if a student has any accommodation in that group marked set [sub]Accom[group]=’1’. Otherwise set it to ‘0’.

c. If a student has the M2 accommodation marked, then set [sub]AccomM2='1'. Otherwise set it to '0'.

d. If a student has the M3 accommodation marked, then set [sub]AccomM3='1'. Otherwise set it to '0'.

6. Exclusion Rules

a. NH & RI: If the student has a StuStatus value of 1,2, or 3 then exclude the student

b. VT: If the student has a StuStatus value of 1, then exclude the student.

c. ME: If the student has a StuStatus is 1, 2, or 3 then exclude the student.

F. School Student Data

1. Testing and teaching CSV files will be created for each state and grade and school.

2. Students with the SchCode will be in the school grade specific CSV file for the testing year.

3. Fall Only: Students with the sprSchcode will be in the school grade specific CSV file for the teaching year.

4. Accommodation Flags

a. If the student has at least 1 standard accommodation marked (excluding M) for a given subject then set [sub]STDaccom flag to '1'. Otherwise set it to '0'.

b. For each group of accommodations (S, T, P, R, and O) if a student has any accommodation in that group marked set [sub]Accom[group]='1'. Otherwise set it to '0'.

c. If a student has the M2 accommodation marked, then set [sub]AccomM2='1'. Otherwise set it to '0'.

d. If a student has the M3 accommodation marked, then set [sub]AccomM3='1'. Otherwise set it to '0'.

5. Exclusion Rules

a. NH & RI: If the student has a StuStatus value of 1,2 or 3, then exclude the student

b. VT: If the student has a StuStatus value of 1, then exclude the student.

c. ME: If the student has a StuStatus is 1, 2, or 3 then exclude the student.

G. Common Item Information

1. An excel file will be created containing item information for common items: grade, subject, released item number, item analysis heading data, raw data item name, item type, key, and point value.

H. State Standard Deviations and Averages Scaled Scores

1. A csv file will be created for each state containing the standard deviations and average scale scores for disaggregated subgroups by subject.

2. Exclude students based on state aggregation StuGrade, StuStatus, and SchType decision rules.

3. Data will be suppressed based on minimum N-size and report type decision rules.

4. Average scaled score will be rounded to the nearest whole number. Standard deviations will be rounded to the nearest tenth.

I. Grade Level Results Report Disaggregated and Historical Data

1. Teaching and testing CSV files will be created for each state containing the grade level results disaggregated and historical data.

2. Data will be suppressed based on minimum N-size and report type decision rules.

3. Private schools are excluded from NH & RI files.

J. Grade Level Results Report Participation Category Data

1. Testing CSV file will be created for each state containing the grade level results participation data.
2. Private schools are excluded from NH & RI files.

K. Grade Level Results Report Subtopic Data

1. Teaching and testing CSV files will be created for each state containing the grade level results subtopic.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. Private schools are excluded from NH & RI files.

L. Summary Results Data

1. Teaching and testing CSV files will be created for each state containing the school, district and state summary data.
2. Data will be suppressed based on minimum N-size and report type decision rules.
3. Private schools are excluded from NH & RI files.

M. Released Item Percent Responses Data

1. The CSV files will only contain state level aggregation for released items.
2. CSV files will be created for each state and grade containing the released item analysis report state data.

N. Invalidated Students Original Score

1. A CSV file will be created for each state including all grades.
2. Original raw scores for students whose responses were invalidated for reporting will be provided.
3. Exclusion Rules
    a. NH: If the student has a testing school type of 'PRI' or StuStatus is 1,2, 3, or 4 then exclude the student
    b. RI: If testing school type is PRI and teaching school type is PRI or blank, then exclude the student.
    c. VT: Do not exclude any students
    d. ME: If the student has a StuStatus is 1,2 ,3,or 4 then exclude the student.

O. Student Questionnaire Summary

1. One CSV file will be created for each state containing percent of students at each response, percent of students at each achievement level, and average scaled score, by student questionnaire response.
2. Only include students who are included in state level aggregations.
3. Data will be suppressed based on minimum N-size and report type decision rules.

P. TCTA Questionnaire Raw Data

1. One CSV file will be created for each state containing raw TC Questionnaire data.
2. One CSV file will be created for each state containing raw TA Questionnaire data.

Q. TCTA Questionnaire Frequency Distribution

1. One CSV file will be created for each state containing the distribution of responses of TC Questionnaire raw data.
2. One CSV file will be created for each state containing the distribution of responses of TA Questionnaire raw data.

R. Scaled Score Lookup

1. One CSV file and one excel file will be created containing the scaled score lookup data.

S. Grade 11Writing Score Distribution Results

1. Testing CSV file will be created for each state containing the grade level results participation data.
2. Private schools are excluded from NH & RI files.

T. Grade 11Writing Historical Subtopic Results

1. Testing CSV file will be created for each state containing the grade level results participation data.
2. Private schools are excluded from NH & RI files.

U. Subtopic Average Points Earned (For Program Management)

1. One excel file will be created containing four worksheets. The first worksheet contains the total possible points for each subtopic as reported on the item analysis report and the range for students who are just proficient. The remaining three worksheets contain state average subtopic scores as reported on the item analysis report.
2. Program management uses this file to create a document which is provided to the schools.

V. Item Stats for Inquiry Task Items (For Program Management)

1. Since Inquiry Task Items are not stored in IABS, one CSV file will be created containing item stats for Inquiry Task items.
2. All three states are included in the calculations.

W. Memo Shipping Files (For Program Management)

1. Provide PM in excel list of schools and districts that tested regardless of grade.

X. CD Print File

Y. Fall Table Data File Deliverables

Data File	**Layout**	**File Name**
Preliminary State Results	N/A	Included in Equating Report
State Student Released Item Data	NECAP1213FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1213FallStateStudentReleasedItem[GS].csv GS=0308 or 11
State Student Raw Data	NECAP1213FallStateStudentRawLayout.xls (one worksheet for each of the 4 unique test designs)	NECAP1213FallStateStudentRaw[GS].csv Gs=0304, 0508, or11
State Student Scored Data	NECAP1213FallStateStudentScoredLayout.xls	NECAP1213FallStateStudentScored.csv
District Student Data	NECAP1213FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1213FallTestingDistrictSlice[GR]_[District Code].csv NECAP1213FallTeachingDistrictSlice[GR]_[District Code].csv
School Student Data	NECAP1213FallStudentReleasedItemLayout.xls(one worksheet for grade 11 and one worksheet for 03-08)	NECAP1213FallTestingSchoolSlice[GR]_[District Code][School Code].csv

		NECAP1213FallTeachingSchoolSlice[GR]_[District Code][School Code]..csv
Common Item Information	NECAP1213FallCommonItemInformationLayout.xls	NECAP1213FallCommonItemInformation.xls
Grade Level Results Report Disaggregated and Historical Data	NECAP1213FallResultsReport DisaggregatedandHistoricaLayout.xls	NECAP1213FallResultsReportTestingDisaggregatedandHistoricalcsv NECAP1213FallResultsReportTeachingDisaggregatedandHistorical.csv
State Standard Deviations and Average Scaled Scores	NECAP1213FallStateStandardDeviationsLayout.xls	NECAP1213FallStateStandardDeviations.csv
Grade Level Results Report Participation Category Data	NECAP1213FallResultsReportParticipationLayout.xls	NECAP1213FallResultsReportTestingParticipation.csv
Grade Level Results Report Subtopic Data	NECAP1213FallResultsReport SubtopicLayout.xls	NECAP1213FallResultsReportTestingSubtopic.csv NECAP1213FallResultsReportTeachingSubtopic.csv
Summary Results Data	NECAP1213FallSummaryResultsLayout.xls	NECAP1213FallSummaryResultsTesting.csv NECAP1213FallSummaryResultsTeaching.csv
Released Item Percent Responses Data	NECAP1213FallReleasedItemPercentResponsesLayout.xls	NECAP1213FallReleasedItemPercentResponses.csv
Invalidated Students Original Score	NECAP1213FallStateInvalidatedStudent OriginalScoredLayout.xls	NECAP1213FallStateInvalidatedStudent OriginalScored.csv
Student Questionnaire Summary	NECAP1213FallStudentQuestionnaireSummaryLayout.xls	NECAP1213FallStudentQuestionnaireSummary.csv
TCTA Questionnaire Raw Data	NECAP1213FallTCQuestionnaireRawLayout.xls NECAP1213FallTAQuestionnaireRawLayout.xls	NECAP1213FallTCQuestionnaireRaw.csv NECAP1213FallTAQuestionnaireRaw.csv
TCTA Questionnaire Frequency Distribution	NECAP1213FallTCTAQuestionnaireFreqLayout.xls	NECAP1213FallTCTAQuestionnaireFreq.csv
Scaled Score Lookup	NECAP1213FallScaleScoreLookupLayout.xls	NECAP1213FallScaleScoreLookup.xls NECAP1213FallScaleScoreLookup.csv
Grade 11Writing Score Distribution Results	NECAP1213FallResultsReportGrade11WriScoreDistributionLayout.xls	NECAP1213FallResultsReportGrade11WriScoreDistributionTesting.xls NECAP1213FallResultsReportGrade11WriScoreDistributionTeaching.xls
Grade 11Writing	NECAP1213FallResultsReportGrade11WriHisoricalSubtopicLayout.xls	NECAP1213FallResultsReportGrade11WriHisoricalSubtopicTesting.xls

Historical Subtopic Results		NECAP1213FallResultsReportGrade11WriHisoricalSubtopicTeaching.xls
Subtopic Average Points Earned (For Project Management)	N/A	NECAP1213FallSubtopicAvgPointsEarned.xls
Memo Shipping Files (For Program Management)	N/A	TBD
CD Print File	N/A	TBD

A. Spring Table Data File Deliverables

Data File	**Layout**	**File Name**
Preliminary State Results	N/A	Included in Equating Report
State Student Released Item Data	NECAP1213SpringStudentReleasedItemLayout.xls	NECAP1213SpringStateStudentReleasedItem.csv
State Student Raw Data	NECAP1213SpringStateStudentRawLayout.xls	NECAP1213SpringStateStudentRaw.csv
State Student Scored Data	NECAP1213SpringStateStudentScoredLayout.xls	NECAP1213SpringStateStudentScored.csv
District Student Data	NECAP1213SpringStudentReleasedItemLayout.xls	NECAP1213SpringDistrictSlice[GR]_[District Code].csv
School Student Data	NECAP1213SpingStudentReleasedItemLayout.xls	NECAP1213SpringSchoolSlice[GR]_[District Code][School Code].csv
Common Item Information	NECAP1213SpringCommonItemInformationLayout.xls	NECAP1213SpringCommonItemInformation.csv
State Standard Deviations and Average Scaled Scores	NECAP1213SpingStateStandardDeviationsLayout.xls	NECAP1213SpringStateStandardDeviations.csv
Grade Level Results Report Disaggregated and Historical Data	NECAP1213SpringResultsReport DisaggregatedandHistoricaLayout.xls	NECAP1213SpringResultsReportDisaggregatedandHistorical.csv
Grade Level Results Report Participation Category Data	NECAP1213SpringResultsReport ParticipationLayout.xls	NECAP1213SpringResultsReportParticipation.csv
Grade Level Results Report Subtopic Data	NECAP1213SpringResultsReport SubtopicLayout.xls	NECAP1213SpringResultsReportSubtopic.csv
Summary Results Data	NECAP1213SpringSummaryResultsLayout.xls	NECAP1213SpringSummaryResults.csv

Released Item Percent Responses Data	NECAP1213SpringReleasedItemPercentResponsesLayout.xls	NECAP1213SpringReleasedItemPercentResponses.csv
Invalidated Students Original Score	NECAP1213SpringStateInvalidatedStudent OriginalScoredLayout.xls	NECAP1213SpringStateInvalidatedStudent OriginalScored.csv
Student Questionnaire Summary	NECAP1213SpringStudentQuestionnaireSummaryLayout.xls	NECAP1213SpringStudentQuestionnaireSummary.csv
TCTA Questionnaire Raw Data	NECAP1213SpringTCQuestionnaireRawLayout.xls NECAP1213SpringTAQuestionnaireRawLayout.xls	NECAP1213SpringTCQuestionnaireRaw.csv NECAP1213SpringTAQuestionnaireRaw.csv
TCTA Questionnaire Frequency Distribution	NECAP1213SpringTCTAQuestionnaireFreqLayout.xls	NECAP1213SpringTCTAQuestionnaireFreq.csv
Scaled Score Lookup	NECAP1213SpringScaleScoreLookupLayout.xls	NECAP1213SpringScaleScoreLookup.xls NECAP1213SpringScaleScoreLookup.csv
Subtopic Average Points Earned (For Project Management)	N/A	NECAP1213SpringSubtopicAvgPointsEarned.xls
Item Stats for Inquiry Task Items (For Program Management)	N/A	NECAP1213SpringInquiryItemStats.csv
Memo Shipping Files (For Program Management)	N/A	TBD